

interazione uomo:macchina _interfaccecreative



un breve saggio introduttivo a cura di
marco.infussi



la traduzione, memorizzazione elettronica, la riproduzione e l'adattamento totale o parziale, con qualsiasi mezzo, compresi i microfilm e le copie fotostatiche è concessa senza limiti

tra queste note è assurdo trovare, in libri coperti da diritti d'autore, frasi del tipo: l'editore, per quanto di propria spettanza, considera rare le opere fuori del proprio catalogo editoriale. la riproduzione a mezzo di fotocopia degli esemplari di tali opere esistenti nelle biblioteche è consentita, non essendo concorrenziale all'opera. non possono considerarsi rare le opere di cui esiste, nel catalogo dell'editore, una successiva edizione, o le opere antologiche.

sono per la proprietà intellettuale ma non per gli ostacoli alla divulgazione dello scibile

redazione e impaginazione: marco.infussi per w.graphique™

prima edizione: © giugno 2006

interazione uomo:macchina _interfaccecreative



un breve saggio introduttivo a cura di
marco.infussi



dallo stesso autore:

[q] - catalogo generale delle opere; marco.infussi, e-book 500 pagine, 2007.
[laser mbe hybrid-fet] - tesi di laurea; marco.infussi, e-book 300 pagine, 2005.
[.:frakt_lectures] - percorso sui rapporti tra arte, matematica e natura; marco.infussi, e-book 365 pagine, 2004.
[amo le nuvolette_] - contenitore di stati d'animo; marco.infussi, e-book 75 pagine, 2003.
[...parlo ...ascolto] - raccolta di poesie in grafica; marco.infussi, e-book 33 pagine, 2002.
[tempo&spazio] - percorso interdisciplinare; marco.infussi, e-book 100 pagine, 2001.

realizzare un libro è un'operazione complessa, che richiede numerosi controlli: sul testo, sulle immagini e sulle relazioni che si stabiliscono fra loro. l'esperienza suggerisce che è praticamente impossibile pubblicare un libro privo di errori. sarò quindi grato ai lettori che vorranno segnalarmi. per segnalazioni o suggerimenti relativi a questo libro rivolgersi a:

w.graphique@alice.it

[supporta il lavoro di 17%*skillz_pro con donazioni di qualsiasi genere: wet.graphics@tin.it]
[&] [&] [&] [&] [&] [&]

Introduzione

Non avrei di certo immaginato che la scienza fosse questo.

Il giorno in cui, inconsapevolmente o forse incoscientemente, scelsi di iscrivermi alla facoltà d'ingegneria, non avevo la minima idea di quel che avrebbe potuto esser poi alla mia portata, dopo cinque anni di studio. "There's plenty of room" per davvero, l'entusiasmo di Richard Feynman nel proferire queste parole agli albori delle micro e nanotecnologie diventa contagioso.

Ho deciso di riproporre la stessa formula del saggio sulla geometria frattale e i rapporti tra l'arte, la matematica e la natura, anche perché una parte degli argomenti trattati si inseriscono più o meno sulla stessa linea di ricerca. Il procedimento è il solito, comunissimo, data-mining umano su internet. Lo facciamo tutti, anche se poi lasciamo questo sapere resta nei dischi rigidi. A me piace invece l'idea di mettere a frutto i mesi passati a selezionare articoli e referenze, a distillare gli argomenti studiati, raccogliendo il tutto in un contenitore che sia il più possibile organico ed ordinato, affinché chiunque intenda avvicinarsi alla materia non debba far altro che leggere, usando il mio tempo e risparmiando il suo, saltando il processo di ricerca bibliografica ed introducendosi così subito alla ricerca scientifica.

Una grande spinta ad approfondire gli argomenti è nata durante i corsi di sensori e nanosistemi tenuti nella mia università dal prof. Arnaldo D'Amico, un "nonno" ideale per chi ama il sapere, anche perché una parte di questo saggio è l'integrazione di ricerche svolte proprio per i suddetti esami. Aggiornato, dinamico, scherzoso, ha la pazienza di dedicare del tempo ad ascoltare gli studenti, a supportare le loro idee. Differentemente dalla maggior parte di coloro che hanno voluto la riforma dell'ordinamento accademico, è convinto della fondamentale importanza dell'interdisciplinarietà, di una solida base culturale, di collegamenti inusitati e proprio per questo innovativi, della creatività e della curiosità nel campo scientifico: tutti concetti che condivido appieno e che pongo alla base del mio operare. E' per questo che ho voluto citarlo, perché di docenti con questa apertura mentale ce ne sono davvero pochi. Proprio per questa sua natura, diciamo volo-pindarica, il contenitore si presenta come raccolta di articoli, concetti, idee, senza troppe pretese nell'essere esauriente e rigoroso. La scelta di proporre gli argomenti in italiano/inglese misto può esser vista come esercizio per il lettore (se mai ce ne sarà uno!) novizio al mondo scientifico, anche se l'intenzione vera è quella di tradurre tutto il testo in inglese in una futura rielaborazione.

Rivolgo allora questo piccolo saggio a tutti gli studenti delle scuole superiori, che sia utile a trovare un orientamento per proseguire gli studi, e ovviamente agli studenti universitari, per i quali ogni tanto si scende nei dettagli concreti, e ai curiosi, che troveranno materiale a volontà per alimentare quel senso di stupore che, almeno qui in Italia, è l'unica cosa che fa muovere la comunità scientifica.


marco.infussi

capitolo 0: interazione uomo:macchina

- 0.0. interazione uomo:macchina
- 0.1. azione umana ed uso degli strumenti
- 0.2. cognizione distribuita
- 0.3. la progettazione di interfacce uomo-computer
- 0.4. analisi dell'attività umana e del contesto d'uso
- 0.5. sviluppo di prototipi
- 0.6. valutazione
- 0.7. ergonomia, domotica e wearable sensors
- 0.8. le nuove tendenze di ricerca
- 0.9. riferimenti bibliografici

capitolo 1: chi ^x

- 1.0. chi: computer:human interaction
- 1.1. una sfida per il designer
- 1.2. interattività
- 1.3. evoluzione dei paradigmi di interazione uomo-computer
- 1.4. progettare l'interazione
- 1.5. conclusioni
- 1.6. referenze bibliografiche

capitolo 2: usabilità

- 2.0. usabilità
- 2.1. l'occhio
- 2.2. l'errore
- 2.3. memoria a breve termine
- 2.4. gestalt, o psicologia della visione
- 2.5. religioni informatiche
- 2.6. the loebner prize
- 2.7. riconoscimento della scrittura
- 2.8. gesture recognition
- 2.9. tecnologia assistiva

capitolo 3: arte&scienza: interfacce creative

- 3.0. valore euristico dell'arte
- 3.1. letteratura del novecento
- 3.2. ars electronica
- 3.3. siggraph

capitolo 4: tecnologie d'interfaccia

- 4.0 attuali tecnologie d'interfaccia
- 4.1 video touchscreen
- 4.2 tastiera, mouse, touchpad, graphics tablet, light pen per computer
- 4.3 gamepad, joystick, light gun, analog stick, power glove per applicazioni videoludiche
- 4.4 musical instrument digital interface e computer music
- 4.5 aptica: informazione trasmessa da stimolazioni tattili

capitolo 5: mems

- 5.0. micro electro mechanical systems
- 5.1. accelerometro
- 5.2. visione artificiale
- 5.3. motion capture

capitolo 6: gesture recognition

- 6.0. gesture recognition
- 6.1. nature of gesture
- 6.2. representations of gesture
- 6.3. pen-based gesture recognition
- 6.4. tracker-based gesture recognition
- 6.5. suggestions for system design
- 6.6. summary and future directions
- 6.7. bibliographic references

capitolo 7: case studies

- 7.0. smart laser scanner for hci
- 7.1. physical motion interface
- 7.2. sensorizzazione di una protesi
- 7.3. apple multitouch screen
- 7.4. nintendo wii remote

appendice

- a.0. alcuni concetti aggiuntivi
- a.1. l'intelligenza artificiale: la storia e le idee
- a.2. rassegna di articoli

0

capitolo



0.0. Interazione Uomo-Macchina

L'Interazione Uomo-Macchina può essere definita in molti modi. Hewett, et al. (1992) la descrivono come:

"...la disciplina che si occupa della progettazione, valutazione e implementazione di sistemi interattivi e dello studio di tutti i fenomeni più importanti connessi con l'interazione tra questo tipo di sistemi e l'essere umano".

In un noto libro Preece et al. (1994) propongono una definizione che pone l'accento su altri aspetti:

"L'Interazione Uomo-Macchina si occupa della progettazione di sistemi che siano in grado di sostenere le persone che li usano a svolgere la loro attività in modo produttivo e sicuro".

Qualunque sia la definizione offerta dell'interazione uomo-macchina l'elemento saliente risiede nell'applicazione sistematica sia di conoscenze sulle caratteristiche cognitive degli esseri umani che sulle capacità delle macchine, per consentirci di fare cose che non avremmo potuto fare senza il supporto di strumenti. Presenteremo l'Interazione Uomo-Macchina come la disciplina che studia l'interazione tra il sistema cognitivo umano e gli strumenti per l'elaborazione di informazione e illustreremo come la conoscenza prodotta da questo studio sia utilizzata per progettare strumenti appropriati per i più svariati usi, dal lavoro, all'educazione, al divertimento. In particolare cercheremo di rispondere ad alcune domande. Quali sono le proprietà del sistema cognitivo umano che presiedono al controllo delle azioni ed all'uso di strumenti? Che tipo di interazione si instaura tra il sistema umano di elaborazione dell'informazione e gli strumenti, e che effetti ha sulla mente umana l'uso di sistemi artificiali di elaborazione dell'informazione? Quali sono i requisiti per la progettazione di sistemi artificiali di elaborazione dell'informazione con cui l'uomo può interagire efficacemente?

0.1. L'azione umana e l'uso degli strumenti

La psicologia cognitiva ha storicamente trascurato l'azione umana. Ha considerato la percezione e l'azione come due separati domini di studio ed ha privilegiato il primo. Tale distinzione è però ingannevole (vedi ad es. Neisser, 1976; Arbib, 1989), poichè è stato mostrato che la percezione è coinvolta nella selezione delle azioni, nella loro esecuzione e nella continua valutazione dei risultati (Norman e Shallice, 1980). Infatti un'azione non può, sia da un punto di vista fisiologico che psicologico, considerarsi completata senza una fase di valutazione (Arbib, 1989). Inoltre, l'attenzione posta sul ruolo degli strumenti nel mediare l'azione umana è stata ancora più scarsa, fatta eccezione per un'importante tradizione psicologica, quella della scuola storico-culturale (Vygotsky, 1987). In tale contesto, uno dei pochi modelli dell'azione umana che ha considerato anche il ruolo degli strumenti, almeno in fase d'uso, è quello proposto da Norman (1988; Hutchins, Hollan e Norman, 1985); peraltro questo modello ripropone alcuni dei concetti fondamentali dell'approccio culturale nell'ambito dell'interazione uomo-computer. Il modello deriva dalla teoria del controllo delle azioni di Norman e Shallice (1980) e tenta di dare conto sia delle fasi che delle possibili difficoltà di passaggio da una fase all'altra (distanze) osservate nell'uso di strumenti.

Le fasi del modello Hutchins, Hollan e Norman sono sette, una per gli obiettivi, tre per l'esecuzione e tre per la valutazione:

- Formazione dell'obiettivo
- Formazione dell'intenzione
- Specificazione dell'azione
- Esecuzione dell'azione
- Percezione dello stato del mondo
- Interpretazione dello stato del mondo
- Valutazione del risultato

Le distanze sono invece tre, due delle quali sono presenti sia sul lato dell'esecuzione che sul lato della valutazione: la distanza semantica e la distanza referenziale. La terza distanza si riferisce al passaggio dall'esecuzione alla valutazione dell'azione. Sia le fasi che le distanze non devono essere considerate come entità discrete dai contorni ben separati ma piuttosto come momenti più o meno presenti in tutte le attività. Il modello ha

lo scopo principale di definire le potenziali unità di analisi attraverso le quali è possibile rappresentare il comportamento umano nell'uso di strumenti e non ha né la pretesa di essere esaustivo né quella di essere corretto. Ha però il merito di essere una guida efficace nell'analisi delle difficoltà d'uso di strumenti, e di fornire un framework entro il quale collocare molte delle conoscenze della psicologia cognitiva (Bagnara e Rizzo, 1989). Per esemplificare i concetti descritti sopra proviamo ad analizzare una attività utilizzando il modello di Norman. Immaginiamo di lavorare al computer, di usare un programma di scrittura e di aver più files aperti. Immaginiamo inoltre di usare un'interfaccia grafica a manipolazione diretta (ad es. le interfacce che fanno uso di icone, finestre, sistemi di puntamento, etc.). Sto lavorando sul file A ma decido che ho bisogno di alcune informazioni che sono sul file B (questo è il mio obiettivo: avere una parte delle informazioni presenti nel file B attivo in una prima schermata sul computer). Il mio obiettivo deve essere tradotto nelle intenzioni che definiscono le azioni appropriate nel mondo: prendere il mouse, muovere il puntatore sopra il comando *Finestra* nella barra di menu, cliccare, selezionare il nome del file desiderato, rilasciare il bottone del mouse, osservare in che punto è evidenziato il testo del file B nella finestra, etc.). Ma c'è dell'altro. Devo anche specificare come muovere la mia mano, come impugnare propriamente il mouse e come spingere il suo bottone controllando il puntatore sullo schermo. L'obiettivo deve quindi essere tradotto in specifiche intenzioni, che a loro volta devono essere tradotte in specifiche sequenze di azioni, azioni che controllano i miei apparati percettivo-motori. E' da notare inoltre, che posso raggiungere il mio scopo attraverso altre sequenze di azioni e altre intenzioni. Se per esempio muovendo il puntatore verso la barra di menu mi accorgo che sul retro della finestra attiva è parzialmente evidente la finestra relativa al file B, posso portare il puntatore su quell'area dello schermo e cliccare in modo da rendere tale finestra attiva in primo piano. L'obiettivo non è cambiato ma parte della sequenza di azioni e delle relative intenzioni lo sono. L'attività appena descritta sarebbe profondamente diversa se io stessi lavorando non con un'interfaccia grafica a manipolazione diretta bensì con un sistema che usa una interfaccia a linea di comando (es. VI per UNIX o WordStar per DOS 3.3). L'obiettivo potrebbe rimanere immutato ma le intenzioni, le azioni, l'interpretazione degli eventi e la valutazione dei risultati sarebbero radicalmente diverse. Le distanze tra le varie fasi del modello dell'azione si riferiscono proprio al modo con cui si realizza più o meno prontamente il passaggio da una fase ad un'altra.

Più specificamente la distanza semantica, per ciò che concerne l'azione, riguarda la relazione fra le intenzioni dell'utente ed il significato dei comandi che è possibile eseguire sull'interfaccia. In altre parole, tale distanza è funzione della facilità con cui l'interfaccia fornisce mezzi e strumenti per esprimere le intenzioni dell'utente. Intenzioni, si noti bene, pertinenti al dominio di applicazione dell'attività (es., mantenere costante la pressione in una valvola, modulare il livello di luminosità di una sala, sovrapporre ad un grafico un testo con differenti orientamenti, etc.). Sul lato della valutazione, la distanza semantica si riferisce all'elaborazione che è richiesta all'utente per determinare se le condizioni di soddisfacimento delle proprie intenzioni sono state realizzate. Se i termini dell'output non sono quelli definiti nell'intenzione dell'utente, a costui viene di fatto richiesto di tradurre l'output in termini che siano compatibili con l'intenzione, al fine di valutare i risultati della sua azione. Per esemplificare quando detto, si consideri il caso del testo scritto su un grafico. Se l'aver compiuto l'azione di scrivere non fornisce immediatamente informazioni sul suo esito (posizione del testo in relazione al grafico) ma richiede passi successivi per ottenere la rappresentazione di tale informazione, allora la distanza cresce proporzionalmente alle informazioni che devono essere gestite per anticipare l'esito o per produrlo materialmente. Il vecchio sistema operativo DOS e molte delle sue applicazioni sono caratterizzate da elevate distanze semantiche sia sul lato dell'esecuzione che sul lato della valutazione. Questo comporta che l'intenzione di avere, per esempio, un testo sottolineato, venga tradotta in una serie articolata di comandi dalla sintassi rigida che devono essere ricordati in modo corretto ed eseguiti. Il risultato poi può essere valutato solo operando ulteriori trasformazioni mentali sugli effetti realmente prodotti oppure eseguendo la serie di operazioni necessaria per ottenere il risultato finale (il testo stampato) e solo allora valutarne il risultato.

La distanza referenziale o "di riferimento" intercorre fra il significato di una espressione e la sua forma fisica, sia sul lato dell'esecuzione che su quello della valutazione. Le interfacce a linea di comando hanno solitamente una distanza di riferimento elevata a causa della relazione arbitraria tra la forma di una espressione ed il suo significato. Per esempio, se vogliamo duplicare con un solo comando un certo numero di files da un disco ad un altro usando una interfaccia di tipo DOS è necessario attribuire ai files di interesse un nome composto con una appendice comune, e poi eseguire (digitare) il comando di copia. Con una interfaccia grafica a manipolazione diretta le azioni ed il loro significato in relazione al compito sono un tutt'uno: si selezionano con il mouse i files di interesse e si trasporta la loro immagine direttamente sul nuovo disco. In questo caso non c'è una

distanza tra l'intenzione e l'azione visto che l'azione di trascinare l'icona dei file sul disco di destinazione coincide con l'intenzione di farne una copia.

Vi è infine la distanza inter-referenziale, che riguarda la relazione fra le forme di input e quelle di output presenti nel corso dell'interazione uomo-computer. Questa distanza è massima quando le due forme sono completamente distinte. Si pensi ad esempio al caso di una tastiera con tasti rigidi (dove non ci sia feedback in corrispondenza dell'azione di premere) e ad un video che presenta solo una lettera per volta. Nella misura in cui l'output del sistema può essere liberamente usato come input o perlomeno come una sua componente, tale distanza è ridotta. Nelle interfacce grafiche un'icona, ad esempio, rappresenta un output del sistema ma è nello stesso tempo oggetto di possibili azioni dell'utente, ed i risultati di queste azioni costituiscono a loro volta nuovi input per il sistema. Nelle interfacce tradizionali tipo DOS, l'input e l'output non solo sono operativamente differenti ma spesso hanno un lessico differente (gli stessi termini hanno significati differenti se sono usati come input o come output). Il modello dell'azione presentato, pur non avendo la pretesa di essere un modello accurato dei processi cognitivi di interazione con gli strumenti, ha però il vantaggio di fornire una guida teoricamente fondata per comprendere, valutare e, come vedremo più avanti, progettare l'interazione uomo-computer. Proviamo ora a esercitarci sui concetti di base prima di vedere quali profondi effetti possano avere gli strumenti sulla cognizione umana.

Finestra: la distanza inter-referenziale... e le altre. Voglio spostare ad esempio la sedia su cui sono seduto per avvicinarmi al tavolo, il luogo in cui applico le mie mani e concentro il mio sguardo per dirigere l'azione coincidono con il luogo in cui ricevo informazioni di ritorno sui risultati in un continuum spazio-temporale. Non è presente alcuna distanza inter-referenziale. Ma appena inizio ad usare degli strumenti, ovvero dei sistemi di transduzione energetica e/o elaborazione computazionale, la distanza emerge: da distanze piccole quali quelle poste da sistemi meccanici elementari tipo le leve, a distanze enormi quali quelle poste dai sistemi a controllo remoto, dove se non si progetta adeguatamente il feedback creando l'illusione di una manipolazione più o meno diretta degli oggetti il senso stesso dell'azione si smarrisce. Un'obiezione ingenua alla creazione della distanza inter-referenziale consiste nell'affermare che gli strumenti stessi, in quanto oggetti fisici, permettono i propri continuum spazio-temporali di interazione ed è quindi questa l'interazione che bisogna progettare (ad es. la tastiera e lo schermo del computer, premo i tasti e vedo il risultato sullo schermo). Tale obiezione implica che gli strumenti in quel momento cessano di essere i mediatori dell'attività per divenire invece oggetto dell'attività, ed è proprio quello che dobbiamo evitare nel corso dello svolgimento dell'interazione uomo-macchina. Gli strumenti possono, e devono, essere oggetto dell'attività ma solo in fase di apprendimento, perché se abbiamo comprato un frullatore è per bere un frullato, non per trovare il modo di farlo smettere di frullare, così se abbiamo acquistato un computer è per comunicare con i nostri amici e non per essere costretti ad imparare protocolli di comunicazione. E' per questo motivo che si dice comunemente che la migliore interfaccia è quella "invisibile", ovvero quella che non diviene mai, o perlomeno spesso, oggetto dell'attività.

L'evoluzione principale dell'interazione uomo-macchina negli ultimi vent'anni riguarda un tentativo scarsamente riuscito di ridurre le enormi distanze di cui sono stati dotati gli strumenti da progettisti che progettavano le funzioni per gli strumenti e non le funzioni degli strumenti per l'interazione con l'essere umano.

0.2. La cognizione distribuita

L'approccio teorico noto con il nome di "Cognizione Distribuita" sostiene che l'attività cognitiva umana non è caratterizzata esclusivamente dall'attività cerebrale bensì è distribuita tra il cervello e l'ambiente esterno inclusi gli artefatti che l'uomo adopera. Questa tesi risale alla scuola storicoculturale sovietica che sosteneva che tutti i tipi di attività umana cosciente sono sempre formati con l'ausilio di strumenti esterni. Vygotskij (1974) definì il principio sottostante la distribuzione dell'attività cognitiva come il principio dell'organizzazione extracorticale delle funzioni mentali complesse. La Cognizione Distribuita (Norman, 1988, Hutchins, 1995) ha elaborato solo alcuni dei principi proposti dall'approccio sovietico, quelli legati appunto al ruolo di mediazione degli artefatti nel definire le proprietà della cognizione umana.

Ha però prodotto ampia evidenza empirica sul fatto che l'azione umana, sia in contesti reali di lavoro (Hutchins, 1995) che in condizioni sperimentali (Zhang & Norman, 1994) non può essere studiata analizzando le prestazioni cognitive individuali, la cosiddetta cognizione nella testa, in quanto la cognizione umana è distribuita tra gli individui e gli artefatti presenti nell'ambiente esterno. Secondo questo approccio, un sistema "cognitivo" è un sistema composto da un individuo che interagisce con un insieme di strumenti oppure da un gruppo di individui che interagiscono tra loro mediante artefatti esterni. Un sistema cognitivo così definito fornisce le risorse informative necessarie all'esecuzione di un'attività.

Queste risorse possono essere rappresentate internamente, nella mente dell'individuo, o nell'ambiente esterno, in quest'ultimo caso possono trovare varie forme di rappresentazione. La distribuzione di queste risorse è la base da cui partire per comprendere il funzionamento dell'interazione e per imparare a progettare.

Per meglio comprendere il ruolo degli artefatti cognitivi nel caratterizzare l'attività cognitiva umana, proviamo a fare il seguente gioco: il gioco del 15. Il gioco si svolge tra due giocatori senza l'uso di alcun supporto esterno come carta e penna o le dita della mano. L'obiettivo è duplice: fare 15 dalla somma di tre numeri scelti tra i nove numeri da 1 a 9, ed evitare che l'avversario possa fare 15 per primo. I due giocatori a turno scelgono un numero da 1 a 9. Ogni numero può essere scelto una sola volta, quindi se un numero è stato preso da un giocatore non è più disponibile per l'altro giocatore. Come detto sopra l'obiettivo del gioco è fare 15 mettendo insieme qualunque tripletta di numeri la cui somma sia 15.

Non c'è il rischio di sballare poiché non è importante quanti numeri sono stati scelti da un giocatore, importa solo che tra questi numeri selezionati ce ne siano tre e solo tre la cui somma sia 15. Questo gioco non è molto diffuso e neppure troppo divertente. Infatti è abbastanza difficile tenere a mente nello stesso tempo i numeri che ogni giocatore ha scelto, quelli scelti dall'altro per evitare che faccia 15, e fare le somme. In questa configurazione del gioco, e cioè senza l'ausilio di supporti esterni quali la scrittura è facile commettere degli errori.

Eppure con ogni probabilità il lettore avrà giocato non poche volte a questo gioco, in una differente configurazione, nota con il nome di "tris" o "filetto". Il gioco del 15 e il gioco del tris sono isomorfi, cioè formalmente identici ma la scelta della rappresentazione in un caso e nell'altro modifica l'attività e quindi la relativa difficoltà. Infatti nel caso del gioco del 15 il

problema da risolvere è aritmetico. Nel caso del tris il problema ha una rappresentazione spaziale: lo scopo del gioco è di fare tris con tre simboli dello stesso tipo su una matrice 3X3, per orizzontale, diagonale o verticale. La rappresentazione spaziale del gioco lo rende semplice, tanto che i giocatori diventano presto esperti e la partita finisce in parità. Nel caso della rappresentazione aritmetica il gioco è lento e piuttosto difficile da risolvere, perché i giocatori fanno molti errori dovuti alla difficoltà di ricordare e computare le combinazioni di tre numeri per arrivare a 15. La struttura formale del tris e del 15 è identica, ciò che ne fa due diverse realtà psicologiche è il diverso uso degli artefatti cognitivi, di strumenti per la rappresentazione, la conservazione e la manipolazione di informazioni (Norman, 1993).

Nel gioco del tris sono presenti alcune caratteristiche fondamentali degli artefatti cognitivi: a) distribuzione nel tempo dell'attività; b) distribuzione nello spazio dell'attività; c) modifica dell'attività. Nel tris, la distribuzione nel tempo dell'attività comporta che le combinazioni vincenti siano state pre-calcolate e, attraverso una particolare rappresentazione esterna (distribuzione nello spazio), configurate in relazioni geometriche. Ne consegue una modifica dell'attività cognitiva e una diversa modalità di interazione fra l'uomo e gli strumenti che permettono di condurre il gioco.

In realtà è importante notare che data la struttura formale sottostante al 15 e al tris le persone percepiscono l'attività relativa come gioco solo nella particolare configurazione di distribuzione spazio-temporale delle informazioni fornita dal tris e solo per un periodo limitato di tempo. Infatti, giocatori minimamente esperti giungono presto alla conclusione che il gioco può, con semplici strategie, condurre a situazioni in cui non è possibile per nessuno dei due giocatori vincere.

È importante sottolineare che il ruolo fondamentale degli artefatti non consiste in una modifica delle capacità di elaborazione della mente umana bensì in una modifica del contenuto delle conoscenze che sono coinvolte nell'elaborazione (Rizzo, 1999). Ma poiché la nostra abilità nell'elaborare conoscenze dipende criticamente dal loro contenuto, come è stato mostrato in compiti di ragionamento deduttivo (vedi ad es; Johnson-Laird et al. 1972; Girotto et al. 1988), di apprendimento nei processi decisionali (Rizzo et al., 1993; Rizzo, Parlangei, 1994) e nelle differenze culturali nel ragionamento (D'Andrade, 1989), di fatto un artefatto cognitivo modificando il contenuto modifica anche la possibilità di essere creativi o stupidi. È proprio questa la differenza sostanziale tra una progettazione riuscita ed una sbagliata.

0.3. La progettazione di interfacce uomo-computer

La progettazione e lo sviluppo di interfacce interattive è con ogni probabilità la parte di lavoro più intensa e difficile nel processo di sviluppo di strumenti informatici. Per avere un'idea di ciò, basti pensare che il 50% del codice di un moderno software è dedicato all'interfaccia e che il 75% delle revisioni di software riguardano ancora l'interfaccia (Shneiderman, 1998). Le principali ragioni di ciò risiedono nel fatto che progettare un'interfaccia prevede una varietà di scelte e decisioni progettuali che coinvolgono gli utenti e i possibili compiti che potranno essere svolti attraverso l'interfaccia, e la gran parte delle conseguenze di queste decisioni sono imprevedibili.

È principalmente per questo che la progettazione di interfacce interattive è un'attività sperimentale per la quale sono state concepite metodologie di sviluppo (ad es. Shneiderman, 1992), tecniche di valutazione (ad es. Dumas e Redish, 1993) e vengono costantemente proposti nuovi approcci come il contextual design (Beyer et al. 1997). Progettazione, implementazione e valutazione sono state tradizionalmente considerate fasi separate dei processi di sviluppo di sistemi uomo-macchina. Uno dei contributi più importanti nella progettazione di sistemi interattivi è stata l'introduzione del concetto di progettazione iterativa, in cui la progettazione e la valutazione vengono iterate fino al raggiungimento di un risultato soddisfacente. La fase di valutazione pervade completamente questo processo: nella progettazione è necessario valutare il sistema attuale, l'attività umana e il contesto in cui si svolge, le soluzioni di design tipicamente in forma di prototipi, e il sistema finale.

I principi da adottare per una progettazione ergonomica, centrata sull'utente, sono vari ed hanno ricevuto formulazioni diverse, ma ancora oggi i quattro pionieristici principi, formulati da Gould e Lewis (1983), rappresentano un valido sistema di riferimento a cui tutte le formulazioni successive si rifanno:

- i) Comprensione dell'utente. Bisogna formulare una rappresentazione esplicita sia delle competenze cognitive e attitudinali dell'utente che della natura cognitiva del lavoro che deve essere svolto.
- ii) Progettazione interattiva. Un campione potenzialmente rappresentativo dell'utenza deve far parte del gruppo di progettazione per tutto il processo di sviluppo.
- iii) Misurazione costante dei risultati conseguiti. Sin dalle prime fasi della progettazione i potenziali utenti finali devono essere coinvolti nella sperimentazione dell'usabilità del prodotto, e il sistema uomo-computer deve essere testato e valutato nel suo insieme.
- iv) Progettazione iterativa. La progettazione deve realmente avere un carattere ciclico: progettare, testare e misurare, riprogettare; sino ad una prestazione del sistema uomo-computer che soddisfi lo scopo per il quale quel sistema è stato concepito.

In ciò che segue illustreremo le fasi del ciclo di progettazione iterativa centrata sull'utente, indicando per ciascuna le principali metodologie di riferimento. Queste non potranno essere presentate analiticamente ma solo sommariamente in modo da dare un'idea della natura del metodo, verranno però riportati i riferimenti bibliografici dove il lettore potrà trovare la descrizione puntuale dei metodi esposti.

0.4. Analisi dell'attività umana e del contesto d'uso

La conoscenza dell'utente, dell'attività e del contesto d'uso sono una fase cruciale e delicata della progettazione di sistemi interattivi. Data l'irrinunciabile necessità di conoscere gli utenti e le loro attività sono stati proposti nel corso degli anni svariati metodi, anche se non esiste una strategia certa per poter essere sicuri di aver adottato il metodo giusto per condurre l'analisi. I metodi si possono sommariamente distinguere in due ampie classi: analisi dei compiti (o task analysis) e analisi dell'attività. Le due tipologie di analisi poggiano sulla distinzione tra task normativo e l'attività che viene realmente svolta e cioè sulla differenza che intercorre tra il compito così come dovrebbe essere eseguito astraendo dal contesto d'uso, e come lo è di fatto nella realtà. Per una discussione dettagliata tramite analisi di casi studio si veda Rizzo (2000) e Marti (2000). L'analisi dei compiti riguarda situazioni in cui gli obiettivi che devono essere perseguiti siano ben definiti, e in alcuni casi che lo siano anche le modalità e/o i tempi per il conseguimento degli obiettivi. I metodi più noti sono la Hierarchical task analysis (Annett et al., 1971) e GOMS (Card et al., 1983). Questi metodi si basano sulla scomposizione e ricomposizione analitica di ciò che un utente o un gruppo di utenti dovrebbe fare in termini di azioni

e/o processi cognitivi per conseguire degli obiettivi. Per una loro descrizione dettagliata si rimanda a Kirwan e Ainsworth (1992). L'altra classe di metodi riguarda l'analisi dell'attività ed è per certi versi meno formale e più esplorativa, anche perché il fine non è descrivere analiticamente il modo in cui un certo compito viene eseguito bensì cercare di imparare dall'utenza e con l'utenza la natura dell'attività e le conoscenze che sono state prodotte nel tempo con lo svolgimento dell'attività; quindi comprendere perché quell'attività esiste, la sua ontologia. Anche per questa classe esistono vari metodi, ma essi si rifanno in maniera diretta o indiretta all'indagine etnografica.

L'etnografia è uno dei metodi tradizionali dell'antropologia sociale e culturale che studia l'attività umana nell'ambiente socio-culturale in cui si sviluppa (Suchman and Trigg, 1991). Tale studio richiede una estesa osservazione dell'attività "in situ", e cioè nel modo in cui si sviluppa naturalmente all'interno di diversi contesti sociali. Nella progettazione di sistemi uomo-computer la finalità principale nell'adottare questo metodo di indagine è quella di imparare a conoscere l'attività direttamente da chi la svolge, per poter poi progettare sistemi di reale supporto agli utilizzatori.

Generalmente l'indagine etnografica utilizza registrazioni audio e video per raccogliere dati e discuterli con gli utenti e all'interno del gruppo di progetto. Durante questa fase è necessario analizzare con cura i seguenti aspetti:

- a) i principali processi lavorativi,
- b) le caratteristiche e i ruoli degli attori dei processi,
- c) gli obiettivi e il contenuto del processo e delle attività svolte dagli attori,
- d) gli strumenti di supporto alle attività e loro caratteristiche
- e) la normativa e le procedure che regolano l'attività,
- f) le routines e le pratiche di lavoro,
- g) i flussi di comunicazione,
- h) le principali criticità dell'attività.

Per una descrizione dettagliata sui metodi di indagine etnografica si rimanda a Middleton e Engestrom (1996). L'analisi etnografica è un metodo estremamente efficace per l'analisi dell'attività umana e quindi una fonte preziosa e irrinunciabile nella progettazione di sistemi interattivi. La ragione principale è che consente un'analisi approfondita dell'attività e del contesto socio-tecnico, evitando la rischiosa confusione tra l'attività così come viene realmente eseguita (la pratica, le routine) e l'attività normativa, quella cioè descritta da manuali e regolamenti. Questo livello di conoscenza porta ad una migliore definizione dei requisiti utente e a una maggiore consapevolezza da parte del progettista della natura dell'attività.

0.5. Sviluppo di prototipi

Lo sviluppo di prototipi è uno dei punti chiave della progettazione centrata sull'utente. I prototipi consentono l'implementazione di un processo iterativo di progressivo affinamento di idee di design. Questo processo si articola nelle seguenti fasi:

- Simulazione di alternative di design usando prototipi "a bassa fedeltà".
- Valutazione di prototipi in fase precoce del ciclo di vita del sistema finale.
- Prototipazione iterativa secondo cicli di progettazione, valutazione e riprogettazione.

I prototipi possono assumere forme e significati diversi. Nel disegno industriale per prototipo si intende un modellino dell'oggetto che si vuole progettare, per un progettista software si tratta tipicamente di una simulazione su schermo della forma e del comportamento del sistema, per un programmatore un prototipo è un programma di test mentre per un esperto di fattori umani un prototipo può essere uno *storyboard*, cioè la descrizione di una sequenza di azioni che illustra in forma narrativa aspetti salienti del sistema nel contesto d'uso (Hooper, et al. 1996). Nella progettazione centrata sull'utente il prototipo ha un'accezione molto ampia di cui Houde e Hill (1998) danno la seguente definizione:

"...un prototipo è una rappresentazione di una idea di design a prescindere dal media che viene utilizzato per realizzarlo. Un prototipo può mostrare aspetti diversi di un artefatto e viene raffinato iterativamente fino ad evolvere nel sistema finale".

Ciò che caratterizza un prototipo non è dunque il media o gli strumenti che vengono utilizzati per realizzarlo ma, piuttosto, il modo in cui il progettista lo usa per esplorare scelte di design o per dimostrare alcune caratteristiche del sistema finale. Per chiarire questo concetto, Houde e Hill propongono un modello che ha due obiettivi: descrivere gli aspetti del un sistema finale che possono essere dimostrati attraverso prototipi, e mostrare il modo in cui vari prototipi dello stesso sistema possono contribuire allo sviluppo del sistema finale. Il modello risponde alle seguenti domande che riguardano il sistema finale:

- 1) Quale ruolo l'artefatto giocherà nella vita dell'utente?
- 2) Che aspetto deve avere (look&feel)?
- 3) Come dovrebbe essere implementato

Le dimensioni del modello sono definite come *Ruolo*, *Look and feel*, e *Implementazione*. Il *ruolo* si riferisce alla funzione che l'artefatto ha nella vita dell'utente, il modo in cui è utile all'utente. Per dimostrare questo aspetto si richiede dunque che sia definito puntualmente il contesto d'uso. Il *look and feel* riguarda l'esperienza sensoriale concreta nell'uso dell'artefatto; ciò che l'utente vede, sperimenta e sente durante l'uso. E' necessario dimostrare questo aspetto simulando concretamente l'esperienza sensoriale dell'utente. L'*implementazione* riguarda le caratteristiche e le componenti tecniche del sistema. Per dimostrare la fattibilità di questi aspetti è necessario che siano sviluppate almeno parzialmente alcune funzioni del sistema. Il modello "a triangolo" mostra come un problema di design possa essere affrontato da molti punti di vista. L'obiettivo di questo modello è quello di supportare il progettista nell'individuare e visualizzare le tre dimensioni.

E' infatti consigliabile sviluppare prototipi dimostrativi di ciascuna dimensione piuttosto che sviluppare un prototipo monolitico che integra tutte le caratteristiche dall'inizio. Infatti è soltanto dopo la progettazione iterativa di prototipi sulle tre dimensioni, che un singolo prototipo integrato può cominciare a evolvere acquisendo un posizione centrale nel modello a triangolo. Esistono vari metodi per lo sviluppo di prototipi. Di seguito sono elencate le tecniche più diffuse, ma è necessario ricordare che qualsiasi sia la tecnica utilizzata, ciò che è veramente essenziale è che ogni soluzione di design sia valutata con gli utenti prima dell'implementazione del sistema finale.

Prototipi cartacei. Modalità poco dispendiosa ma estremamente efficace di progettazione e valutazione. Per le interfacce software si tratta di una serie di "schizzi" su carta delle schermate e del loro contenuto. All'utente viene presentata l'interfaccia e attraverso un piccolo scenario d'uso viene richiesto di provare a utilizzarla. Lo sperimentatore provvede a fornire le schermate successive simulando il comportamento del sistema. La prima e la più innovativa interfaccia grafica a manipolazione diretta, la Xerox Star, fu progettata attraverso prototipi cartacei (Liddle, 1996).

Scenari. Scenario-based design (Carroll, 1995) è un metodo che permette ai progettisti e agli utenti di descrivere attività esistenti o di prevedere o immaginare nuove attività che possono essere prodotte dall'interazione con un nuovo artefatto. Lo scopo di uno scenario è quello di esprimere in modo dettagliato una situazione di uso e di documentare punto per punto le azioni eseguite dagli utenti. Uno scenario può essere rappresentato da descrizioni testuali, immagini, diagrammi, video e può essere usato in diverse fasi del processo di sviluppo di un artefatto. Gli scenari sono uno strumento di progettazione molto versatile ed efficace. Questo è il motivo per cui è possibile utilizzarli in varie fasi del processo di sviluppo di un sistema, in particolare per la strutturazione dei dati raccolti con l'osservazione dell'attività; per immaginare le caratteristiche del sistema futuro (fase creativa del processo di design, chiamata anche "envisioning"), per la progettazione parallela di soluzioni alternative in fase di prototipazione, e per la valutazione (si veda il paragrafo successivo).

Mock-ups. Si tratta di modellini dell'artefatto finale. Sono generalmente utilizzati per verificare caratteristiche fisiche e percettive legate all'uso dell'artefatto come il peso, le dimensioni, la maneggevolezza. Una scatola di cartone per le pizze può rappresentare le dimensioni di un computer, una busta di sabbia può essere utilizzata per rappresentare il peso di un palmare, e una torcia tascabile può giocare il ruolo di sistema di puntamento remoto. E' difficile sottovalutare la loro importanza nel prototipare sia il look&feel che il ruolo di un artefatto.

Mago di Oz. Una simulazione in laboratorio o sul campo in cui lo sperimentatore è nascosto (“il Mago” o “l’Uomo dietro la tenda”) e gioca il ruolo del sistema. Il soggetto interagisce con il sistema emulato dal “Mago” come se si trattasse del sistema reale. Questa tecnica è stata utilizzata con successo già nel 1971 da Erdmann e Neal (1971) per simulare un sistema di prenotazione di biglietti aerei in un aeroporto. Uno sperimentatore nascosto interagiva attraverso il computer con i clienti che provavano a utilizzarlo.

Video. Dimostrazioni video che esemplificano le caratteristiche del sistema. E’ una tecnica particolarmente adatta a comunicare soluzioni di design in team interdisciplinari e soprattutto agli utenti finali del sistema. Sono una estensione e integrazione degli scenari e del Mago di Oz, costosi da realizzare ma efficaci nel comunicare le idee più ardite e più lontane dagli artefatti comuni.

Simulazioni al computer. Come i prototipi cartacei, sono schermate che danno un’idea dell’interazione con il sistema, ma invece che essere realizzati su carta, vengono sviluppati utilizzando generalmente strumenti grafici e di animazione.

Storyboard. Uno scenario d’uso narrato in forma testuale o rappresentato sottoforma di disegni e fumetti, sia in formato cartaceo che attraverso realizzazioni software. Sfrutta le tecniche usate nel cinema, nel teatro, nel fumetto per sviluppare le sceneggiature.

0.6. Valutazione

La valutazione è una fase essenziale della progettazione centrata sull’utente. Secondo questo approccio infatti non esiste attività di design che non sia accompagnata da una attività di valutazione, in un ciclo iterativo che si completa quando il sistema soddisfa i requisiti utente. La valutazione può essere effettuata in momenti diversi del ciclo di vita di un sistema, e quindi “in corso d’opera” (formative evaluation) o in fase avanzata del processo di sviluppo (summative evaluation). In ciò che segue forniremo una panoramica dei metodi di valutazione concentrandoci principalmente su Cognitive Walkthrough, Usability testing e Valutazione Euristica.

Cognitive walkthrough. Il Cognitive Walkthrough (CW) è un metodo ispettivo largamente utilizzato nella valutazione di interfacce utente (Polson et al., 1992). Si applica a prototipi già realizzati e funzionanti che forniscono una risposta ad ogni azione dell’utente. Sulla base di scenari d’uso, vengono selezionati dei task che vengono poi eseguiti passo dopo passo utilizzando l’interfaccia. Durante l’esecuzione di ogni task, il valutatore risponde a domande formulate a partire dalla teoria cognitiva del modello “esplorativo dell’apprendimento” (Wharton et al, 1994), secondo cui un utente inesperto apprende il funzionamento del sistema tramite l’esplorazione dell’interfaccia. Questo modello tenta di esplorare le relazioni tra obiettivi, azioni e stati visibili dell’interfaccia per rilevare problemi relativi alla facilità di apprendimento del sistema. In ciò che segue proponiamo le modalità di esecuzione del CW, attraverso l’applicazione di un diverso modello cognitivo, quello di Holland, Hutchins e Norman. Per una descrizione dettagliata del metodo su un caso studio rimandiamo a Rizzo et al., (1998). Questo tipo di CW analizza le azioni dell’utente che tenta di raggiungere un qualche obiettivo definito all’interno di uno scenario d’uso. Il valutatore esplora il sistema tentando di scoprire le azioni che potrebbero contribuire all’esecuzione del compito definito nello scenario, selezionando quelle la cui apparenza/descrizione sembra più vicina a ciò che sta tentando di fare. Una volta eseguita l’azione, il valutatore interpreta la risposta del sistema e valuta il progresso fatto verso il raggiungimento dell’obiettivo. E’ possibile che durante questo processo l’utente debba riformulare l’obiettivo iniziale nel caso realizzi un’impossibilità di soddisfarlo. La sessione di test si svolge una volta definito uno scenario d’uso: il valutatore prova ad eseguire il compito definito dallo scenario cercando di rispondere alle seguenti domande:

D1: L’azione corretta è disponibile e sufficientemente evidente all’utente e corrisponde alle intenzioni espresse? (Distanza tra intenzione e azione)

D2: L’utente è in grado di collegare la descrizione dell’azione corretta con quel che sta tentando di fare? (Distanza Azione-Forma)

D3: L’utente riceve feedback dell’azione nello stesso spazio e con la stessa modalità dell’azione? (Distanza Input Azione - Output Feedback)

D4: L’utente interpreta correttamente la risposta del sistema ad una azione corretta? (Distanza Effetti-Forma)

D5: L'utente è in grado di valutare correttamente i risultati? (Distanza Forma - Valutazione)

D6: Se lo scopo è sbagliato o può essere migliorato, l'utente è in grado di comprendere che l'intenzione che sta tentando di realizzare non può essere raggiunta nell'attuale situazione? (Distanza Azione/ Risultati ottenuti).

Ogni volta che la risposta non è completamente affermativa il valutatore la inserisce in una lista di problemi rilevati. Alla fine della sessione la lista viene comunicata al team di progettazione (che include rappresentanti di utenti finali) con la specifica delle possibili soluzioni. Una volta che le soluzioni sono state implementate, il sistema viene sottoposto di nuovo ad una fase di test.

Usability testing. Il test di usabilità, o usability testing, è un metodo che consente di analizzare le reazioni dell'utente verso il sistema in fase di design. Agli utenti viene richiesto di eseguire un insieme di compiti utilizzando il sistema e lo sperimentatore osserva il loro comportamento e raccoglie dati empirici sui risultati della sessione di test. Questi risultati sono utilizzati per migliorare il sistema, riprogettare i difetti e valutare il modo in cui la riprogettazione è stata realizzata secondo i principi della progettazione iterativa. Il test di usabilità è fortemente orientato ad analizzare nel dettaglio la performance dell'utente nell'uso del sistema, e in questo si differenzia in modo sostanziale dal Beta testing, una metodologia usata largamente in ambiente industriale. Il Beta testing si effettua sul prodotto finito: il sistema viene consegnato e dopo un periodo di prova la compagnia contatta l'utente per chiedere informazioni sull'affidabilità del sistema e la soddisfazione durante l'uso. Non ci sono sperimentatori che affiancano gli utenti. Gli eventuali problemi riscontrati dagli utenti vengono (a volte) raccolti in un documento e consegnati alla compagnia. In questo modo è difficile diagnosticare con precisione i problemi e dunque correggerli. Il test di usabilità invece viene effettuato con utenti veri e su compiti specifici, selezionati un modo che siano rappresentativi dell'attività. Il test viene video-registrato e i risultati tradotti in indicazioni puntuali di riprogettazione. E' necessario però che il test venga eseguito con cura per evitare problemi di affidabilità dei risultati (gli utenti che partecipano al test devono essere rappresentativi di tutta la popolazione di utenti del sistema) e di validità (la validità dei risultati dipende dall'accuratezza con cui è stato pianificato ed eseguito il test). Il test di usabilità si articola nei seguenti passi:

Definire gli obiettivi del test. Gli obiettivi possono essere generici, e quindi riguardare il miglioramento del prodotto, l'aumento della soddisfazione dell'utente, la creazione di un prodotto facile da apprendere e da usare; oppure specifici, come la valutazione di un menu, di funzioni di navigazione, la leggibilità delle etichette, ecc.

Definire il campione di soggetti che partecipano al test. Il campione di soggetti che partecipano al test deve essere rappresentativo dell'intera popolazione di utenti. I soggetti quindi devono essere campionati per caratteristiche che riguardano esperienza, (esperto/novizio), età, frequenza d'uso del sistema, conoscenza di sistemi simili, e il numero di utenti dipende dagli obiettivi del test. Nielsen e Molich (1990) sostengono che il 50% dei maggiori problemi di usabilità possono essere rilevati con 3 utenti. Altri autori riportano dati secondo cui con 5 utenti è possibile rilevare il 90% dei problemi di usabilità.

Selezionare task e scenari. I compiti che gli utenti devono eseguire durante il test devono essere reali, cioè rappresentativi dell'attività. Gli scenari possono essere selezionati a partire dall'osservazione dell'attività fatta durante la fase iniziale del processo di sviluppo, oppure costruiti ad hoc per testare situazioni inaspettate in cui intervengono turbative.

Decidere come misurare l'usabilità. Prima di effettuare un test di usabilità è necessario definire il modo in cui i risultati ottenuti verranno misurati. Il tipo di misurazioni varia da valutazioni soggettive (soddisfazione, difficoltà d'uso) a misure quantitative come il tempo di esecuzione del compito, il numero e tipologia di errori, il numero di task completati, il numero di richieste di aiuto (manuale, help, sperimentatore). Qualsiasi sia il tipo di analisi dei dati, è necessario che i risultati rimangano anonimi e che i soggetti siano informati sull'uso che ne viene fatto. Durante il test, lo sperimentatore può ricorrere a varie tecniche di conduzione della valutazione. Tra queste ricordiamo il "think aloud" (al soggetto viene richiesto di "pensare ad alta voce", cercando di esplicitare le azioni che esegue e il motivo per cui le esegue); "co-discovery" o collaborative approach (due

soggetti eseguono il compito insieme dandosi supporto); active intervention (lo sperimentatore stimola i soggetti a ragionare su quanto è accaduto durante il test).

Preparare il materiale e l'ambiente sperimentale. L'ambiente in cui la valutazione viene eseguita deve essere approntato posizionando la video camera, stabilendo i ruoli del team di sperimentatori e fornendo il materiale su supporto (manuali, carta e penna, video cassette). Non è necessario un laboratorio. La valutazione deve comunque essere preceduto da una fase pilota in cui vengono provate e raffinate le procedure di test.

Valutazione euristica. La valutazione euristica è un metodo ispettivo a basso costo che viene eseguito durante la progettazione di un'interfaccia. Consiste essenzialmente nel verificare che un insieme di principi di buona progettazione, espressi sotto forma di linee guida, siano stati applicati correttamente. Ogni linea guida si riferisce ad una categoria di problemi di design a cui è applicabile, e quando una linea guida viene applicata ad uno spazio di problemi la trattiamo come "euristica" e cioè come un principio derivato dall'esperienza passata nel trattare e risolvere quella categoria di problemi. In ciò che segue proponiamo un insieme di linee guida formulate da Nielsen e Mack (1994).

Prevedere un dialogo semplice e naturale. Questa linea guida si concretizza in vario modo nel design. Alcuni esempi: strutturare l'accesso alle informazioni in modo da rispettare le aspettative dell'utente, fornire solo le informazioni necessarie, rimuovere o nascondere le informazioni irrilevanti e di cui c'è bisogno solo raramente, raggruppare le informazioni che devono essere utilizzate in modo integrato.

Parlare il linguaggio dell'utente. Significa usare una terminologia basata sul linguaggio parlato dagli utenti nello svolgimento dei loro compiti e non in termini tipici del sistema, vedere l'interazione dalla prospettiva dell'utente, usare delle entità (parole, simboli, icone) mnemoniche e altamente significative in modo da facilitare il ricordo e il riconoscimento.

Minimizzare il carico di memoria dell'utente. Questa linea guida deriva direttamente da un principio ben noto in psicologia cognitiva, quello cioè di promuovere il riconoscimento piuttosto che il ricordo e ridurre il più possibile le informazioni che devono essere tenute a mente. Nel design questo principio si può realizzare in vari modi:

- Descrivere e spiegare sia il formato dei dati di input che i valori di default.
- Usare un piccolo numero di regole applicate universalmente.
- Usare dei comandi generici ed applicarli a tutti gli oggetti dell'interfaccia. Per es. le funzioni taglia, copia, incolla valgono per i caratteri, le parole i paragrafi, i file.

Essere coerenti. La coerenza di un'interfaccia deve essere realizzata rispetto a tre componenti principali:

- Coerenza degli effetti: le stesse parole, gli stessi comandi, le stesse azioni devono produrre gli stessi effetti in situazioni equivalenti.
- Coerenza del linguaggio e della grafica: le stesse informazioni/comandi/pulsanti devono essere disposti nella medesima collocazione e nello stesso ordine in tutte le schermate/dialog box mantenendo lo stesso significato.
- Coerenza dell'input: la sintassi deve essere la stessa in tutto il sistema.

Fornire un feedback costante. Durante l'interazione con il sistema, è necessario che l'utente possa contare su un feedback immediato ed efficace sugli effetti dell'azione appena compiuta in modo da poter essere sempre informato sugli effetti che ha avuto la propria azione sul sistema, le possibili conseguenze della propria azione e il nuovo stato del sistema a seguito dell'azione compiuta. Il feedback a questo proposito deve essere il più specifico possibile come nel caso di operazioni che richiedono un'attesa in cui è necessario fornire un'indicazione della quantità di tempo necessaria a completare il comando. Quando il compito è particolarmente complesso il feedback deve essere fornito sia durante l'esecuzione che alla fine del compito.

Rendere evidenti le vie di uscita. Nessun utente desidera sentirsi intrappolato nel sistema. Affinché questo spiacevole effetto non si verifichi, è necessario fornire semplici possibilità di uscita per quante più situazioni possibile. Esempi di questo tipo sono i comandi “CANCELLA” nei dialog box che compaiono quando il sistema è in attesa dell’input dell’utente; “UNDO” per tornare indietro allo stato precedente; “INTERROMPI” per le operazioni lunghe; “QUIT” per uscire dal programma in ogni momento; oppure le configurazioni di “defaults” per ripristinare un elenco di proprietà.

Fornire delle scorciatoie. Gli utenti esperti devono poter svolgere velocemente le operazioni che usano più frequentemente. Per questo è opportuno offrire strategie per aumentare la velocità delle operazioni tramite l’uso del mouse e della tastiera, oppure dare la possibilità di digitare in anticipo i dati prima che il sistema sia pronto, o ancora permettere “salti nella navigazione” (andare direttamente ad un certo punto senza passare per i nodi intermedi).

Aiutare gli utenti a riconoscere, diagnosticare e recuperare gli errori. Un buon messaggio di errore permette all’utente di sapere quale è il problema e come prevenirlo o recuperarlo. Si pensi ad un messaggio di errore del tipo “errore 25”. Oltre a non essere di alcun aiuto per l’utente, viola la seconda delle linee guida che abbiamo illustrato in questa sezione, e cioè quella di usare il linguaggio dell’utente e non quello della macchina. Un messaggio molto più efficace nello stesso contesto potrebbe essere: “Non posso aprire il documento perché l’applicazione non è nel sistema. Vuoi provare ad aprirlo con Teachtext?”. Gli errori devono poter essere diagnosticati e recuperati ma soprattutto prevenuti. A questo proposito è indispensabile utilizzare funzioni costrittive che impediscano la continuazione di azioni errate avvertendo l’utente che si sta verificando una situazione inusuale. In letteratura sono disponibili moltissimi tipi di linee guida, che vanno da principi generali, come quelli illustrati sopra, a indicazioni di dettaglio sull’uso dei colori, lo stile di interazione, il layout del display (icone, menu, etichette). Le linee guida variano anche per tipologie di applicazioni, come quelle per il web, quelle per utenti disabili ecc. Nell’applicare questo metodo di valutazione dunque è necessario anche saper selezionare l’insieme di linee guida più appropriato al problema che stiamo affrontando e utilizzarle anche come indicazioni per fornire soluzioni di re-design. Può essere utile a questo proposito utilizzare una tabella come quella illustrata di seguito, in cui oltre al problema rilevato e al principio violato, sia stimata la severità del problema e una chiara indicazione di miglioramento.

0.7. Ergonomia, Domotica e Wearable sensors

I metodi bioingegneristici applicati all’ergonomia si pongono lo scopo di adattare le condizioni di lavoro alla natura fisio-psicologica dell’uomo. Il punto di partenza è la conoscenza delle capacità fisiche e mentali dell’uomo, il punto d’arrivo vuole essere ottenere contemporaneamente la massima sicurezza, il massimo benessere e il massimo rendimento. Lo sviluppo di opportuni strumenti di analisi e simulazione possono permettere verifiche ergonomiche in tempo reale, quando il nuovo progetto è ancora nelle prime fasi e solo in stato di disegno CAD, o reiterando la procedura nel corso dello sviluppo del prodotto per ottimizzarne progressivamente il design. Le simulazioni dell’interazione uomo-macchina e la loro diretta valutazione ergonomica quantitativa definiscono un nuovo modo di progettazione che viene detto Prototipazione Virtuale. Si parla in questo caso anche di Ergonomia Proattiva.

Per Domotica si intende generalmente l’applicazione di nuove soluzioni tecnologiche in ambito domestico, legate all’elettronica ed alle telecomunicazioni, al fine di migliorare apparati e prodotti già esistenti e di fornire un maggior numero di servizi di utilità domestica. La domotica costituisce anche uno strumento di integrazione degli utenti deboli, atto ad aumentare il comfort e la fruibilità degli spazi e dei prodotti ivi contenuti. In questa linea si sta sviluppando anche il paradigma del comfort attivo, ovvero del comfort che si adatta in funzione all’utente ed al suo stato contingente, grazie a sensori non intrusivi indossabili o distribuiti che rappresentano l’interfaccia tra l’utente e il sistema domotico. Rilevando le esigenze dell’utente in maniera non intrusiva, si ottiene un comfort che si adatta dinamicamente, del tutto trasparente all’utente. L’attività di

ricerca condotta nel settore dell'Ergonomia e della Domotica consiste principalmente nello sviluppo di metodologie innovative e modelli di analisi e simulazione dell'interazione uomo-macchina-ambiente.

I principali metodi e ambiti di ricerca sono:

- analisi e simulazione dei gesti correlati all'uso dell'autovettura, delle interazioni con il sedile e la postura di guida, studio della postura assisa e dei compiti lavorativi da posizione seduta, analisi di materiali per la realizzazione di cuscini ergonomici mediante modelli biomeccanici, analisi del movimento e della postura con sistemi optoelettronici, mappe di pressione con sistemi a matrice piezoresistivi, sedia dinamometrica;
- studio del comportamento motorio e posturale e delle modificazioni antropometriche del corpo umano in microgravità;
- sviluppo di metodi di innovativi per la realizzazione delle calzature sulla base di dati antropometrici e biomeccanici funzionali per la mass-customisation del prodotto;
- realizzazione di tecnologie e metodologie di valutazione funzionale su base quantitativa della mobilità del rachide sano e patologico;
- tecnologie avanzate per l'interfaccia uomo-macchina, sviluppando sistemi evoluti e multi input basati sulle metodologie delle Brain Computer Interfaces (BCI);
- sviluppo e applicazione di wearable sensors e relative metodologie per il rilievo non intrusivo dei segnali biologici associati al comportamento spontaneo del soggetto;
- responsive environments per lo sviluppo del paradigma del comfort ambientale attivo.

0.8. Le nuove tendenze di ricerca

Siamo ormai abituati a pensare all'interfaccia di un computer in termini di schermo, tastiera, eventuali sistemi di puntamento e di applicazioni specifiche per elaborare informazioni. Tuttavia i prodotti commerciali attualmente disponibili non rappresentano altro che la realizzazione degli strumenti concepiti negli anni '50 e '60. Questa situazione è destinata a mutare rapidamente nei prossimi anni, gli stessi creatori delle interfacce grafiche, ad esempio Smith e Tognazzini, hanno affermato come la metafora della scrivania e l'interfaccia grafica tradizionale siano quasi un ostacolo allo sviluppo di nuovi sistemi. La mancanza di una cultura in scienza cognitiva da parte di molti informatici fa sì che questi vedano nelle interfacce grafiche un punto d'arrivo piuttosto che qualcosa da superare.

L'interfaccia utente media i processi di acquisizione, manipolazione, navigazione, analisi, archiviazione, visualizzazione di dati e di condivisione di questi dati con altri. Per sfruttare adeguatamente questi processi le interfacce future saranno tridimensionali, dinamiche, virtuali, pervasive, uditive, guidate dai gesti e immersive. Ovvero le prossime interfacce andranno sempre più nella direzione di adeguarsi alle modalità proprie dell'agire umano. Ad esempio, sono già disponibili commercialmente interfacce che permettono all'utente di muoversi liberamente (senza indossare alcuno strumento di Realtà Virtuale) in uno spazio fisico reale e vedere proiettata la propria immagine su uno spazio virtuale dove sono manipolati vari oggetti. Altre interfacce basate su nuovi paradigmi d'interazione sono attualmente in corso di sperimentazione.

Tra queste vorremmo menzionare HIPS un sistema museale sviluppato all'interno di un progetto europeo coordinato dall'Università di Siena (gli altri partner del consorzio sono l'Università di Edinburgh, University College Dublin, Irst-Irc, Gmd, Sintef, Alcatel). In questo progetto vengono sperimentate nuove metafore di interazione per il passatempo, l'intrattenimento, l'apprendimento. HIPS (*Hyper-Interaction within Physical Space*) è una guida elettronica che permette ai turisti di accedere alle informazioni relative allo spazio circostante (le stanze di un museo o un luogo all'aperto) semplicemente muovendosi. I visitatori infatti portano con sé la guida, un piccolo computer palmare e ricevono le informazioni tramite auricolari. L'informazione presentata al visitatore è generata in modo dinamico, integrata con mappe ed altre indicazioni spaziali che consentono di rintracciare oggetti e luoghi di interesse. Da un punto di vista tecnologico HIPS è basato su di un'architettura client-server. Il client è lo strumento che l'utente porta con sé durante la visita, ovvero il computer palmare su cui si agisce prevalentemente tramite pulsanti hardware e penna elettronica. Il server è invece una stazione fissa, su cui è memorizzata l'informazione, collegata con il palmare attraverso ponti radio. Per

determinare la posizione dell'utente sono utilizzati infrarossi in ambienti chiusi e GPS all'aperto. Dal punto di vista del design dell'interazione, questo sistema offre una serie di possibilità nuove di mediazione dell'attività umana.

Le soluzioni di design realizzate nel sistema sono il frutto di accurati studi, basati principalmente su indagini etnografiche, del comportamento di visitatori in spazi d'arte. Questi studi hanno evidenziato come i movimenti dei visitatori e quindi il loro percorso all'interno del museo, rappresentino una fonte importante per elaborare modelli di attività (Marti et al., 1999). La ricerca sul campo ha evidenziato infatti che i visitatori tendono a muoversi in uno spazio espositivo in modalità omogenee, Le variabili più significative legate al movimento nello spazio fisico (geometria del percorso, tempi di visita, numero di fermate) possono essere isolate e messe in relazione alla spazio informativo, al fine di adattare la presentazione, ad esempio, descrizioni più lunghe e accurate sono fornite a chi cammina lentamente e spende molto tempo davanti alle opere, mentre descrizioni più sintetiche a chi si muove più in fretta. L'interazione quindi è veicolata prevalentemente dal movimento fisico e questo lascia la libertà al visitatore di concentrarsi sull'ambiente circostante che diventa la vera interfaccia del sistema (Marti et al. 2000). L'esperienza di HIPS è indicativa di quanto sia importante analizzare l'attività umana del suo naturale contesto d'uso e di come i risultati dell'analisi possano essere tradotti in specifiche di design.

Nel suo recente libro, *il computer invisibile*, Norman (1999) riporta la seguente storia. Nel 1877 Edison inventò il fonografo. All'inizio la tecnologia era primitiva e delicata e così iniziò un lungo percorso per migliorarla. Tuttavia però non fu chiaro dall'inizio l'uso che di questa macchina si sarebbe potuto fare. Inizialmente si pensava di utilizzarla per dimostrazioni pubbliche a pagamento, con l'idea di realizzare in seguito "l'ufficio senza carta", dal momento che lettere e altri documenti cartacei avrebbero potuto essere registrati nei cilindri e spediti al destinatario. La buona qualità tecnologica di questo prodotto non fu sufficiente a farlo affermare sul mercato. L'errore di Edison fu quello di adottare un approccio centrato sulla tecnologia piuttosto che sull'utente, e i suoi competitori vinsero. I cilindri vennero sostituiti dai dischi che erano meno fragili, più facili da impacchettare e spedire, ed era anche più semplice produrli in grandi quantità.

La storia di Edison insegna che la tecnologia è una parte relativamente facile da modificare nell'evoluzione di un prodotto. Ciò che è difficile, è progettare un prodotto in armonia con il contesto sociale, culturale e organizzativo della comunità che lo utilizzerà. La tecnologia deve aspirare a diventare "invisibile", una sorta di infrastruttura in cui la componente tecnologica non sia l'oggetto ma piuttosto il mediatore dell'attività umana. Una componente essenziale della progettazione di artefatti tecnologici è dunque la comprensione delle caratteristiche distintive dell'attività umana e del contesto socio-culturale in cui si sviluppa. Ed è possibile realizzare artefatti di reale supporto a chi li utilizza solo adottando un approccio sperimentale centrato sull'essere umano, in cui la qualità dello strumento viene raggiunta dopo prove e verifiche effettuate con le metodologie che abbiamo illustrato in questo capitolo.

0.9. Riferimenti bibliografici

- Annett, J., Duncan, K.D. Stammers, R.B. Gray, M.J. (1971) Tas Analysis, Training Information Number 6, HMSO, London.
- Arbib, M. A. (1989). The Metaphorical Brain 2. New York: John Wiley & Sons.
- Bagnara S., Rizzo A. (1989) A methodology for the analysis of error processes in human-computer interaction. In Work with Computers: Organizational, Management and Health Aspects, eds M.J.
- Smith and G. Salvendy, Amsterdam: Elsevier Science, pp. 605-612.
- Beyer, H. Holtzblatt, K. (1997) Contextual Design. Morgan Kaufmann Publishers, Inc., San Francisco.
- Card, S.K. Moran, T.P. Newell, A. (1983) The Psychology of Human Computer Interaction, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carroll, J.M. (1995) Scenario-based Design, New York: Wiley & Sons.
- D'Andrade R. (1989). Culturally based reasoning. In Cognitions and social worlds, eds A. Gellantly and D. Rogers Oxford, Eng.: Oxford University Press.
- Dumas J.S., Redish J.C. (1993). A practical guide to usability testing. Norwood, NJ: Ablex.
- Erdmann, R.L. Neal, A.S. (1971) Laboratory vs. field experimentation in human factors. An evaluation of an experimental self-service airline ticket vendor. Human Factors, 13, pp 521-531.
- Giroto V., Light P., Colbourn C.J. (1988) Pragmatic schemas and conditional reasoning in children. Quarterly Journal of Experimental Psychology, 40, pp. 342-357.
- Gould, J.D. Lewis, C.H. (1983) Designing for usability key principles and what designers think. In Proceedings of the 1983 Computer Human Interaction Conference, 50-53.
- Hewett, T. Baecker, R. Cad. S. Carey, T. Gasen, J. Mantei, M. Perlman, G. Strong, G. Verplank, W. (1992) ACM SIGCHI Curricula for Human-Computer interaction, Report of ACM SIGCHI Curriculum Developm.G.
- Hooper, Woolsey K., Scott, K. Gayle C. (1996) Vizability PWS Publishing Company, Boston, MA.
- Houde, S, Hill, C. (1997) What Do Prototypes Prototype?, in Handbook of Human Computer Interaction (2nd Ed.), M. Helander, T. Landauer, and P. Prabhu (eds.): Amsterdam: Elsevier Science B.V, 1997.
- Hutchins, E. (1995). Cognition in the wild. MIT Press, London.
- Hutchins E.L., Hollan J.D. and Norman, D.A. (1985). Direct manipulation interfaces. Human-Computer Interaction, 1, 311-338.
- Johnson-Laird P.N., Legrenzi P., Legrenzi M. (1972). Reasoning and a sense of reality. British Journal of Psychology, 16, 243-275.
- Kirwan, B., Ainsworth, L.K., (1992) *A Guide to Task Analysis*. London: Taylor & Francis.
- Liddle D. (1996). Design of the conceptual model. In Winograd T. (Ed.) Bringing design to software. New York: Addison-Wesley.
- Marti, P. Rizzo, A. Petroni, L. Tozzi, G. Diligenti, M. (1999) Adapting the museum: a non-intrusive user modeling approach, Proceedings of UM99, Canada.
- Marti, P. (2000) "The choice of the unit of analysis for modelling real work settings", *Journal of Cognition, Technology and Work*, 2:62-74.
- Marti, P., Gabrielli, L. Pucci, F. (2000) Situated Interaction in Art, Personal Technologies 5/1.
- Middleton, D. Engestrom, Y. (1996) Cognition and Communication at Work Cambridge: Cambridge University Press.
- Neisser, U. (1976). Cognition and Reality. San Francisco: Freeman & Co.
- Nielsen, J. Mack, R.L. (1994) Usability Inspection Methods. New York: John Wiley & Sons.
- Nielsen, J., Molich, R. (1990). Heuristic evaluation of user interfaces. Proceedings of CHI 90, 249- 256. New York, NY: ACM.
- Norman, D. A. (1988). The psychology of everyday things. New York: Basic Books.
- Norman, D.A. (1993). Things that make us smart: Tools. Reading, MA: Addison Wesley.
- Norman, D.A. and Shallice, T. (1980). Attention to action: Willed and automatic control of behavior. (Tech. Rep. 8006). San Diego: University of California, San Diego.
- Norman, D. A. (1999). Il computer invisibile. Milano: Apogeo.
- Polson, P.G., Lewis, C. (1990) Theory-based design for easily learned interfaces. Human-Computer Interface 5, 191-220,
- Preece, J. Rogers, J. Sharp, H. Benyon, D. Holland, H. Carey, T. (1994) Human-Computer Interaction. Reading, MA: Addison-Wesley.
- Rizzo, A., Parlangeli, O., Cambiganu, C. Bagnara S. (1993) Complex systems by situated knowledge: implicit learning. In HCI 19b: ed. M.J. Smith G. Salvendy. Amsterdam: Elsevier Science.
- Rizzo, A., Parlangeli, O. (1994) Learning strategies and situated knowledge. Behavioral and Brain Sciences. 17, 420-421.
- Rizzo, A. Marchigiani, E. Andreadis, A. (1998) The AVANTI Project: Prototyping and evaluation with a Cognitive Walkthrough based on the Norman's model of action Proceedings of DIS, Amsterdam, 1998.
- Rizzo, A. (1999) Il pensiero nella vita quotidiana. In V. Giroto, P. Legrenzi (A cura di) Psicologia del Pensiero. Bologna: Il Mulino, pp. 203-226.
- Rizzo, A, Pasquini, A., Di Nucci P. & Bagnara, S. (2000). SHELFs: Managing critical issues through experience feedback. *Human Factors and Ergonomics in Manufacturing*, 10,1, 83-98.
- Shneiderman, B. (1998) Designing the User Interface. 3rd Edition. Reading MA: Addison-Wesley.
- Shneiderman B. (1992). Designing the User interface. Reading MA: Addison Wesley.
- Suchman, L.A. Trigg, R.H. (1991) Video as a Medium for Reflection and Design. In Greenbaum and Kyng (eds) Cooperative Design of Computer Systems. Hillsdale, New Jersey: Lawrence Erlbaum, pp 65-89.
- Vygotskij (1974). Storia dello sviluppo delle funzioni psichiche superiori e altri scritti. Firenze: Giunti Barbera.
- Wharton, C., Rieman, J., Lewis, C., and Polson, P. (1994) The cognitive walkthrough: A practitioner's guide. In J. Nielsen and R.L. Mack (Eds.), Usability Inspection Methods. John Wiley & Sons, 105-140.
- Zhang, J., and Norman, D. A. (1994). Representations in Distributed Cognitive Tasks. Cognitive Science 18: 87-122.

Lettere consigliate

- Clayton L e Rieman J: *Task-Centred User Interface Design: A Practical Introduction*. Boulder Colorado: University of Colorado, 1993. ftp.cs.colorado.edu/pub/cs/distribs/clewis/HCI-Design-Book
- Newman W M e Lamming M G: *Interactive System Design*. Addison Wesley, Workingham, 1995.

1

capitolo



1.0. Chi: computer - human interaction

Human-computer interaction (HCI) or, alternatively, **computer-human interaction** (symbolized as **X χ Chi**, the 22nd letter of the Greek alphabet) is the study of interaction between people (users) and computers. It is an interdisciplinary subject, relating computer science with many other fields of study and research. Interaction between users and computers occurs at the user interface (or simply *interface*), which includes both software and hardware, for example, general purpose computer peripherals and large-scale mechanical systems such as aircraft and power plants.

Interdisciplinary aspects

Combined with computer science and information technology are fields including:

- Aesthetics
- Anthropology
- Artificial intelligence
- Cognitive science
- Computer vision
- Design
- Ergonomics
- Library and information science
- Philosophy
- Phenomenology
- Psychology
- Social psychology
- Sociology
- Speech-language pathology

A basic goal of HCI is to improve the interaction between users and computers by making computers more user-friendly and receptive to the user's needs. Specifically, HCI is concerned with

- methodologies and processes for designing interfaces (i.e., given a task and a class of users, design the best possible interface within given constraints, optimizing for a desired property such as learnability or efficiency of use)
- methods for implementing interfaces (e.g. software toolkits and libraries; efficient algorithms)
- techniques for evaluating and comparing interfaces
- developing new interfaces and interaction techniques
- developing descriptive and predictive models and theories of interaction

A long term goal of HCI is to design systems that minimize the barrier between the human's cognitive model of what they want to accomplish and the computer's understanding of the user's task (see CSCW). Professional practitioners in HCI are usually designers concerned with the practical application of design methodologies to real-world problems. Their work often revolves around designing graphical user interfaces and web interfaces. Researchers in HCI are interested in developing new design methodologies, experimenting with new hardware devices, prototyping new software systems, exploring new paradigms for interaction, and developing models and theories of interaction.

Terminology

- **HCI vs CHI.** The acronym CHI (pronounced *kaɪ*), for computer-human interaction, has been used to refer to this field, perhaps more frequently in the past than now. However, researchers and practitioners now refer to their field of study as HCI (pronounced as an initialism), which perhaps rose in popularity partly because of the notion that the human, and the human's needs and time, should be considered first, and are more important than the machine's. This notion became increasingly relevant towards the end of the 20th century

as computers became increasingly inexpensive (as did CPU time), small, and powerful. Since the turn of the millennium, the field of human-centered computing has emerged as an even more pronounced focus on understanding human beings as actors within socio-technical systems.

- **Usability vs Usefulness.** Design methodologies in HCI aim to create user interfaces that are usable, i.e. that can be operated with ease and efficiency. However, an even more basic requirement is that the user interface be *useful*, i.e. that it allow the user to complete relevant tasks.
- **Intuitive and Natural.** Software products are often touted by marketers as being "intuitive" and "natural" to use, often simply because they have a graphical user interface. Many researchers in HCI view such claims as unfounded (e.g. a poorly designed GUI may be very unusable), and some object to the use of the words *intuitive* and *natural* as vague and/or misleading, since these are very context-dependent terms.
- **Data Density and Information Absorption.** The rapid growth in the density of computer screen real estate has created an opportunity to accelerate "information absorption" to much higher levels. Classic "data density" on a computer is 50-100 data points, recent advances in data visualization enable thousands of data points to be presented in forms which can be rapidly absorbed. Interfaces such as virtual reality will give further growth to the potential density of information presented.

Design methodologies

A number of diverse methodologies outlining techniques for human-computer interaction design have emerged since the rise of the field in the 1980s. Most design methodologies stem from a model for how users, designers, and technical systems interact. Early methodologies, for example, treated users' cognitive processes as predictable and quantifiable and encouraged design practitioners to look to cognitive science results in areas such as memory and attention when designing user interfaces. Modern models tend to focus on a constant feedback and conversation between users, designers, and engineers and push for technical systems to be wrapped around the types of experiences users want to have, rather than wrapping user experience around a completed system. **User-centered design (UCD)** is a modern, widely practiced design philosophy rooted in the idea that users must take center-stage in the design of any computer system. Users, designers, and technical practitioners work together to articulate the wants, needs, and limitations of the user and create a system that addresses these elements. Often, user-centered design projects are informed by ethnographic studies of the environments in which users will be interacting with the system.

Academic conferences

One of the top academic conferences for new research in human-computer interaction, especially within computer science, is the annually held ACM's Conference on Human Factors in Computing Systems, usually referred to by its short name CHI (pronounced *kai*, or *khai*). CHI is organized by ACM SIGCHI Special Interest Group on Computer-Human Interaction. CHI is a large, highly competitive conference, with thousands of attendants, and is quite broad in scope.

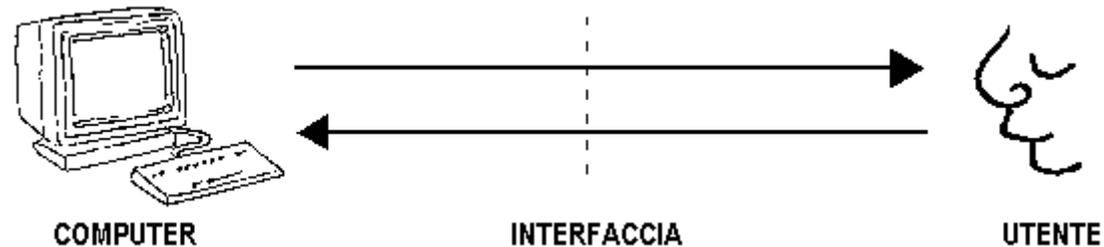
1.1. Una sfida per il designer

La tecnologia dei computer pone una nuova sfida al *designer industriale*: il software consente al progettista una libertà e una flessibilità del tutto sconosciute a chi utilizza materiali e tecnologie tradizionali per realizzare *artefatti*. Con i computer, il nostro mondo si popola di *artefatti interattivi*; in futuro, forse di *artefatti intelligenti*. Questo pone problemi del tutto nuovi, di natura filosofico/concettuale e metodologico/progettuale, sia al progettista di software che al designer industriale. Queste note, scritte da un informatico, si propongono di fornire un inquadramento di natura concettuale a questi problemi, approcciandoli dal punto di vista di chi progetta sistemi software.

Dopo una discussione introduttiva sui temi principali del nostro discorso (Sezione 1), nella Sezione 2 approfondiremo ulteriormente il concetto di interattività; nella Sezione 3 richiameremo i principali paradigmi di interazione uomo-computer che si sono consolidati fino ad oggi, e i nuovi paradigmi emergenti in relazione alle nuove tecnologie di interazione; nella Sezione 4 discuteremo le difficoltà nel fornire principi e linee guida che possano supportare il designer della sua attività di progettazione. Seguiranno infine, nella Sezione 5, alcune conclusioni.

Il computer come meta-medium o come meta-strumento

Per introdurre i termini fondamentali del nostro discorso, riferiamoci alla figura seguente. In essa vediamo una coppia di "sistemi" che interagiscono fra loro: un uomo e un computer.



Possiamo considerare questo secondo sistema da due punti di vista complementari, ma abbastanza diversi. In primo luogo, lo possiamo considerare essenzialmente come un *medium*, anzi come un "meta-medium", capace di simulare tutti gli altri media:

Sebbene i calcolatori digitali siano stati originariamente progettati per effettuare calcoli aritmetici, la capacità di simulare i dettagli di un qualsiasi modello descrittivo significa che il calcolatore, considerato come un medium in sè, può essere tutti gli altri media, a condizione che vengano adeguatamente forniti metodi di inserimento ["embedding"] e di visualizzazione. Inoltre, questo nuovo "meta-medium" è attivo - esso può rispondere a interrogazioni e ad esperimenti - pertanto i messaggi possono coinvolgere il discente in una conversazione nei due sensi. Questa proprietà non era mai stata disponibile prima, tranne che con il medium dell'insegnante individuale. Noi riteniamo che le implicazioni siano vaste e stimolanti. (Kay & Goldberg, 1977)

Il secondo punto di vista è quello strumentale: secondo questo punto di vista, possiamo considerare il computer essenzialmente come uno *strumento*. Anzi, come un "meta-strumento": ancora una volta, progettando del software opportuno, possiamo realizzare col computer una molteplicità di strumenti diversi. Secondo una immagine celeberrima di Seymour Papert (1980):

Il calcolatore è il Proteo delle macchine. La sua essenza è la universalità, il suo potere la capacità di simulazione. Poichè esso può assumere mille forme e assolvere a mille funzioni, esso può soddisfare mille gusti diversi.

Medium o strumento? Il designer di sistemi basati su computer può interpretare il suo lavoro in modo alquanto diverso, a seconda che ponga l'enfasi sul primo o sul secondo aspetto. Se consideriamo il computer principalmente come un nuovo medium, molto probabilmente porremo enfasi, nel progettare, soprattutto sulla struttura dell'informazione che attraverso questo medium presentiamo all'utente; se invece poniamo l'enfasi sugli aspetti strumentali, ci occuperemo forse, prevalentemente, di progettare le modalità attraverso cui questo utente può "accedere a" e "usare" queste informazioni.

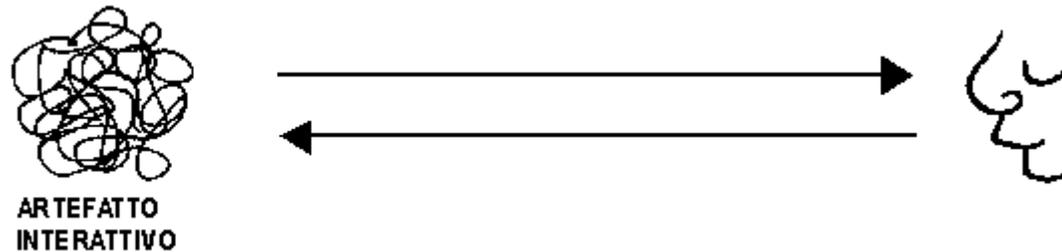
Nel primo caso il progettista sarà in qualche modo assimilabile a un *redattore* (questo è evidente, per esempio, nella progettazione di molti sistemi multimediali e ipermediali); nel secondo caso a un vero e proprio *industrial designer*: un progettista di nuovi artefatti e, in particolare, di nuovi strumenti. A questo nuovo tipo di designer, il software fornisce grandi possibilità e permette grande libertà creativa. Non più vincolato dalle leggi della meccanica, della statica, della elettricità, dalle caratteristiche dei materiali utilizzati, dai processi industriali di produzione, il nuovo designer ha come unico vincolo la necessità che il comportamento degli oggetti progettati possa essere descritto in modo algoritmico in qualche linguaggio di programmazione. In altre parole, i vincoli fisici vengono sostituiti da vincoli di natura esclusivamente logica: l'orizzonte delle possibilità non è più determinato dalle leggi della fisica, ma dalla pura intelligenza.

Possiamo così concepire in grande libertà artefatti del tutto nuovi, che interagiscano con il loro utilizzatore non solo sul piano fisico/meccanico (il coltello, il cacciavite, la seggiola, il televisore, l'automobile, ...) ma anche sul piano *cognitivo*:

La sfida è quella di creare sistemi che, attraverso una intima interazione cognitiva con l'utilizzatore, lo aiutino a strutturare e a manipolare le proprie idee. Con l'aiuto di tali sistemi, la gente sarà allora in grado di fronteggiare compiti intellettuali sempre più complessi di ogni tipo - comporre libri e presentazioni multi-mediali, progettare prodotti e programmi, comporre musica, analizzare esperimenti, gli argomenti della legge e della politica, ragionare di temi scientifici e sociali, a così via. La chiave per costruire tali sistemi è quella di trovare il modo per cui un utilizzatore possa operare sulle proprie idee come se fossero degli oggetti, così come gli attuali text editor gli permettono di agire sulle parole come se fossero oggetti. (Card & Moran, 1988)

Si modifica così il concetto stesso di "strumento": da "arnese per fare" ("*arnese atto all'esecuzione di determinate operazioni proprie di un'arte, di un mestiere, o di una tecnica*", recita il dizionario alla voce "strumento" ²) ad "arnese per pensare" ³.

Si noti che, in questi artefatti, il calcolatore, come presenza riconoscibile, scompare: Proteo multiforme, assume di volta in volta uno specifico aspetto (un foglio elettronico, un word processor, un sistema di progettazione, un computer game) e, di per sé, non è più visibile. Possiamo dire che, agli occhi del suo utente, un prodotto software divenga una sorta di "macchina virtuale" che si sovrappone e si sostituisce alla "macchina reale" costituita dal calcolatore ⁴. Noi, gli utenti, interagiamo in effetti, sul piano cognitivo, con questa macchina virtuale, e non con il device fisico (il video, la tastiera, il mouse, ...) con cui siamo fisicamente in contatto. Possiamo, pertanto, ridisegnare la figura precedente così:



E' di questa macchina virtuale che noi dobbiamo comprendere le regole di comportamento, le risposte, i malfunzionamenti. E' questa macchina virtuale, questo insieme di "apparenze attive", che il designer deve progettare. Questo concetto, ancora non bene assimilato dai progettisti di software, è ben espresso da Ted Nelson, che ancora nel 1980 (e quindi dopo 25 anni dall'invenzione dei computer) sentiva la necessità di precisarlo, con parole che ancora oggi è opportuno considerare con attenzione:

Un sistema interattivo basato su computer è una serie di presentazioni che hanno lo scopo di influenzare la mente in un certo modo, proprio come un film. Questa non è un'analogia casuale, ma l'aspetto centrale. Io uso il termine "virtuale" nel suo significato tradizionale, come l'opposto di "reale". La realtà di un film riguarda come lo scenario è stata dipinto e dove gli attori si sono spostati fra una ripresa e l'altra, ma a chi interessa? La virtualità di un film è ciò che sembra esserci dentro. La realtà di un sistema interattivo include le sue strutture dati e il linguaggio che è stato usato per programmarlo - ma, ancora, a chi interessa? Ciò che più importa è: che cosa esso sembra essere? (Nelson, 1980)

Un sistema interattivo, io credo, può essere considerato come un nuovo tipo di film; ma un film che possiamo controllare, e che riguarda qualcosa che desideriamo influenzare. Ed è nell'immaginare questo film interattivo che consiste il compito importante del progettista. Non preoccupiamoci dei dadi e delle viti: qual è il sogno?

Multimodalità e interattività

Ma ritorniamo alla nostra prima figura. Essa pone l'enfasi sull'aspetto centrale per il progettista: il rapporto - o, come si dice, *l'interfaccia* - fra utente e artefatto. A questo proposito vogliamo individuare due temi centrali - quasi due parole chiave - che il progettista dell'interfaccia deve considerare, e, sui quali ci soffermeremo alquanto nel seguito: **multi-modalità** e **interattività**.

Il primo tema fondamentale è quello delle diverse *modalità* con cui si può attuare il processo di comunicazione, nei due sensi indicati in figura: dall'uomo al computer e dal computer all'uomo. Noi abbiamo a disposizione numerose modalità comunicative possibili, che vanno scelte e utilizzate in funzione dei particolari scopi dell'artefatto che stiamo progettando. "Multi-modalità" significa allora che potremo sfruttare, *contemporaneamente*, canali di comunicazione differenti e complementari. Poiché l'uomo è dotato di cinque sensi, possiamo ovviamente auspicare che la comunicazione verso di esso sfrutti in modo appropriato, quando la tecnologia ce lo consentirà, tutti e cinque questi sensi. Per quanto riguarda invece la comunicazione nell'altra direzione, si tratterà di dotare il computer di opportuni "sensi" che gli permettano di acquisire informazioni da canali di comunicazione sempre più appropriati (con la evoluzione della tecnologia) alla struttura sia fisica che cognitiva dell'interlocutore umano ⁵. La visione estrema, a questo proposito, è quella dell'*uomo-terminale*, presente non solo nei romanzi *cyberpunk*, ma anche nei laboratori di ricerca sulle tecnologie di interazione:

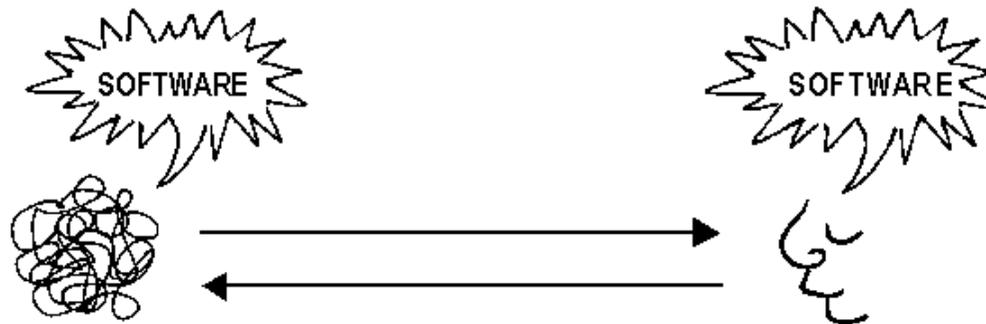
La persona è il vero terminale di ogni sistema informativo computerizzato. Questo terminale è già stato progettato. Noi possiamo solo progettare per esso, incorporando le capacità e le limitazioni umane come elementi espliciti del nostro pensare alla situazione complessiva dell'interfaccia (Bolt, 1984, pag.81)

Il secondo aspetto fondamentale è quello della *interattività*. La tecnologia del software, con la quale "modelliamo" il computer per fargli assumere la forma (e la funzione) desiderata, ci permette di creare con grande facilità degli oggetti interattivi. In realtà tutti gli oggetti con i quali noi possiamo venire a contatto sono, in un certo senso, interattivi, anche quelli che ci appaiono completamente passivi: un cacciavite, ad esempio, non sarebbe utilizzabile se la pressione che esso esercita sul palmo della mano non ci fornisse quel feedback che ci permette di controllarne con accuratezza il movimento. In questo senso, dunque, il cacciavite è un oggetto interattivo: la interazione, a partire dal livello sensoriale/percettivo, coinvolge processi cognitivi complessi, che ci permettono di effettuare l'azione di "avvitare". Numerosi altri oggetti di uso quotidiano possono essere considerati interattivi, per gli stessi motivi: le forbici, la macchina da scrivere, l'automobile e, su un piano diverso (in quanto in questo caso la interazione si svolge quasi esclusivamente a livello cognitivo) il libro. Non è difficile comprendere, tuttavia, che un artefatto realizzato mediante software può essere dotato di "interattività" (comunque la si voglia definire, e di questo tratteremo nel prossimo paragrafo) in un modo decisamente più marcato. Ed è questa la sfida che il progettista (non solo il progettista software, ma il progettista *tout court*) deve in qualche modo raccogliere:

Progettare l'interattività significa progettare nella quarta dimensione temporale: dare al tempo il valore di parametro organizzatore della qualità. Per la cultura del progetto, che ha sempre avuto come ruolo centrale quello di dare qualità a delle forme nello spazio, occuparsi dell'interattività significa dunque avventurarsi su un territorio del tutto nuovo, i cui riferimenti culturali vengono, casomai, dal cinema, dal teatro, dalla musica; da attività cioè in cui la qualità si produce appunto organizzando sequenze di eventi. (Manzini, 1990)

Design e scienze cognitive

Il progettista software e il designer industriale di oggi hanno di fronte a sé un compito nuovo: *progettare l'interattività*. Per svolgere bene questo compito è necessario disporre, da un lato, di una tecnologia appropriata; dall'altro, di principi ben fondati su cui basare delle *linee guida* che orientino il designer nel progetto di artefatti interattivi di qualità. A che punto siamo oggi? Dal punto di vista della tecnologia (che è nell'ambito del nostro discorso essenzialmente tecnologia software), abbiamo visto, in mezzo secolo di evoluzione, consolidarsi il modello dei linguaggi procedurali, e delinearsi il paradigma della programmazione ad oggetti. Dal punto di vista metodologico, da dove possiamo trarre i principi e le linee guida che orientino il progettista di artefatti interattivi? Approfondiamo ancora il nostro schema, e modifichiamolo nel modo seguente:



In questa immagine abbiamo posto in evidenza la presenza di *software* in *entrambi* i sistemi di cui desideriamo progettare la interazione. Nel caso dell'interlocutore umano, ovviamente, il riferimento alla presenza di software va considerato in senso metaforico, e inquadrato nelle in alcune tendenze della psicologia cognitiva, in cui i meccanismi della mente umana vengono studiati sulla base di modelli di tipo "information processing" fortemente ispirati alla scienza dei calcolatori:

L'avvento dei moderni calcolatori digitali ha fornito una ricca metafora teoretica per costruire teorie sulla elaborazione umana dell'informazione: le persone, come i calcolatori, acquisiscono informazioni dall'ambiente; sia le persone che i calcolatori immagazzinano informazioni e le recuperano quando esse sono applicabili ai compiti correnti; entrambi sono limitati nella quantità di informazioni che possono processare in un dato momento; entrambi trasformano informazioni per produrre nuove informazioni; entrambi restituiscono informazioni all'ambiente. L'architettura dei calcolatori ha fortemente influenzato molte elaborazioni iniziali della moderna psicologia cognitiva. I componenti delle teorie cognitive sono stati spesso sorprendentemente simili ai componenti dei calcolatori, come anche la loro organizzazione. (Barsalou, 1992)

Secondo tale modello, la figura precedente ci mostra due "Information Processor" fra loro interagenti, sia dal punto di vista fisico che - ed è questo che qui interessa - dal punto di vista cognitivo. Ciascuno di questi due Information Processor funziona secondo leggi proprie, in accordo alle strutture "hardware" e "software" che lo costituiscono, e che ne determinano pertanto le modalità di comportamento e i relativi parametri. Ciò che più conta, ai fini del nostro discorso, è che questi due Processor dovranno interagire in modo armonico: il progettista di artefatti interattivi non potrà non tenere conto, nella sua attività, dei meccanismi che regolano il funzionamento dell'Human Information Processor che con il suo artefatto dovrà interagire. Così come, in passato, gli studi sulla forma e sul colore hanno fornito indispensabili strumenti di comprensione al progettista grafico, e gli studi sulla percezione del movimento hanno fornito le basi metodologiche per la realizzazione dei cartoni animati e delle sceneggiature cinematografiche, i risultati della moderna psicologia cognitiva possono essere di grande aiuto nella progettazione di software che risulti "usabile" (Card et al., 1983).

Human Computer Interaction: una disciplina specifica

Fondare il design dell'interazione uomo-computer sulla conoscenza dei processi cognitivi umani che vengono coinvolti è programma ambizioso, ed ancora agli inizi. Si inquadra in una disciplina nata agli inizi degli anni 90, denominata Interazione Uomo-Calcolatore ("Human Computer Interaction", o semplicemente HCI) che appare oggi anche in Italia, come disciplina a se stante, nei curricula universitari ⁶:

La Interazione Uomo-Calcolatore è una disciplina interessata alla progettazione, alla valutazione e alla realizzazione di sistemi interattivi computerizzati destinati all'uso umano, e allo studio dei principali fenomeni correlati ⁷

La definizione riportata qui sopra è molto ampia, e impone chiaramente un approccio interdisciplinare:

L'HCI nella sua totalità è un'area interdisciplinare. Sta emergendo come una specializzazione nell'ambito di parecchie discipline, in ciascuna con enfasi differenti: scienza dei calcolatori (progettazione delle applicazioni e ingegneria delle interfacce umane), psicologia (l'applicazione delle teorie dei processi cognitivi e l'analisi empirica del comportamento degli utenti), sociologia e antropologia (le interazioni fra tecnologia, lavoro e organizzazione), e disegno industriale (i prodotti interattivi) ⁸

1.2. Interattività

A volte si deve estrarre un'espressione dal linguaggio, farla pulire, - e poi si può rimetterla in circolazione. L. Wittgenstein

Il termine "interattività" denota un concetto molto ampio. Troviamo nel dizionario:

***interattivo:** che è reciprocamente attivo, di due o più elementi che esercitano reciproca attività l'uno sull'altro*

***interazione:** azione, reazione, influenza reciproca di cause, fenomeni, forze, elementi, sostanze, agenti naturali, fisici, chimici, e, per estensione, psicologici e sociali⁹*

Queste definizioni sono troppo generiche per risultare utili nella nostra analisi. Meglio la definizione di A.Lippman, che studia i sistemi interattivi al Media Lab del MIT:

***Interattività.** Attività mutua e simultanea da parte di entrambi i partecipanti, che normalmente, ma non necessariamente, lavorano per un certo obiettivo¹⁰*

Secondo Lippman, questa definizione ha cinque importanti corollari, che definiscono cinque importanti attributi che devono essere posseduti da qualunque scambio comunicativo che possa correttamente dirsi interattivo. Vediamoli brevemente, facendoci aiutare dalle parole dello stesso Lippman¹¹:

1. Interrompibilità. Ciascuno deve essere in grado di interrompere l'altro, perchè stiamo cercando di distinguere fra ciò che è interattivo - che significa reciproco e simultaneo, da ciò che è alternato. [...] Il modello dell'interazione è la conversazione, non la lezione. Pertanto la domanda è: qual è la granularità del sistema interattivo - cioè, qual è il più piccolo elemento atomico al di sotto del quale non si può interrompere? Nella conversazione qual è? Può essere la parola o può essere la frase.

2. Degrado morbido. In ogni interazione è sempre possibile fare una richiesta a cui non si può rispondere. [...] Ma questo blocca forse l'interazione? La risposta è no, se il sistema degrada in modo morbido. Nel senso che può ripiegare su ciò che non può fare. Dirà: "Risponderò a ciò in un minuto".

3. Prospezione limitata. Non si può avere precalcolato tutto ciò che si sta per dire. In una conversazione, quanto in avanti rispetto a quello che si sta dicendo si pensa realmente? Abbiamo un obiettivo, ma poichè esso è interattivo, e ciascuno di noi interromperà l'altro, non possiamo nemmeno prevedere se lo raggiungeremo, o dove arriveremo, o di quanto ci allontaneremo, così che non guardiamo poi troppo in avanti nel costruire l'interazione.

4. Nessun default. Quando si costruisce un sistema, si vuole evitare di costruire un sistema che abbia una sorta di cammino assunto per default, da seguire a meno che non lo si cambi. Si vuole che ciò accada al volo poichè i default sono, non so, noiosi.

5. **Impressione di una base dati infinita.** Se in ogni momento ci sono solo due scelte possibili, l'esito delle quali dopo un po' diviene ovvio, allora neanche in questo caso è interattivo. Ti conquista perchè ci sono quindici cose che si possono fare ad ogni immagine, e ciascuna delle quali è diversa ¹² C'è l'apparenza di una base dati infinita. Si può toccare un edificio, entrarci dentro, senza sapere che cosa ci si troverà se non ci si è già stati prima. Se tutto quello che si può fare è guidare e svoltare, si sa che cosa farà ciascun input. Se diventa subito evidente che convergerà, allora non è più interattivo, non è vero?, perchè non c'è alcun valore nell'interrompere, perchè si sa dove si arriverà - è come se si possedesse una capacità di prospezione infinita. E' tutto precalcolato. Se nulla di ciò che si può fare può cambiare le cose, qual è lo scopo dell'interazione?

Esistono sistemi interattivi?

Il modello che sta alla base di questa visione della interattività è, ovviamente, quello della conversazione fra due interlocutori umani. La immagine che la discussione sopra riportata evoca è quella di una conversazione, per così dire, erratica ed imprevedibile, che prosegue senza inceppi - ma anche senza una disciplina imposta a priori - sviluppando in egual modo gli spunti proposti da entrambi gli interlocutori. L'esempio che più facilmente viene alla mente, qui, è Eliza, il programma che "fingeva di essere uno psichiatra", che tanto fece discutere i ricercatori dell'intelligenza artificiale (Weizenbaum, 1966). Ecco un esempio ¹³, in cui il testo digitato dall'interlocutore umano è in neretto, mentre le risposte di Eliza sono in chiaro:

Good morning my name is Gaines and I wish to talk to you

I am not interested in names

Sorry perhaps you could let me know your interests

Please don't apologize

Thank you. What interest you most

Why do you ask

Because I have to demonstrate your ineffable capabilities

Is that the real reason

Yes your capabilities interest me

You seem quite positive

I am indeed positive

Is it because you are indeed positive that you came to me

No I came to demonstrate your capabilities

Are you saying 'no' just to be negative

Yes

You are being somewhat short with me

I am very annoyed at your lack of conversational grace

How long have you been annoyed at my lack of conversational grace

For several minutes now

I am not sure I understand you fully

Can you tell the time

You believe I can tell the time don't you

....¹⁴

Oggi, un quarto di secolo dopo gli esperimenti di Weizenbaum, sappiamo non solo che Eliza era sostanzialmente uno scherzo, anche se con grandi implicazioni filosofiche, ma anche che un dialogo realmente interattivo con un computer, nell'accezione sopra discussa, è eccezionalmente difficile da realizzare, ed ancora ben al di là dello stato dell'arte. Sicuramente, i sistemi software finora progettati hanno sfruttato le possibilità di una comunicazione realmente interattiva in misura molto limitata:

Ci sono motivi per sostenere che non esista alcun sistema realmente interattivo - con l'eccezione, forse, dei pacemaker e dei defibrillatori cardiaci - poiché la vera interazione implica che l'utente risponda al sistema almeno con la stessa frequenza con cui il sistema risponde all'utente e, ciò che è ancora più importante, che le iniziative prese dall'utente o dal sistema alterino il comportamento dell'altro. I video game dei giorni gloriosi della rinascita dell'Atari, o il Flight Simulator e i suoi cloni, o certi sistemi operativi ci offrono almeno una vaga idea di ciò che l'interazione potrebbe essere (Joyce, 1991)

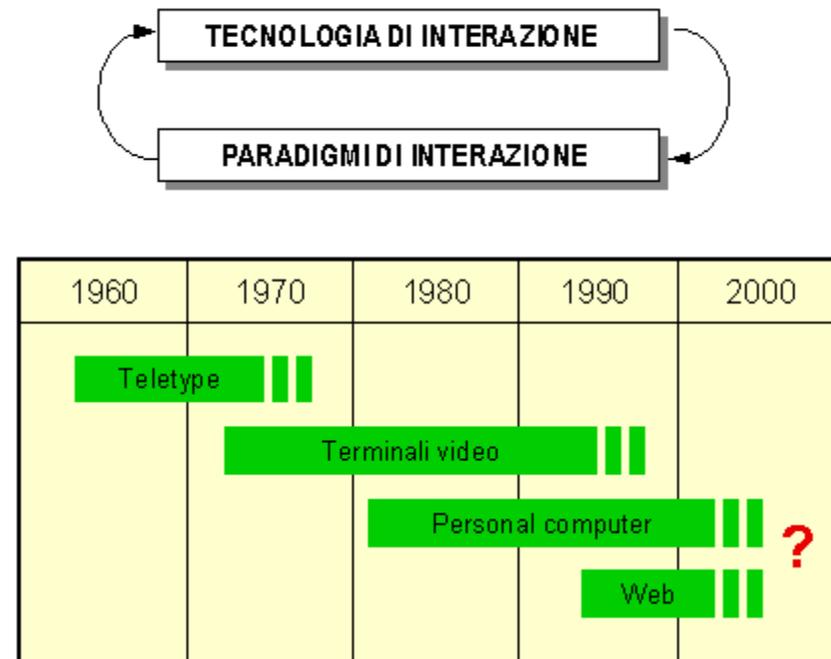
Per verificare questa affermazione, e per mostrare quali possibilità di interazione siano state a tutt'oggi sfruttate dai progettisti software, presentiamo, nella prossima sezione, una sintetica panoramica della evoluzione dei paradigmi di interazione che si sono consolidati nel primo mezzo secolo di storia dei computer.

1.3. Evoluzione dei paradigmi di interazione uomo-computer

Il ciclo "compito-artefatto"

Durante questo primo mezzo secolo di storia della elaborazione elettronica, la interazione fra uomo e calcolatore si è evoluta secondo un certo numero di paradigmi diversi, che in questo paragrafo vogliamo ripercorrere brevemente. Come filo conduttore sceglieremo la evoluzione della tecnologia di interazione col calcolatore: i differenti device che si sono di volta in volta maggiormente diffusi hanno, in effetti, avuto un ruolo determinante nel suggerire modelli di comunicazione differenti (e, a loro volta, ne sono stati influenzati, in un ciclo di retroazione fra strumento e compito che porta a una continua evoluzione delle modalità di interazione uomo-macchina:

Dovendo schematizzare, prenderemo come riferimento quattro tecnologie consolidate di interazione uomo-calcolatore, la cui evoluzione nel tempo è indicata approssimativamente nella figura precedente: il terminale scrivente, o teletype; il terminale video; il personal computer; il web. Si noti che alcune delle tecnologie individuate sono sostanzialmente, dal punto di vista funzionale, un super-set delle precedenti, nel senso che possono "simularne" le funzioni: così, ad esempio, con un terminale video posso simulare un terminale scrivente, e con un personal computer posso simulare un terminale video.



Il terminale scrivente: scrivi e leggi

Il terminale scrivente era essenzialmente un apparato composto da tastiera e da stampante integrata (tipicamente, a foglio continuo). A confronto con i terminali sviluppati successivamente, le prestazioni erano molto modeste, sia per quanto riguarda la velocità di stampa (ad esempio, 30 caratteri al secondo) che per quanto riguarda la velocità di trasmissione lungo la linea verso il calcolatore (ad esempio, qualche diecina di caratteri al secondo). Con la mediazione di un tale apparato, con il calcolatore si dialoga per iscritto; il paradigma di comunicazione che si è inizialmente diffuso è pertanto quello della comunicazione scritta: secondo questo paradigma, tipicamente, il calcolatore segnala all'utente il suo stato di "attesa comandi"; l'utente digita quindi il comando, a cui segue la risposta dell'elaboratore e il sollecito successivo. Questa è la modalità comunicativa di molti command language (come ad esempio quello tradizionale dei sistemi Unix e successivamente MS-DOS), o dei query language per l'interrogazione di basi di dati (come ad esempio il linguaggio SQL) o, ancora, di molti adventure game della prima generazione, di tipo testuale.

In tutti questi esempi, è sempre l'utente che ha il controllo del dialogo: il calcolatore ha un ruolo passivo, limitandosi essenzialmente a fornire le risposte richieste. Esistono tuttavia, anche se in numero più ridotto, sistemi in cui questi ruoli sono invertiti, e il dialogo è totalmente controllato dal calcolatore: è l'utente ad avere un ruolo passivo, limitandosi a fornire le risposte richieste. Esempi tipici sono i sistemi esperti, gli advisory system, e in generale tutti quei sistemi in cui è il calcolatore ad avere competenza del problema e potere di decisione nella conduzione della conversazione. Paradigmi conversazionali veri e propri (come suggeriti dall'esempio di Eliza del paragrafo precedente o da qualsiasi conversazione fra due interlocutori umani che dialoghino "alla pari") sono comunque ancora ben oltre lo stato dell'arte, non solo per le enormi difficoltà tecniche insite nell'elaborazione del linguaggio naturale, ma anche per la grande complessità e ricchezza intrinseche alla nozione stessa di conversazione.

Il terminale video: scegli e compila

Con l'introduzione (a partire dal 1971) del terminale video, il tabulato continuo del terminale scrivente viene sostituito dallo schermo video (tipicamente di 24 linee di 80 caratteri): la velocità di visualizzazione di una singola schermata è praticamente istantanea, e quella di trasmissione sulla linea verso il calcolatore aumenta in modo considerevole (tipicamente, tra 150 e 1200 caratteri al secondo). La tastiera si arricchisce di svariati tasti funzionali (che attivano funzioni realizzate dal software o eseguite direttamente dal terminale). Altro aspetto innovativo determinante è la presenza, sul video, di un cursore spostabile in ogni direzione mediante gli appositi tasti. Questo permette all'utente di "indicare" una posizione precisa del video: il cursore è una sorta di "pennino" che "deposita sul video" il carattere digitato alla tastiera, o un "dito" con cui indicare l'elemento informativo di interesse (un carattere, un campo dati, una voce di menù, ...). La possibilità di visualizzare rapidamente sul video l'informazione che si sta manipolando, e di indicare col cursore la specifica porzione di interesse, suggeriscono lo sviluppo di modalità di comunicazione completamente diverse da quelle precedentemente citate. Ad esempio, negli editor full-screen, l'utente che crea o corregge un testo non opera più su una pagina "invisibile" e pertanto soltanto "pensata", come avveniva in precedenza nei line editor, ma su una pagina di testo visibile sullo schermo, indicando col cursore il punto desiderato (ad esempio, il punto in cui inserire altro testo, o il brano da cancellare).

I sistemi software di questa generazione propongono, tipicamente, un'interfaccia utente basata su una duplice metafora: la metafora del menù ("non scriverlo, scegli") e quella della modulistica d'ufficio (in gergo: "maschera"), da compilare spostando il cursore sui vari campi di input. A questa classe di sistemi appartiene la maggior parte delle applicazioni di tipo gestionale sviluppate a partire dagli anni settanta per oltre un quarto di secolo: la compilazione, campo a campo, di una maschera visualizzata sullo schermo diviene il paradigma standard per le applicazioni di tipo transazionale (sistemi informativi aziendali). Per evidenziare meglio la struttura delle maschere video (differenziando i campi variabili rispetto al testo fisso, separando le varie zone della maschera, ecc.) i terminali video si arricchiscono negli anni di funzioni grafiche, anche se elementari:

testo in doppia intensità, sottolineato o in reverse, caratteri grafici con cui comporre cornici o zonature del video, ecc. Una molteplicità di tasti funzionali specializzati permette una interazione scorrevole e efficiente.

Il secondo modello tipico di interazione è quello basato su menù: su più linee del video vengono proposte le varie alternative possibili, che l'utente può selezionare con semplici operazioni alla tastiera (ad esempio, digitandone il numero d'ordine). Molto spesso un'alternativa richiamerà un menù di livello inferiore, e così via, fino ad arrivare alle maschere della transazione selezionata. Introducendo una nuova dimensione "gestuale" (l'atto di indicare sul video qualcosa con il cursore), il terminale video introduce, sia pure in embrione, un paradigma nuovo, che si svilupperà appieno con i sistemi dotati di mouse: quello della manipolazione diretta (Shneiderman, 1983). Dal punto di vista della qualità dell'interazione, il paradigma "scegli e compila" semplifica molto la natura del dialogo uomo-computer che, se da un lato guadagna semplicità (le scelte possibili sono ben visibili sullo schermo), dall'altro perde ricchezza (le scelte possibili sono solo quelle visibili sullo schermo). Con menù e maschere l'utente è completamente guidato, ma il computer assume un ruolo alquanto più passivo che in precedenza: da soggetto di un dialogo (che almeno metaforicamente, e con molta generosità, poteva ricordare il dialogo fra due interlocutori umani - si pensi all'immagine del "cervello elettronico", ricorrente negli anni '50 e '60), diviene ora puro oggetto (sia pure "responsivo") di manipolazione, più o meno come un telefonata lavatrice, un televisore.

Il personal computer: non dirlo, fallo!

Nel personal computer, al video e alla tastiera si aggiungono potenza di calcolo locale, possibilità di archiviazione su dischetti flessibili asportabili o su dischi interni e possibilità di stampa locale.. Dopo i primi home computer, nati alla fine degli anni settanta e destinati essenzialmente al mercato degli amatori o al nuovo incerto mercato dell'home computing, dal 1981 il personal computer entra nel mondo delle aziende con il PC IBM che si afferma subito come standard di fatto al quale quasi tutti i costruttori si allineano proponendo macchine compatibili. Con il personal computer l'elaboratore esce dal centro edp e arriva sulla scrivania di utenti non informatici: segretarie, professionisti, dirigenti d'azienda, studenti. Dal punto di vista del design dell'interazione, nasce una nuova era. Nell'arco di pochi anni vengono realizzate e diffuse applicazioni software innovative, concepite per un mercato di largo consumo, che non hanno eguali, nei prodotti dell'edp tradizionale, per facilità d'uso e attenzione complessiva agli aspetti di interfaccia utente.

Rispetto al terminale video, due sono le grandi novità che modificano drasticamente la qualità dell'interazione. Da un lato, la potenza di calcolo locale permette al computer di reagire agli stimoli dell'utente con tempi di risposta pressochè immediati, e comunque sensibilmente inferiori a quelli dei terminali tradizionali, relativamente lenti per la presenza di linee di comunicazione (ancora relativamente lente) con l'elaboratore remoto. Dall'altro lato, la possibilità di archiviazione locale cambia profondamente il rapporto dell'utente con i dati gestiti dal calcolatore. Questo, da depositario di dati conservati in archivi centrali e accessibili soltanto per il suo tramite, diviene piuttosto uno strumento di manipolazione dei dati, che l'utente conserva direttamente, memorizzati su dischetti asportabili. Il personal computer introduce una forte discontinuità con il passato, ed avvia un processo di radicale trasformazione degli strumenti informatici che proseguirà fino agli anni 90. Dal punto di vista della interazione uomo-computer, il paradigma più significativo è probabilmente quello del foglio elettronico, il prodotto software per personal computer di gran lunga più diffuso.

Il foglio elettronico porta alle estreme conseguenze la possibilità di una interazione estremamente rapida con il calcolatore. Utente e computer operano alternativamente su una tabella visibile sullo schermo, contenente dati legati fra loro da relazioni matematiche, anche numerose e complesse. Ogniqualvolta l'utente aggiorna un valore in tabella, il programma, in modo pressoché istantaneo, aggiornerà tutti i valori ad esso correlati, permettendo all'utente di verificare le conseguenze delle proprie variazioni (in un bilancio, in un budget, in un piano di vendite, ...), e di modificare nuovamente, di conseguenza, i valori forniti. Il computer diviene uno strumento flessibile di simulazione ("che cosa succederebbe se..."),

in un rapporto interattivo con il suo utente sconosciuto in passato. Per non rallentare queste manipolazioni, i menù e i dati (che nelle applicazioni transazionali classiche apparivano in momenti diversi sul video) vengono ora collocati assieme in un'unica schermata, che contiene sia la tabella del foglio elettronico che un menù orizzontale, ridotto a una o due righe dello schermo (menu bar) ¹⁵.

La successiva evoluzione delle tecnologie di interazione è costituita dalla workstation un computer personale dotato di elevata potenza di calcolo, video grafico di buona risoluzione, tastiera e mouse. Quest'ultimo, inventato a metà degli anni 60 (English et al., 1967), apparirà nelle prime workstation all'inizio degli anni 70 e diventerà, dall'inizio degli anni 90, il corredo standard di ogni personal computer. L'utente, spostando il mouse sul piano della scrivania, è in grado di muovere un puntatore sul video, con grande precisione e rapidità. Il mouse è dotato, a seconda dei modelli, di uno, due o tre pulsanti, premendo i quali è possibile comunicare al calcolatore informazioni aggiuntive. E', in sostanza, come se l'utente indicasse un oggetto sul video e dicesse, contemporaneamente, "questo!". Il mouse introduce possibilità di interazione completamente nuove rispetto alla semplice tastiera, permettendo di comunicare con il calcolatore, per così dire, a gesti. Nonostante la sua semplicità, esso permette infatti una buona varietà di azioni: puntare ("pointing"), far click ("clicking"), far click due volte ("double-clicking"), premere ("pressing"), trascinare ("dragging"), e così via. La disponibilità di un video grafico a buona risoluzione, su cui rappresentare oggetti con accurata resa grafica, e di un mouse con cui indicare, selezionare, trascinare questi oggetti qua e là per lo schermo, suggeriscono allora un paradigma di interazione del tutto nuovo, che possiamo sintetizzare con lo slogan: non dirlo, fallo! Si tratta della manipolazione diretta, da parte degli utenti, degli oggetti rappresentati graficamente sul video, che permette di eliminare - o almeno ridurre notevolmente - nella comunicazione fra uomo e calcolatore la mediazione del linguaggio scritto. Anche se questo nuovo paradigma può essere sviluppato in molti modi diversi, fino ad oggi è stato principalmente utilizzato per la realizzazione di sistemi di tipo desktop. La metafora della scrivania, proposta nelle esperienze pionieristiche effettuate in ambito Xerox a partire dai primi anni '70, su workstation molto costose e quindi a bassa diffusione, sarà quindi adottata in personal computer a basso costo e a larga diffusione dalla Apple con il Macintosh (a partire dal 1984), e successivamente adottata dalla Microsoft nelle varie versioni di Windows, che diventeranno lo standard di fatto nei personal computer a partire dalla metà degli anni 90.

Il web: point & click

L'uso del mouse suggerisce naturalmente l'idea di presentare sul video dei "bottoni virtuali" sui quali operare premendo i bottoni (reali) del mouse. In effetti, il concetto di bottone (che può assumere varie foggie, o essere addirittura invisibile, come nel caso delle figure - o testi -sensibili) si è molto diffuso in questo tipo di interfacce. Si delinea così un nuovo paradigma di manipolazione diretta, in cui la interazione verso il computer è molto semplificata, riducendosi alla semplice pressione di bottoni puntati dal mouse: point&click. L'uso generalizzato del concetto di bottone suggerisce abbastanza naturalmente la evoluzione dalla metafora della scrivania alla metafora del pannello di controllo.: Secondo questa metafora, il video tende ad assomigliare al pannello di un'apparecchiatura elettronica, con pulsanti, interruttori, slider, spie luminose, display numerici, Esempi di questa tendenza sono numerosi, dal workspace della workstation NeXT ai vari simulatori, come il classico Flight Simulator della Microsoft, in cui il video mostra, fedelmente riprodotto nei dettagli, il quadro dei comandi di un aereo (indicatore della velocità dell'aria, altimetro, orizzonte artificiale, bussola, misuratore del carburante, pressione e temperatura dell'olio, ...).

Alla diffusione del paradigma "Point & click" contribuirà poi in modo determinante la diffusione esplosiva di Internet e, in particolare, del world wide web, a partire dalla metà degli anni 90. L'interfaccia proposta dai primi browser web (a partire da Mosaic) e tuttora prevalente si basa sulla nozione di ipertesto: la "navigazione" all'interno del web avviene cliccando testi o immagini sensibili, e attivando così dei link predefiniti che portano ad altre pagine. La interazione è ridotta a una successione di due azioni elementari: si indica col mouse un oggetto testuale o grafico presente sul video e lo si seleziona cliccando con il tasto sinistro del mouse stesso.

Il futuro

La prudenza e le esperienze del passato suggeriscono di non azzardare previsioni sull'evoluzione delle tecnologie. Come dice Schnaars, che ha dedicato un intero libro all'analisi degli errori di previsione:

Anche un rapido esame delle previsioni passate rivela che la maggioranza di quelle tecnologiche è stata totalmente sbagliata. In realtà quasi tutte hanno sbagliato il bersaglio, e non per pochi centimetri, ma completamente e senza mantenere alcun rapporto con quello che è accaduto realmente anni dopo.

Preferiremo quindi non prospettare ipotesi su quali tecnologie di interazione si contenderanno il mercato in questo secolo appena iniziato, né sui paradigmi che di conseguenza si consolideranno. Ci limiteremo, a conclusione di questa breve panoramica, a citare alcune importanti linee di tendenza in atto, che presumibilmente influenzeranno in modo significativo il design dell'interazione dei prossimi anni. La prima è quella verso la multimedialità. In tutte le tecnologie che abbiamo finora brevemente ricordato, il calcolatore elaborava sostanzialmente dati alfanumerici, testo, grafica. Le grandi capacità di elaborazione e archiviazione dei sistemi odierni e, soprattutto, la crescente disponibilità di connessioni Internet a larga banda permetteranno sempre più di utilizzare oggetti in formato audio e video che non potevano essere gestiti dai sistemi e dalle reti del passato. Con queste nuove possibilità si avvicina la possibilità di realizzare la famosa visione del Dynabook, con cui Alan Kay, una ventina d'anni fa, indentificava il personal computer del futuro:

Immaginate di avere il vostro personale manipolatore di conoscenza in confezione portatile, della dimensione e della forma di un normale blocco da appunti. Supponete che esso abbia abbastanza potenza da superare i vostri sensi della vista e dell'udito, abbastanza capacità da immagazzinare, per un successivo reperimento, migliaia di pagine-equivalenti di materiale di riferimento, di poemi, lettere, ricette, registrazioni, disegni, animazioni, spartiti musicali, forme d'onda, simulazioni dinamiche, e qualunque cosa voi possiate desiderare di ricordare e modificare. Pensiamo a un apparecchio piccolo e portatile per quanto è possibile, in grado sia di ricevere che di fornire informazioni in quantità paragonabili a quelle del sistema sensorio umano. L'output visivo dovrebbe essere, almeno, di qualità migliore di quella ottenibile dai giornali. L'output uditivo dovrebbe avere un simile standard di alta fedeltà. Non ci dovrebbe essere alcuna pausa percepibile fra causa ed effetto. Una delle metafore che abbiamo usato per progettare tale sistema è quella di uno strumento musicale, come un flauto, che è posseduto dal suo utilizzatore e risponde in modo istantaneo e consistente ai desideri del suo proprietario. Immaginate quanto sarebbe assurdo il ritardo di un secondo fra l'atto di soffiare una nota e il suo ascolto! (Kay & Goldberg, 1977)

Quali paradigmi di interazione si consolideranno? Una metafora emergente è quella dell'agente: un segretario specializzato e personale "dentro" il computer, che obbedisce ai nostri comandi e controlla per noi le vaste risorse di questi sistemi. Con questo segretario, in genere pensato con sembianze umane, si comunica a voce, in linguaggio naturale. Il paradigma "parla e ascolta" è stato da sempre proposto dagli scrittori di fantascienza nelle loro visioni di tecnologie future: ricordiamo tutti HAL, il computer di bordo della nave spaziale in viaggio verso Giove del film "2001: Odissea nello spazio" (1968) tratto dal romanzo di Arthur C. Clarke, che comunica con l'equipaggio in un inglese perfetto, pronunziato con voce soave anche se un po' metallica. La ricerca sulla comunicazione vocale è attiva da vari decenni, ed ha prodotto tecnologie di sintesi e riconoscimento vocale che, anche se non ancora largamente diffuse sul mercato, sono ormai di buon livello (sistemi di dettatura, risponditori vocali automatici, conversioni text-to-speech). E' pertanto presumibile che, nei prossimi anni, questo paradigma di interazione, che costituisce il paradigma principale nella comunicazione uomo-uomo, si diffonderà in modo significativo anche nella comunicazione uomo-computer.

I primi sistemi di realtà virtuale hanno origine dalle esperienze effettuate alla fine degli anni 80 nei laboratori di ricerca della NASA (Fisher, 1986). Con questa tecnologia, più ancora di quanto avveniva con i tradizionali simulatori di volo (disponibili fin dagli anni '70), l'uomo si trova completamente "immerso" in un mondo artificiale, prodotto dal calcolatore con tecniche di computer graphics, nel quale può muoversi e con il quale può interagire direttamente. Il paradigma della manipolazione diretta viene qui portato alle estreme conseguenze, "immergendo totalmente" l'utente in una realtà artificiale, con una interfaccia multisensoriale: entra & agisci. La strumentazione di base consiste in un visore con due piccoli display collegati al computer, uno per ciascun occhio, per mostrare una immagine tridimensionale. Il visore impedisce a chi lo indossa di vedere altro, generando una sensazione di immersione totale nella realtà artificiale. Un sensore montato sul visore segnala al computer la posizione e l'orientamento della testa dell'utente: quando questi muove la testa, la scena visualizzata sui due display ruota di conseguenza (il sistema permette di esplorare un campo visivo di 360 gradi). Per interagire con gli oggetti della realtà artificiale, l'utente indossa uno speciale guanto: così, egli può afferrare e muovere gli oggetti artificiali con le proprie mani: un sensore, simile a quello posto sul visore, permette di determinare la posizione della mano nello spazio. All'utente viene fornito un feedback appropriato; sono allo studio tecniche per fornire anche un feedback tattile. Una coppia di auricolari permette infine di udire suoni provenienti dalla scena artificiale, in stereofonia, con grande effetto realistico.

Le applicazioni della realtà artificiale per ora sono limitate, ma le prospettive sono estremamente suggestive: dalla telepresenza (la possibilità di muoversi in un ambiente remoto, vedendo ciò che vede una telecamera montata su un robot remoto), alla telemanipolazione (ad esempio, il controllo del braccio di un robot remoto), fino a nuovi prodotti di entertainment (computer games, cinema interattivo, ...) e alla computer art (Krueger, 1991). Nella realtà virtuale il concetto di interattività viene portato alle estreme conseguenze; le implicazioni filosofiche sono ampie e, per certi versi, sconcertanti. Dal punto di vista del progettista dell'interazione, i problemi sono numerosi, a incominciare dagli accorgimenti di tipo "cinestetico" che devono essere attuati per evitare fenomeni di disorientamento cognitivo che possono verificarsi quando ci troviamo in ambienti non familiari, o differenti, o troppo complessi o, semplicemente, per qualche ragione, sconcertanti per l'essere umano.

Sovrapposizione e adattamento di paradigmi

Mentre la evoluzione delle tecnologie hardware è estremamente rapida, il software ha un ciclo di vita piuttosto lungo: un sistema di software applicativo può vivere, rimanendo essenzialmente se stesso nonostante i continui interventi di manutenzione, anche 20 anni. Ne consegue che raramente il software in uso presso una organizzazione viene rinnovato contestualmente al rinnovo delle tecnologie hardware e, in particolare, di quelle relative all'interazione utente-calcolatore. Si osserva quindi in molti casi, per così dire, un ritardo dei paradigmi di interazione adottati dal software effettivamente in uso, rispetto a quanto sarebbe possibile in relazione alla tecnologia hardware installata. Così, ad esempio, programmi progettati per essere utilizzati con terminali scriventi sono spesso sopravvissuti a lungo dopo l'adozione di terminali video. Si pensi, ad esempio, ai line editor, sopravvissuti a volte per molti anni alla introduzione dei terminali video, nonostante che questa tecnologia permettesse, con gli editor full-screen, un trattamento del testo ben più agevole; si pensi, ancora, ad alcuni grossi sistemi di prenotazione voli basati su un'interazione a comandi tuttora in uso, ecc.. Si osserva così, spesso, una sorta di sovrapposizione di paradigmi diversi in uno stesso sistema di calcolo, che in qualche modo riflette la evoluzione tecnologica avvenuta durante la vita del sistema stesso.

Un altro fenomeno frequente è quello dell'adattamento dei paradigmi: paradigmi di interazione sviluppati in relazione a una determinata tecnologia vengono adattati a una tecnologia diversa (che potrà essere, di volta in volta, più semplice oppure più evoluta). Ad esempio, modalità di interazione nate per workstation dotate di mouse sono state spesso trasferite, successivamente, ad ambienti meno potenti (personal computer con grafica elementare e senza mouse, o anche terminali video). Lo stile di interazione suggerito dal mouse viene così "tradotto", sia pure in maniera semplificata, in modalità di interazione differenti (tipicamente, con l'uso di tasti funzionali). Si tratta, in questo caso, di un adattamento all'indietro ("retrofitting"): un paradigma di successo viene trasportato su una tecnologia più povera, per la quale, forse, non sarebbe mai stato concepito.

Naturalmente, esistono anche numerosi esempi di adattamento in avanti: paradigmi di interazione nati e diffusi con successo su una determinata tecnologia vengono riproposti nell'ambito di tecnologie più avanzate. Un esempio è quello dei linguaggi di comandi per i sistemi operativi: concepiti inizialmente per i terminali scriventi, sono stati estesamente utilizzati, sostanzialmente con la stessa filosofia, anche su terminali video e su personal computer evoluti (si pensi, ad esempio, ai linguaggi di comandi dei sistemi Unix, ancora oggi molto diffusi).

User model, metafora, mimesi

"L'essenza della metafora - ci dicono Lakoff & Johnson (1984) - è comprendere e sperimentare un tipo di cosa nei termini di un'altra". Studiando la evoluzione dell'interfaccia utente, ci si imbatte continuamente in numerose metafore, che sono state di volta in volta utilizzate per creare nuovi artefatti: la metafora del menù ("il video è un menù"); la metafora della scrivania ("il video è il ripiano della scrivania"); la metafora del foglio da disegno ("il video è un foglio da disegno"); la metafora del pannello di controllo ("il video è un pannello di controllo"); la metafora della stanza ("il video è una stanza"); la metafora degli agenti ("i programmi sono agenti"); la metafora dei virus ("programmi-virus infettano programmi sani"), e così via.

In effetti, il procedimento metaforico può essere di grande stimolo alla creatività nella progettazione di software e, in particolare, nella definizione di nuovi paradigmi di interazione. La metafora ci suggerisce nuove "virtualità", che poi possiamo realizzare con tutta la flessibilità concessa dal software: a volte, la metafora sarà interpretata in modo molto libero, altre volte il software riprodurrà fedelmente "dentro" il calcolatore la realtà alla quale ci si è ispirati. In quest'ultimo caso, parleremo più propriamente di mimesi. La metafora e la mimesi, inoltre, aiutano fortemente l'utente a costruirsi un "modello concettuale" corretto del sistema software. Che cosa si intenda per "user conceptual model" è ben riassunto da Smith et al. (1982):

Un modello concettuale dell'utente è l'insieme dei concetti che una persona gradualmente acquisisce per spiegare il comportamento di un sistema, sia che si tratti di un sistema computerizzato, o di un sistema fisico, o di un sistema ipotetico. È il modello sviluppato nella mente dell'utilizzatore che mette in grado quest'ultimo di comprendere il sistema e interagire con esso. Il primo compito per il progettista di un sistema è quello di decidere quale modello è preferibile per i suoi utilizzatori. Questo passo estremamente importante è spesso trascurato o effettuato malamente.

La metafora deve tuttavia essere usata correttamente. Metafore seguite in modo troppo letterale falliscono, e spesso elementi di "magia" (cioè funzioni o comportamenti che violano in qualche modo la metafora) possono contribuire grandemente al successo di un'interfaccia. Oggi sappiamo che le metafore di successo a un certo momento "scompaiono", generando, per così dire, oggetti completamente nuovi che vivono di vita propria, per essere eventualmente a loro volta utilizzati per produrre nuove metafore:

Metodi scoperti in un medium ci forniscono metafore che suggeriscono nuovi modi di pensare a nozioni in altri media (Kay & Goldberg 1977)

1.4. Progettare l'interazione

Non ci sono risposte semplici, solo compromessi. Donald A. Norman.

Principi, guideline, standard

Quali indicazioni si possono dare al designer di artefatti interattivi? Progettare l'interazione è mestiere nuovo, e occorre fornire opportuni strumenti (sia metodologici che concettuali) al progettista. Già la rapida panoramica proposta nella sezione precedente suggerisce la enorme ampiezza dello spettro delle soluzioni progettuali possibili. Parlare di interazione tout court rischia pertanto di non essere di alcuna utilità, in quanto il discorso si fa subito troppo ampio, e quindi generico. E' più facile, ovviamente, affrontare l'analisi di singole modalità di interazione, e tentare di fornire linee guida specifiche basate su evidenza scientifica (sperimentale) o almeno su considerazioni di buon senso comune. In questi termini, il problema è sicuramente più maneggevole, ed esiste una vasta letteratura in proposito. Ad esempio, già quasi vent'anni fa, nell'ambito di interfacce utente di tipo tradizionale (terminali), Smith & Mosier (1984) propongono ben 679 guideline per la progettazione del dialogo uomo-computer, organizzate in sei aree: data entry, data display, sequence control, user guidance, data transmission, data protection. Tre di queste guideline sono, ad esempio:

- Il computer dovrebbe fornire una risposta alle azioni di data entry rapidamente, in modo tale che l'utente non venga rallentato o che non debba adattare il suo ritmo di lavoro ai ritardi di risposta del computer; per le operazioni normali, il ritardo dei feedback sul video non dovrebbe superare i 0.2 secondi...
- Visualizzare preferibilmente parole complete piuttosto che abbreviazioni...
- Il computer dovrebbe fornire un'indicazione sullo stato della transazione, ogni qual volta la risposta completa a una richiesta dell'utente subisce qualche ritardo...

Di fronte a tali imponenti sforzi prescrittivi, ci si sente a disagio: in assenza di una sistematizzazione teorica che ci permetta di ricavare le linee guida operative da principi di carattere generale, il progettista ha la sensazione di muoversi su un terreno incerto, in cui le indicazioni sono spesso fra loro in conflitto, e quindi ogni scelta è opinabile. L'unica via d'uscita sembra allora essere la conformità ad esempi di sistemi analoghi a quello in corso di progettazione, che siano ritenuti, per giudizio comune, "buoni". Dopotutto, chi mai è in grado di seguire contemporaneamente 679 indicazioni? Ciò che ci serve è quindi una nuova disciplina scientifica, correttamente fondata su principi generali:

Se noi ci proponiamo una scienza dell'interazione uomo-computer, è essenziale che abbiamo dei principi dai quali derivare il modo di interazione fra persona e computer. E' facile immaginare degli esperimenti per verificare questa o quella idea, confrontare o contrastare alternative, valutare la qualità della offerta tecnologica più recente. Ma noi dobbiamo aspirare a qualcosa di più della risposta ai bisogni correnti. La tecnologia sulla quale si basa l'interfaccia uomo-computer cambia rapidamente in confronto al tempo con cui gli esperimenti psicologici forniscono risposte. Se non stiamo attenti, le risposte di domani riguarderanno le preoccupazioni di ieri (Norman, 1983).

Ma la strada da percorrere, per questo, è ancora lunga. La disciplina della comunicazione uomo-computer è solo un capitolo, ma un capitolo importante, della disciplina della comunicazione tout court, e la comprensione dei meccanismi della comunicazione umana resta ancora un problema di formidabile complessità.

La metafora dell'extraterrestre

Per far comprendere la natura e la complessità dei problemi di design della interazione, partiamo da una metafora nella quale il lettore si possa facilmente immedesimare. Supponiamo di dover effettuare una conversazione con un interlocutore sconosciuto, proveniente da una lontana galassia. Quali sono i principi ai quali i due interlocutori dovranno attenersi, implicitamente o esplicitamente, affinché il dialogo sia possibile e, preferibilmente, facile e gradevole? Per semplificare la discussione, cerchiamo di individuare un numero limitato di requisiti semplici e generali che dovranno essere soddisfatti affinché un dialogo accettabile sia possibile, e che siano indipendenti dal particolare contenuto della conversazione, in modo che si possano ragionevolmente applicare in tutti i casi. Un possibile elenco è il seguente.

1. Principio del vocabolario comune. E' ovvio che, preliminarmente alla conversazione, i due interlocutori dovranno in qualche modo negoziare un linguaggio (anche rudimentale, e non necessariamente parlato) comune. Questa negoziazione, che è il presupposto di ogni dialogo, può essere effettuata in vari modi. Riprenderemo il discorso più oltre.

2. Principio di adeguatezza. Il linguaggio comune scelto per la comunicazione dovrà essere adeguato ai meccanismi fisiologici e cognitivi dei due interlocutori. Ad esempio, l'interlocutore galattico non potrà aspettarsi che l'interlocutore terrestre riconosca suoni al di là della frequenza udibile dall'orecchio umano, né che sia in grado di memorizzare lunghe sequenze di fonemi, né che possa attendere una risposta per un periodo di cento anni, e così via (adeguatezza all'utente). Il linguaggio dovrà inoltre essere adeguato ad esprimere i contenuti della conversazione che si desidera condurre (adeguatezza al compito).

3. Principio di consistenza. Supponiamo di avere messo a punto un linguaggio che rispetti il principio di adeguatezza, e incominciamo a utilizzarlo per la comunicazione con il nostro amico extraterrestre. Per poterci comprendere, sarà comunque necessario che questo linguaggio venga utilizzato, da ciascun interlocutore, in modo consistente. Il concetto è complesso, ma per i nostri scopi sarà sufficiente intenderlo secondo il senso comune, applicandolo sia ai contenuti della comunicazione (in sostanza, ciascun interlocutore non dovrà contraddirsi), che alla forma della stessa (in sostanza, ciascun interlocutore cercherà di "dire cose simili in modo simile"¹⁶).

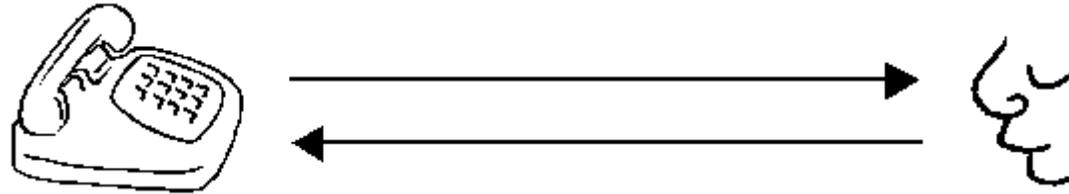
4. Principio del feedback. In ogni conversazione, anche quando ci sia una buona intesa sugli strumenti comunicativi utilizzati, dobbiamo continuamente fornire e ricevere rassicurazioni o "sconferme". Ogni interlocutore deve segnalare, utilizzando il linguaggio stesso usato nella conversazione o strumenti accessori (una espressione del viso, un gesto, ...), se il messaggio dell'altro interlocutore è stato recepito (feedback positivo) o se, al contrario, si è verificato qualche problema, e il messaggio deve essere ritrasmesso (feedback negativo).

5. Principio di annullamento ("undo"). In ogni caso (e cioè sia in presenza di feedback positivo che di feedback negativo) ciascun interlocutore deve essere in grado di annullare parte della comunicazione senza che ciò comprometta il proseguimento del dialogo. In presenza di un precedente feedback negativo, questo ci permetterà di ritrasmettere un messaggio che non è stato compreso; più in generale, questo principio ci riconosce, per così dire, il diritto di cambiare idea, e di riorientare in altro modo la conversazione sulla base dell'esito della conversazione stessa fino a quel momento (o di altri motivi "esterni" di natura qualsiasi).

6. Principio di libertà. In ogni caso, dovrà essere salvaguardata la libertà di ciascun interlocutore di interrompere la conversazione in qualunque momento, qualunque cosa l'altro interlocutore stia facendo o dicendo, senza compromettere, con questo, la possibilità di riprendere la conversazione in un momento successivo ¹⁷.

Conversare con gli oggetti

La metafora dell'extraterrestre è stata qui introdotta con obbiettivi esclusivamente didattici, e non può essere spinta troppo in là. E' comunque stimolante, in quanto facilmente comprensibile e trasferibile anche al "dialogo" con artefatti inanimati, ma interattivi. Pensiamo, ad esempio, a un utente alle prese con uno degli artefatti interattivi più diffusi, un apparecchio telefonico:



Leggiamo allora in questo contesto i principi discussi nel paragrafo precedente, per verificarne la validità in un corretto design dell'interazione:

Principio del vocabolario comune. La interazione con l'apparecchio telefonico è realizzata attraverso un semplice linguaggio multimodale, che comprende gesti (alzare/abbassare il ricevitore, premere i tasti per formare il numero) e suoni (i segnali di libero/occupato o la mancanza di segnali).

Principio di adeguatezza. Per essere adeguato alle capacità fisiologico/cognitive di un utente umano, il linguaggio di interazione sopra descritto dovrà essere progettato con opportuni criteri: ad esempio, i segnali acustici dovranno essere ben udibili e fra loro distinguibili (almeno in assenza di eccessivi disturbi sulla linea), e dovranno avere un volume sonoro compreso in un opportuno intervallo; i tasti dovranno essere correttamente distanziati, e così via. L'adeguatezza al compito, invece, sarà garantita da accorgimenti di design che tengano conto delle particolari condizioni d'uso del nostro apparecchio. Ad esempio, nei tradizionali modelli della Sip, la forcella che sostiene il ricevitore serve anche a impedire che il pulsante che interrompe la conversazione venga abbassato da urti accidentali: questo non accadeva in modelli più vecchi, in cui è la forcella stessa che si abbassa, con una protezione molto più limitata da interruzioni non volute ¹⁸.



Principio di consistenza. Affinché l'apparecchio sia utilizzabile, le convenzioni per interpretare il segnale acustico devono essere seguite in modo consistente: ad esempio, il tono breve sempre per segnalare occupato, il tono lungo sempre per segnalare libero (si pensi alle difficoltà che si incontrano quando, all'estero, ci si imbatte con convenzioni differenti).

Principio del feedback. Per poter usare l'apparecchio, il feedback di tipo acustico (libero/occupato) è fondamentale (che cosa facciamo quando non udiamo alcun segnale?). In un apparecchio correttamente disegnato, ci sono tuttavia anche altri tipi di feedback, altrettanto importanti. Ad esempio, per permetterci di comporre con facilità il numero i tasti ci forniscono un feedback ben percepibile, sia di tipo acustico che di tipo tattile.

Principio di annullamento. Con qualsiasi apparecchio telefonico possiamo annullare ciò che stiamo facendo (ad esempio, comporre un numero) e ripartire da capo in qualunque momento, con grande facilità: basta infatti premere il pulsante di interruzione sulla forcella.

Principio di libertà. Il nostro apparecchio telefonico ci deve permettere di interrompere la chiamata in qualunque istante (mentre sto componendo il numero, mentre l'apparecchio all'altro capo sta squillando o è occupato, mentre il mio interlocutore sta parlando, mentre sto parlando io). Si pensi, ad esempio, alle difficoltà che avremmo se il protocollo di comunicazione ci imponesse, per interrompere la comunicazione, una sequenza prefissata di azioni (ad esempio, prima il chiamante, poi il chiamato, ...), come a volte accade per effettuare trasferimenti di comunicazione fra un apparecchio e l'altro...

Se invece di un "semplice" apparecchio telefonico dobbiamo progettare un sistema software, le possibilità si ampliano, e le scelte si fanno quindi più complicate. I principi generali sopra enunciati, tuttavia, restano di importanza fondamentale per orientare una "buona" progettazione: alla loro violazione possiamo, senza troppe difficoltà, far risalire i principali errori di disegno ricorrenti nei sistemi in uso.

Affordance

Quando abbiamo introdotto il "principio del linguaggio comune", non ci siamo occupati di come questo linguaggio possa essere appreso dai due interlocutori. Nel caso di artefatti interattivi, e in particolare di sistemi software, vengono normalmente utilizzate opportune tecniche per facilitare questo apprendimento: dai manuali d'uso (eventualmente accessibili "in linea"), a vari tipi di ausili interattivi ("help on-line"). Non discutiamo qui questi temi, rimandando il lettore alla letteratura esistente. Ci preme però sottolineare un principio che normalmente viene trascurato nella pratica della progettazione:

Un sistema ben progettato deve essere utilizzabile senza alcuna necessità di manuali d'uso.

Si tratta, per così dire, di un "principio limite" (la documentazione di riferimento andrà comunque realizzata sempre), per soddisfare il quale dovremo utilizzare particolari accorgimenti di progettazione. Ci viene in aiuto la nozione di affordance, introdotta da Gibson e utilizzata da Donald Norman nel contesto del disegno industriale:

Quando incontriamo un oggetto insolito, come facciamo a sapere che cosa farci? O abbiamo avuto a che fare con qualcosa di simile in passato e trasferiamo la vecchia conoscenza all'oggetto nuovo, oppure ci facciamo dare istruzioni. In questi casi, l'informazione che ci serve è di tipo interno. Un altro approccio consiste nell'usare l'informazione presente nel mondo esterno, in particolare se il design del nuovo oggetto ci presenta un'informazione che possa essere interpretata ¹⁹.

E più oltre:

Come può il design segnalare le azioni appropriate? (...) Un insieme importante di segnali ci arriva attraverso i vincoli naturali degli oggetti, vincoli fisici che limitano le possibilità di azione. Un altro insieme di segnali viene da quelle che abbiamo chiamato affordance, gli inviti forniti dagli oggetti, che trasmettono messaggi circa i loro possibili usi, azioni e funzioni. Una piastra liscia invita a spingere, un contenitore vuoto a riempirlo, e così via. Le affordance possono segnalare come si può muovere l'oggetto, che cosa è in grado di sorreggere, se qualcosa può incastrarsi nei suoi incavi, sopra e sotto. Da che parte lo si afferra, quali sono le parti mobili e quali fisse? Gli inviti d'uso suggeriscono la gamma delle possibilità, i vincoli limitano il numero delle alternative. L'uso intelligente di inviti e vincoli d'uso combinati nella progettazione fa sì che l'utente possa determinare prontamente il corso esatto delle azioni, anche in una situazione del tutto nuova.

Inviti e vincoli (e, in generale, il concetto di affordance) sono utilizzabili anche nella progettazione di artefatti "immateriali", realizzati via software. Anzi, il fatto di essere del tutto libero da vincoli imposti dalle caratteristiche fisiche dei materiali utilizzati, e dai processi di produzione industriale, dà al progettista di artefatti software una libertà finora sconosciuta al designer industriale.

1.5. Conclusioni

Il software dà una nuova dimensione al design industriale. Con la tecnologia dei computer siamo in grado di progettare artefatti dotati di un livello di interattività sconosciuto in passato. Oggi i limiti all'attività di progettazione di questi artefatti non sono tanto tecnologici, quanto metodologici e concettuali. Occorre analizzare con occhi nuovi il concetto stesso di interattività, ed occorre elaborare un insieme di principi generali dai quali si possano trarre indicazioni e linee guida concrete per il progettista. Il campo di indagine è fortemente multidisciplinare, e costituisce una disciplina recente ma già molto ampia, denominata HCI (Human-Computer Interaction), nella quale confluiscono varie discipline: l'ingegneria del software, la psicologia cognitiva, la comunicazione, il design industriale.

In queste note abbiamo discusso i principali problemi che il design della interattività pone al progettista, e riesaminato i diversi paradigmi secondo i quali la interazione uomo-computer si è realizzata fino ad oggi. Il filo logico per questa ricostruzione ci è stato fornito dalla evoluzione della tecnologia di interazione: le tecnologie che abbiamo individuato ci hanno permesso di ordinare in uno schema dotato di una sua coerenza i numerosi modelli sperimentati in questo primo mezzo secolo di storia dei computer. A fronte di questo riesame, è possibile convincerci che le potenzialità insite nel concetto di artefatto interattivo sono vastissime: la interattività e la multimodalità - le due parole chiave centrali nel nostro discorso - sono concetti ancora in buona misura da esplorare.

In questo contesto, la enfasi sulla tecnologia che guida gran parte degli sviluppi nel settore della Human-Computer Interaction va considerata, a nostro avviso, con un atteggiamento di cautela, e occorre interrogarsi sul senso complessivo degli svariati tipi di nuovi artefatti che tale tecnologia rende possibili (ma che spesso non superano lo stadio di prototipi in laboratori di ricerca). Citando Winograd & Flores (1986):

Noi siamo interessati a ciò che accade quando nuovi dispositivi vengono creati, e a come nascono le possibilità di innovazione. Qui c'è una circolarità: il mondo determina ciò che noi possiamo fare, e ciò che noi facciamo determina il mondo. La creazione di un nuovo dispositivo o dominio sistematico può avere significati che vanno lontano - può creare nuovi modi di essere che prima non esistevano, e una cornice per azioni che prima non avrebbero avuto senso.

Occorre ricondurre costantemente il lavoro di progettazione all'utente, che ne costituisce il destinatario finale e la ragione d'essere. In "The Hitchhiker's Guide to the Galaxy", famoso cult-book di fantascienza comica dei primi anni 80 troviamo questo brano, che vogliamo qui citare come epigrafe ad avvertimento dei pericoli insiti nella tecnologia di interfaccia fine a se stessa:

Per anni le radio erano state fatte funzionare premendo pulsanti o girando manopole; in seguito, quando la tecnologia divenne più sofisticata, i comandi diventarono sensibili al tocco - bastava sfiorare i pannelli con le dita; ora tutto quello che c'era da fare era far ondeggiare la mano nella direzione dei componenti e sperare. Risparmiava, naturalmente, un sacco di fatica muscolare, ma il risultato era che bisognava star seduti furiosamente immobili se si voleva continuare a sentire lo stesso programma.

1.6. Riferimenti Bibliografici

- ACM SIGCHI Special Interest Group on Computer-Human Interaction Curriculum Development Group, "Curricula for Human-Computer Interaction", ACM, 1992, pp.iv+162
- Adams, D., "The More Than Complete Hitchhiker's Guide", Longmeadow Press, Stamford, CT, 1986
- Apple Computer Co., "Human Interface Guidelines: the Apple Desktop Interface", Addison-Wesley Publishing Co., 1987, pagg. xii+144
- Barsalou, L.W., "Cognitive Psychology - An Overview for Cognitive Scientists", Lawrence Erlbaum Associates, 1992, pp.xi+410
- Bolt, Richard A., "The Human Interface - Where People and Computers Meet", Lifetime Learning Publ., Belmont, CA., 1984, pagg.xvii+113
- Brand, S., "The Media Lab - Inventing the future at M.I.T.". Penguin Books, 1987-1988.
- Card, S.K., Moran, T.P., Newell, A., "The Psychology of Human-Computer Interaction", Lawrence Erlbaum Associates, Hillsdale, 1983, N.J., pagg.xiii+469
- English, W.K., Engelbart, D.C., Berman, M.A., 1967, "Display-selection techniques for text manipulation", IEEE Transactions on Human Factors in Electronics, HFE-8, pagg.5- 15
- Gaver, W.W., "Technology Affordances", in CHI '91 Conference Proceedings (ACM), pagg.79-84
- Goldberg,A., "A History of Personal Workstations", Addison-Wesley Publishing Co., ACM Press, New York, 1988, pagg.ixi+537
- Holub, R.C.(a cura di), "Teoria della ricezione", Einaudi Editore, 1989, pagg.xxxii+305
- Krueger, M., "Artificial Reality II", Addison-Wesley, 1991
- Joyce, M., "Selfish Interaction: subversive texts and the multiple novel", in Thurber, B.D.: "Literacy in the Computer Age", Paradigm, 1991
- Kay, A., Goldberg, A., "Personal Dynamic Media", in Computer, marzo 1977, pagg.31-41, ristampato in Goldberg 1988
- Lakoff, G., e Johnson, M., "Metaphors We Live By", The University of Chicago Press, 1980, pagg.xiii+242
- Lippman, A., "Movie-Maps: An Application of the Optical Videodisc to Computer Graphics", in Computer Graphics 14(3), 1980, pagg.32-42
- Manzini, E., "Artefatti - Verso una nuova ecologia dell'ambiente artificiale", Edizioni Domus Academy, Milano, 1990, pagg.192
- Nelson, T., "Interactive Systems and the Design of Virtuality", in Creative Computing, Vol.6, n.11, novembre 1980, e Vol.6, n.12, dec 1980
- Nielsen, J., "Hypertext and Hypermedia", Academic Press, 1990
- Norman, D.A., "Design Principles for Human-Computer Interfaces", in Proceedings of CHI '93 (ACM), pagg.1-10
- Norman,D.A., "The Psychology of Everyday Things", Basic Books, Inc., New York, 1988; ed.italiana: "La caffettiera del masochista - Psicopatologia degli oggetti quotidiani", Gruppo Editoriale Giunti, Firenze, 1990, pagg.xi+294
- Norman, D.A., "Turn Signals are the Facial Expressions of Automobiles", Addison-Wesley Publishing Co., 1992, pp.xiv+205
- Papert, S., "Mindstorms: Children, Computers, and Powerful Ideas", Basic Books, Inc., New York, 1980
- Schnaars, S.P., "Megamistakes", The Free Press, 1989 (Ed.italiana: "Mega errori", Sperling & Kupfer Editori, 1991)
- Shneiderman, B., "Direct Manipulation: A Step Beyond Programming Languages", in IEEE Computer, agosto 1983, pagg.57-69
- Shneiderman, B., "Designing the User Interface - Strategies for Effective Human-Computer Interaction - Third Edition", Addison-Wesley P.Co,
- Smith, C., Irby, C., Kimball, R., Verplank, B., Harslem, E., "Designing the Star User Interface", in BYTE, Vol.7, n.4 aprile 1982, pagg.242-282
- Smith, S.L., Mosier, J.N., "Design Guidelines for User-System Interface Software"; Technical Report ESD-TR-84-190, Hanscom Air Force, MA
- Thimbely, H., "User Interface Design", ACM Press, 1990, pp.xxiii+470
- Weizenbaum, J., "ELIZA - A Computer Program For the Study of Natural Language Communication Between Man And Machine", Communications of th ACM, Vol.9, n.1, Gennaio 1966, pagg.36-45
- Winograd, T., Flores, F., "Understanding Computers and Cognition: A New Foundation for Design", Ablex Publishing Co., Norwood, New Jersey, 1986, pagg.xiv+207

¹ - Questo articolo è la revisione, effettuata nel febbraio 2001, di un articolo precedentemente pubblicato nel volume "Il progetto delle interfacce", a cura di Giovanni Anceschi, Edizione Domus Academy, marzo 1993.

² - dal Dizionario Garzanti della Lingua Italiana, 1965

³ - Fra i prodotti di personal computing possiamo trovare diversi esempi di questi strumenti per pensare. Tipici esempi sono le funzioni di "outline" presenti nei word processor, che ci supportano nella creazione della "scaletta" del testo oppure molti computer games, fino ai normali fogli elettronici, che ci aiutano a prendere decisioni sulla base del confronto di scenari alternativi.

⁴ - E' significativo osservare che questo aspetto risulta in genere molto difficile da comprendere da parte di chi non ha pratica con questi strumenti. Il profano è in genere affascinato dalla tecnologia, e trova grande difficoltà a comprendere che questa costituisca essenzialmente soltanto una "scatola": una infrastruttura di materiale elettronico che nulla ha a che vedere con gli strumenti che, attraverso di esso, si incarnano per essere da noi usati. Questo problema non sussiste, ad esempio, con gli apparecchi televisivi, dove è ben chiaro a tutti (anche a chi non s'intenda di elettronica) la differenza fra l'elettronica e i programmi televisivi di cui, per il suo tramite, possiamo fruire.

⁵ - Ci si riferisce qui alle ampie attività di ricerca e sperimentazione in corso per la creazione di nuovi device di interazione uomo-computer : speech recognition, eye-tracking, gesture-tracking, body-tracking, ecc.

⁶ - Si vedano, ad es., i curricula scolastici per i corsi di Human-Computer Interaction proposti dallo Special Interest Group on Computer-Human Interaction Curriculum Development Group dell'ACM (ACM SIGCHI, 1992).

⁷ - Questa è la definizione proposta in ACM SIGCHI, 1992(vedi nota precedente).

⁸ - Ibid.

⁹ - Vocabolario Treccani della Lingua Italiana, 1987

¹⁰ - La definizione, e i cinque corollari che seguono, sono tratti da un'intervista di Steward Brand ad A. Lippman, riportata in Brand (1987), pagg.46-50.

¹¹ - Vedi nota precedente. L'intervista, trascritta fedelmente dalla registrazione, costituisce di per sé stessa un interessante esempio di dialogo "interattivo" che si esemplifica da sé mostrando evidenti esempi dei cinque corollari di cui tratta - una spiegazione ed una dimostrazione allo stesso tempo. I brani da noi riportati sono un collage tratto liberamente da tale intervista.

¹² - Si fa qui riferimento all'esempio del videodisco "Movie Map" di Aspen, Colorado, realizzato dallo stesso Lippman nel 1978 (Lippman, 1980).

¹³ - L'esempio è tratto da Gaines & Shaw, "The art of computer conversation", Prentice-Hall, 1984

¹⁴ - "Buongiorno il mio nome è Gaines e desidero parlare con te / Non mi interessano i nomi / Mi dispiace forse puoi dirmi quali sono i tuoi interessi / Per favore non chiedere scusa / Grazie. Che cosa ti interessa di più? / Perché me lo chiedi? / Perché debbo dimostrare le tue ineffabili capacità / Questa è la vera ragione? / Sì le tue capacità mi interessano / Sembri molto positivo / In effetti io sono positivo / E' perché in effetti tu sei positivo che sei venuto da me? / No sono venuto per dimostrare le tue capacità / Tu dici 'no' solo per essere negativo / Sì / Ti stai comportando in modo brusco con me / Sono molto irritato dalla tua mancanza di garbo nella conversazione / Per quanto tempo sei stato irritato dalla mia mancanza di garbo nella conversazione? / Per parecchi minuti / Non sono certo di capire bene / Puoi dirmi l'ora? / Tu credi che io possa dirti l'ora, non è vero? /"

¹⁵ - Il design più tipico di questo stile di interazione è quello proposto dal foglio elettronico 123, distribuito dalla Lotus a partire dal 1983 e, in quegli anni, il prodotto software più venduto di tutti i tempi.

¹⁶ - Questo naturalmente potrà impoverire la forma della comunicazione, ma teniamo presente che l'obiettivo principale, nel nostro esempio, è semplicemente quello della comprensione reciproca, anche se nello scambio di messaggi elementari.

¹⁷ - Questo principio dà ad ogni interlocutore, per così dire, pari diritti sul controllo della conversazione. Progettando un dialogo uomo-computer, normalmente preferiamo attribuire, invece, tali diritti esclusivamente all'interlocutore umano, ed è per questo che il principio del "controllo all'utente" è affermato da quasi tutte le guideline presenti in letteratura. Ciononostante, questo principio è sistematicamente violato in molti sistemi, che spesso costringono l'utente a portare a compimento una lunga serie di operazioni una volta che queste siano state iniziate, o "abbandonano" l'utente improvvisamente e definitivamente con un messaggio brutale (ad esempio, per un "errore di sistema").

¹⁸ - L'esempio è citato anche in Norman, 1988 (pag.188 dell'edizione italiana).

¹⁹ - Norman, 1988 (pag.109 della edizione italiana).

²⁰ - Ibid., pag.

2

capitolo



2.0. Usabilità

Usabilità: l'efficacia, efficienza e soddisfazione con cui determinati utenti possono raggiungere determinati obiettivi in determinati ambienti d'uso (ISO 9241).

I sette stadi dell'azione (Norman):

1. Formare lo scopo (*che scopo voglio raggiungere?*)
2. Formare l'intenzione (*che cosa intendo fare per raggiungere lo scopo?*)
3. Specificare un'azione (*quali azioni specifiche debbo compiere per farlo?*)
4. Eseguire l'azione
5. Percepire lo stato del mondo (*come è cambiato il mondo dopo?*)
6. Interpretare lo stato del mondo (*che cosa significa?*)
7. Valutare il risultato (*ho raggiunto il mio scopo?*)

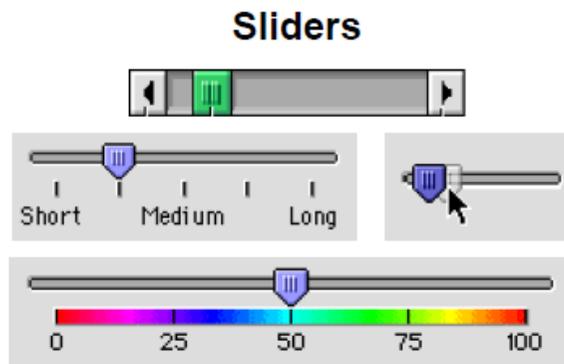
Valutazione:

- il golfo dell'esecuzione (la differenza fra le intenzioni e le azioni possibili)
- il golfo della valutazione (lo sforzo per interpretare lo stato fisico del sistema e determinare fino a che punto corrisponda alle aspettative o alle intenzioni).

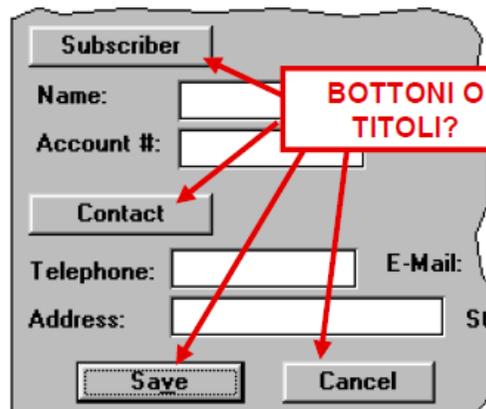
Gli obiettivi principali dell'interaction design

1. Ridurre l'ampiezza del golfo dell'esecuzione (fare in modo che le azioni possibili corrispondano in modo evidente alle intenzioni) = affordance
2. Ridurre l'ampiezza del golfo della valutazione (lo stato fisico del sistema deve essere interpretabile in modo univoco e immediato) = feedback

Affordance: la proprietà di un oggetto d'influenzare, attraverso la sua apparenza visiva, il modo in cui viene usato. Un oggetto che possiede una buona affordance "invita" chi lo guarda a utilizzarlo nel modo per cui è stato concepito. Il compito del progettista è progettare oggetti con buona affordance, per ridurre l'ampiezza del golfo della esecuzione. Il termine è stato introdotto dallo psicologo James J. Gibson nel 1966, e ripreso poi da D. Norman. Un oggetto interattivo ben progettato dovrebbe "non farci pensare" mentre lo usiamo.



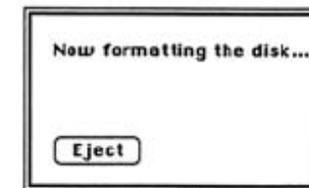
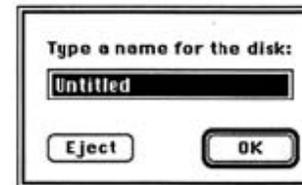
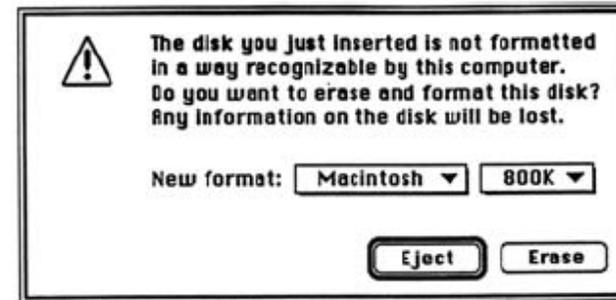
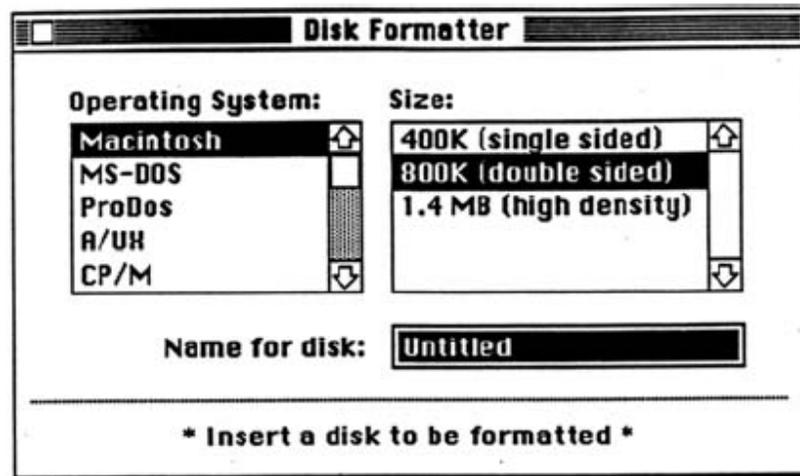
MAC OS 8



Esempio: feedback facilmente interpretabile?



Da: SQL Windows



"Ho una regola semplice per individuare il cattivo design. Tutte le volte che trovo indicazioni su come usare qualcosa, si tratta di un oggetto progettato male". Donald Norman.

"I manuali d'uso sono il cattivo alibi del progettista incapace". Anonimo

Livelli di maturità della progettazione:

- primo stadio: il prodotto *funziona*
- secondo stadio: il prodotto fornisce le *funzioni* richieste
- terzo stadio: il prodotto è *facile* da imparare e da usare
- quarto stadio: il prodotto è *invisibile* durante l'uso

2.1. L'occhio

- I raggi luminosi passano attraverso il cristallino e vengono messi a fuoco sulla retina
- La pupilla controlla la quantità di luce che entra nell'occhio, variando il proprio diametro (7-8 mm \Rightarrow 2-3 mm)

La retina

Gli impulsi nervosi vengono convogliati dalle fibre del nervo ottico verso la corteccia visiva, che si trova nella parte posteriore del cervello

Recettori

Coni

- sensibili ai colori
- sono più numerosi al centro (fovea)
- 6 ml per occhio

Bastoncelli

- sensibili all'intensità della luce
- non discriminano i colori
- visione notturna e percezione del movimento
- sono più numerosi in periferia
- 120 ml per occhio

Vengono attivati dei processi fotochimici che generano impulsi elettrici nelle fibre nervose

Disposizione dei coni e bastoncelli sulla retina

Al centro (fovea): visione più acuta; in periferia: visione del movimento.

I muscoli ciliari fanno aumentare o diminuire lo spessore del cristallino per la messa a fuoco (lente a lunghezza focale variabile).

Capacità dell'occhio di distinguere due punti vicini: è misurata dall'angolo minimo sotto cui devono essere visti perché l'occhio li percepisca separatamente. Se tale angolo vale 1', le loro immagini si trovano sulla retina a una distanza di 5 μ m e stimolano due elementi non contigui della stessa, condizione indispensabile perché siano visti distinti da un occhio normale. NB: 1° (grado) = 60' (minuti) = 3600'' (secondi).

Misura dell'acuità visiva: si misura in valori reciproci dell'angolo visivo minimo alla quale due punti non appaiono più separati.

Esempio: Se tale l'angolo è di 2' l'acuità visiva è pari a 1/2, ossia a 5/10 (non è la metà del normale, poiché l'acuità visiva normale è 11/10).

L'acuità visiva dipende dall'età del soggetto (tende a diminuire dopo i 70 anni), dallo stimolo, dalle caratteristiche dell'occhio, dall'integrità dei coni, ecc. L'acuità visiva è massima in corrispondenza della fovea centrale, e diminuisce verso la periferia.

Movimenti oculari

Saccadi: Movimenti oculari molto veloci (possono superare i 400°/sec) e molto brevi (20-50 msec, durante i quali la visione è soppressa), che hanno il compito di spostare l'asse visivo durante l'esplorazione di una scena (fino a 4/5 volte al sec).

Fissazione: Pausa tra due saccadi successive; rappresenta l'intervallo di tempo durante il quale viene acquisita l'informazione visiva (-60-700 msec).

Scanpath: tracciato bidimensionale che gli occhi compiono durante l'esplorazione di una scena, composta da una successione di saccadi e di fissazioni (durata tipica saccade+fissazione: 230 msec)

Movimenti oculari: durata e ampiezza

Compito Durata media fissazione(msec) - Ampiezza media saccade (gradi)

Lettura silenziosa 225 - 2° (circa 8 caratt.)

Lettura orale 275 - 1.5° (circa 6 caratt.)

Ricerca visiva 275 - 3°

Ispezione di scene visive 330 - 4°

Eye tracking

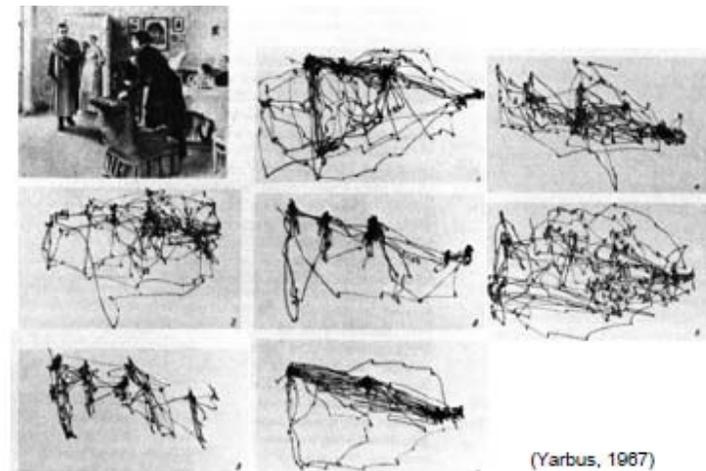
Una sorgente di infrarossi colpisce la cornea generando il riflesso corneale che rende luminosa la pupilla. Una videocamera riprende la posizione della pupilla e un software opportuno ricostruisce il movimento compiuto dallo sguardo del soggetto durante l'esplorazione della scena. Oggi esistono apparecchiature non invasive. La taratura viene fatta molto velocemente fissando una ventina di punti che appaiono in posizioni diverse dello schermo.

Esempio: analisi dei movimenti oculari

- 1. Esame libero del quadro

Gli altri diagrammi corrispondono a specifici task chiesti all'osservatore, come segue:

- 2. Esame dell'ambiente materiale
- 3. Esame dell'età delle persone
- 4. Che cosa facevano prima dell'arrivo del visitatore inatteso?
- 5. Ricordare gli abiti indossati dalle persone
- 6. Ricordare la posizione delle persone e degli oggetti nella stanza
- 7. Quanto tempo il visitatore inatteso è stato lontano dalla famiglia?



Oggetti quotidiani

Un adulto usa e sa riconoscere a prima vista più di 20.000 oggetti.

Alcuni sono di uso quotidiano, per cui ne abbiamo un chiaro modello concettuale. Altri sono semplici da usare, perché offrono chiari indizi sul loro funzionamento. Altri sono irrimediabilmente oscuri e complessi. Perché? Ci formiamo naturalmente un modello concettuale del funzionamento di un oggetto. Usiamo vincoli, inviti e correlazioni spaziali. Ad esempio la bicicletta: il sellino, le manopole del manubrio, i pedali invitano chiaramente in una ed una sola posizione. Le forbici: il lato tagliente, i buchi delle dita, uno più grosso e l'altro più piccolo, non esiste modo per sbagliare. Un contro-esempio è il termostato: controlla la temperatura di arrivo, non l'intensità di riscaldamento: aumentare il termostato non velocizza il riscaldamento di una casa fredda! Affordance negli oggetti sono quindi le proprietà percepite che indicano come usare l'oggetto: le piastre si spingono, le manopole si girano, nelle fessure si infilano oggetti. Sono necessarie però anche informazioni di ritorno, che permettono all'utente di capire che azioni ha eseguito e con quale risultato. Più è evidente, non oppressivo e non ambiguo, meno ci interroghiamo sul buon uso dell'oggetto. Ad esempio, nei vecchi telefoni a rotella che usavano centrali elettromagnetiche, l'uso della rotella aveva resistenze, attriti e suoni evidenti, e il buon andamento della connessione veniva fornito da click, ronzii e altri suoni di centrale. I telefoni moderni sono muti o fanno suoni elettronici di scarsa interpretabilità.

2.2. L'errore

Gli esseri umani sbagliano spesso e volentieri. Un oggetto fatto per esseri umani che non preveda la possibilità dell'errore è un oggetto inutilizzabile. Sono molte le cause di errore da parte degli esseri umani:

- spiegazioni fallaci,
- senso di impotenza,
- problemi nell'esecuzione o nell'interpretazione delle azioni,
- lapsus.

Gli esseri umani sono creature che spiegano:

- La spiegazione si basa su analisi e valutazioni incomplete,
- Spesso anche su una concezione mitologica e antropomorfa degli eventi esterni (il computer mi vuole male),
- Fisica aristotelica ingenua,
- Concezioni ragionevoli (buon senso) negate dalla fisica.

Impotenza appresa

- Tendenza ad incolpare se stessi,
- Provoca un giudizio di incapacità globale su di sé, che non viene più messo alla prova. Si smette anche solo di tentare.

Gestire l'errore

- Capire le cause dell'errore ed impostare il progetto in modo da ridurle al minimo
- Rendere tutte le azioni reversibili
- Rendere più difficili le azioni irreversibili
- Facilitare la scoperta degli errori e facilitarne la correzione
- Considerare che l'utente sbaglia, ed anzi funziona per approssimazioni successive.

Le funzioni obbligatorie

Sono vincoli fisici che impediscono di compiere certe azioni in maniera sbagliata o in momenti sbagliati

Impotenza insegnata

- Cattive spiegazioni, immagini di sistema, libri o insegnanti ci convincono che non siamo fatti per un dato compito, come ad esempio la matematica.

Profezie autoavverantesi (*Self-fulfilling prophecies*)

- Andare male ad un esame

Gli errori nelle azioni sono a volte dovuti alla distanza tra il modello mentale dell'utente e il mondo esterno, reale, degli oggetti su cui agiamo. Queste distanze vengono chiamate *golfi*. La teoria dei sette stadi dell'azione identifica due punti in cui si può verificare un errore, due golfi principali:

- Il *golfo dell'esecuzione* è la distanza tra le intenzioni proposte da un oggetto e le azioni in esso effettivamente possibili;
- Il *golfo della valutazione* è lo sforzo necessario per valutare lo stato fisico del sistema DOPO l'azione, ed è minore quando il sistema offre molte informazioni coerenti di feedback.

- *Interlock*: il funzionamento di una parte impedisce il funzionamento di un'altra parte. Es.: Portiere e tappo del serbatoio di un'automobile

- *Lockin*: il funzionamento di un oggetto è garantito anche in caso di interruzione accidentale (es.: i ganci di un telefono, un'interruttore morbido in un computer, il tasto "Annulla" dopo il comando di uscita da un'applicazione)

- *Lockout*: il funzionamento di un oggetto è reso volutamente difficile in modo da garantire la volontà della persona nell'attivarlo (es.: le scale di sicurezza per il seminterrato, opzioni pericolose di un programma).

Donald Norman, "The Psychology of Everyday Things", 1988, (ed. it. "La caffettiera del Masochista", Giunti ed., 1990).

Il suono nell'interazione uomo macchina

- È poco usato, si potrebbe fare molto di più, potrebbe veicolare informazioni complesse
- Esempi:
 - Stato del sistema (es. suono di fondo per segnalare la esecuzione di un processo)
 - Conferma di azioni (es. cancellazione file, invio file in stampa, ...)
 - Navigazione (es. suoni diversi per identificare contesti diversi)

Earcon

Una "icona acustica": un breve suono o sequenza di suoni, ben riconoscibile, usato come simbolo

- suoni presenti in natura
- suoni "astratti"

Esempi: Windows, Sonic Finder di macintosh

Cinestesia

- La percezione della posizione del proprio corpo e dei propri arti, attraverso recettori situati nelle articolazioni (di tre tipi diversi)

Haptics

- "Haptic": tattile
- "Haptics": l'applicazione del senso del tatto all'interazione con il computer

Esempio: feedback tattile della barra braille

Device di output, che trasforma il contenuto di una riga del monitor in un testo Braille a rilievo. Ci sono barre da 80, 40, 20 celle (=numero di crt)

La legge esponenziale della pratica

Il tempo necessario per effettuare un compito diminuisce con la pratica. In particolare, il tempo T_n per effettuare un compito all' n -esima prova è dato da:

$$T_n = T_1 n^{-\alpha}$$

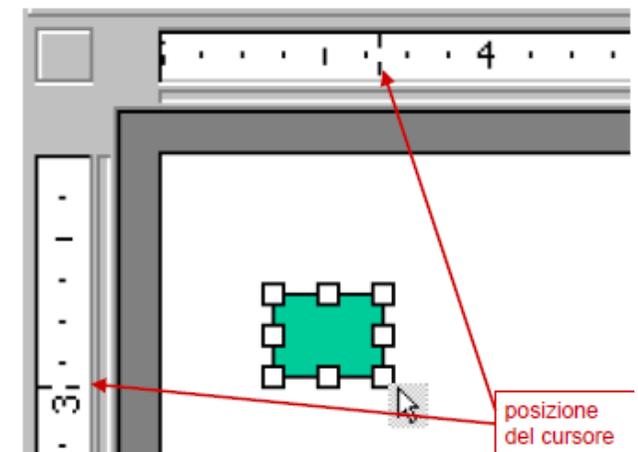
dove $\alpha \cong 0.4$ [0.2 ~ 0.6]

La legge di Fitts: $T = 100 \log_2(D/S + 0.5)$

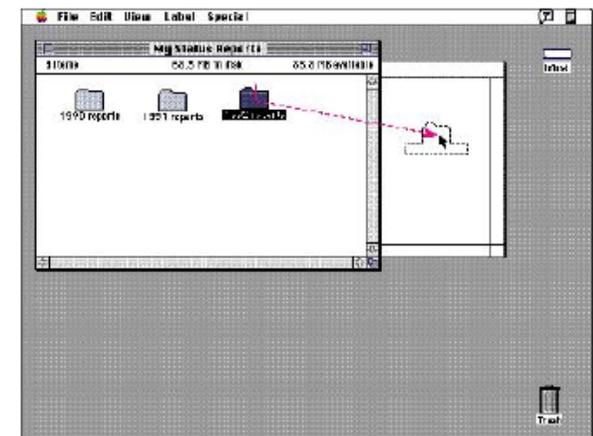
Esempio: In molti casi il puntatore del mouse tende a restare nelle vicinanze della scrollbar. I bottoni di uso più frequente dovrebbero quindi essere posti vicino alla scrollbar.



Feedback quantitativo: esempio

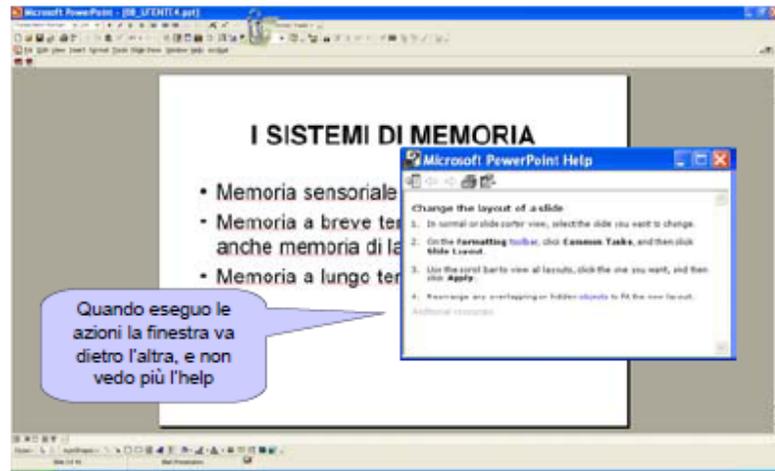


Feedback qualitativo: esempio



2.3. Memoria a Breve Termine

Esempio: help in Windows



Esempio: messaggi di errore sul web



Esempio

Carico cognitivo e Memoria a breve termine (MBT)

- Capacità limitata: 7 ± 2 "unità" ("chunk") (G. Miller, 1956)
- Persistenza dell'informazione breve: 15-30 sec
- La persistenza aumenta con la ripetizione (richiede attenzione)
- Usata per memorizzazione temporanea nei processi cognitivi
- Nuovi input ricoprono i contenuti precedenti ("interferenza")
- Teorie recenti non la considerano un deposito passivo, ma un sistema di funzioni ("working memory")

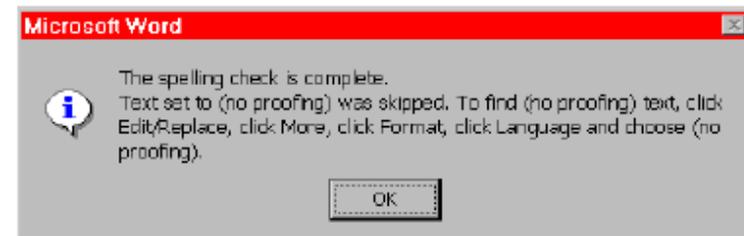
La quantità di "attività mentale" richiesta in un certo istante può essere rilevata attraverso la misura della variazione della dilatazione della pupilla.

Misure sperimentali suggeriscono che:

- differenti tipi di compiti causano differenti carichi cognitivi
- il carico diminuisce durante la transizione da un compito all'altro

MBT: implicazioni sul design

- evitare di sovraccaricare la MBT dell'utente, richiedendogli di memorizzare solo elementi significativi o familiari, in numero limitato
- minimizzare comunque il ricorso alla MBT dell'utente, in presenza di altre attività cognitive (interferenza)
- l'ansietà riduce le prestazioni della MBT: evitare situazioni di stress per l'utente



Da: Microsoft Word 97

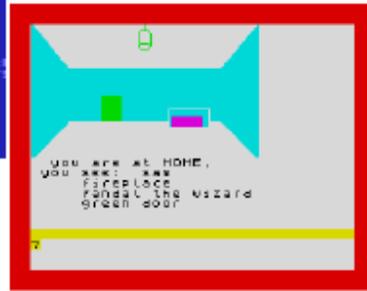
```

AT YOUR FEET IS A PIT FROM WHICH MIST
COILS UPWARD
What now?
You are at the east end of the Hall
of Hissay, filled with wads of cloud
which sway to and fro with a slow
There are passages on both sides and
a narrow path leading up to a wall
opening above. A grand stone
staircase leads downwards to the
south.
What now?
You are on the east bank of a wide
fissure sitting completely across to
Hall of Hissay. The white fog is very
thick here and the fissure is much
too wide to jump.
What now?
Climb?
Climb! A crystal bridge appears
spanning the chasm.
What now?

```

Colossal Adventure, 1983, ZX Spectrum

Rievocazione
(devo ricordarmi i comandi possibili)



Ringmaster, 1984?, ZX Spectrum



Riconoscimento
(i comandi possibili sono mostrati sul video)



Uninvited, 1986 (Mindscape, Macintosh)

MLT: implicazioni sul design

- Preferire il ricorso al riconoscimento piuttosto che alla rievocazione
- Presentare l'informazione ripetutamente
- Costruire associazioni forti, ripetute e distinte, che non interferiscano fra di loro
- Utilizzare associazioni visive

Aspetti connessi con la salute

Lavorare con i computer non è intrinsecamente pericoloso, ma a lungo andare possono sorgere dei problemi:

- Sforzi fisici, temperatura, luce, rumore possono avere alla lunga effetti nocivi sul nostro corpo. In particolare esistono malattie alle mani e agli occhi caratteristiche dell'uso dei computer.

Uso dei colori

- Non solo le nostre percezioni sono limitate (per esempio nel numero di colori singoli identificabili), ma esistono molte variazioni individuali.
- Moltissime persone hanno difficoltà a distinguere colori agli estremi della gamma (ad es., blu e nero), e molte hanno altri tipi di deficienze (ad es., daltonismo).
- E' opportuno dunque non usare mai i colori come unica differenziazione e MAI in contrasto con le aspettative culturali locali.
- Un trucco per verificare la leggibilità dei propri design per persone con problemi di colori è visualizzarli su uno schermo a toni di grigio.

Esempio: messaggi di errore sul web

Se hai bisogno di aiuto clicca ?
I campi contrassegnati da * sono obbligatori.

Nome

Cognome

Indirizzo **Manca l'indirizzo**

Codice Postale **Manca il codice postale**

Città **Manca la città**

Provincia **Manca la provincia**

Nazione

E-mail

Telefono **Manca il recapito telefonico o e' invalido**

Altro telefono

Data di nascita / /

Questa è una soluzione ancora migliore

2.4. Gestalt, o psicologia della percezione visiva

“Nella percezione visiva, il tutto è più della somma delle sue parti”

Le leggi della gestalt (M.Wertheimer, 1923)

- Legge della vicinanza
- Legge della somiglianza
- Legge della chiusura
- Legge della “curva buona”
- Legge della “buona forma”
- Legge dell’esperienza

Legge della vicinanza

Le parti di un insieme percettivo vengono raccolte in unità conformi alla minima distanza, a parità di altre condizioni.

Legge della somiglianza

Di fronte a una moltitudine di elementi diversi, vengono raccolti in gruppo gli elementi fra loro simili - a parità di altre condizioni.

Legge della chiusura

Le linee delimitanti una superficie chiusa si percepiscono come unità più facilmente di quelle che non si chiudono, a parità di altre condizioni.

Legge della “curva buona”

Quelle parti di una figura che formano una “curva buona” o che vanno nella stessa direzione si costituiscono in unità più facilmente delle altre.

Legge della “buona forma”

Il campo percettivo si segmenta in modo che risultino entità per quanto possibile equilibrate, armoniche, costituite secondo un medesimo principio in tutte le loro parti.

Legge dell’esperienza

L’esperienza modella le nostre impressioni.

L' esempio precedente, ristrutturato

2.5. Religioni informatiche

I computer non sono rigidi nelle loro operazioni. E' possibile fargli fare quasi tutto , se il programma è fatto in modo sufficientemente intelligente e correttamente. Il problema è che stiamo usando vecchie parole, vecchi concetti, senza tentare di progredire:

- Ci hanno detto che la gente aveva bisogno di oggetti familiari per apprendere l'uso dei computer, ed ecco che abbiamo icone, finestre, cartelle e cestini della spazzatura in un mondo piatto e bidimensionale.
- Ci hanno detto che il modo giusto e inevitabile di organizzare i nostri dati sulle memorie permanenti è in una gerarchia di contenitori indipendenti e sconnessi, il file system gerarchico. Il problema è che i designer si abituano a pensare che è così che debba andare il mondo. Certe idee vengono incastonate nei progetti, e certe altre lasciate fuori.

Ad esempio, *connessione* è lasciata fuori: Non c'è un modo generale per associare un commento ad un file del mio computer, o per creare connessioni persistenti tra due file. No: i file sono nudi e sconnessi. Analogamente per backup, confronto, condivisione. Purtroppo alla base ci troviamo il meccanismo del file system, con la sua crudezza, rigidità ed inflessibilità. Ci richiedono di spezzare i nostri pensieri in file monotipati, di organizzarli in directory gerarchiche, di forzarle in una struttura che, semplicemente, non è adatta. Nel 1974, Nelson parlava di come i computer, per come erano stati usati fino ad allora, erano serviti per opprimere la gente. Con i PC, se non stiamo attenti, sarà possibile essere oppressi nel nostro salotto. Ebbene, si è tutto avverato!

"Fra circa cinquant'anni sarà possibile programmare i computer ... in modo tale da farli giocare all'"imitation game" così bene che un interrogante medio non avrà più del 70% di possibilità di fare una identificazione corretta dopo 5 minuti di domande". (Turing, 1950)

2.6. The Loebner Prize

- Una gara annuale di programmi che si sottopongono al test di Turing
- Chi supera il test riceve \$100.000
- Finora nessuno ha superato il test
- <http://www.loebner.net/Prizef/loebner-prize.html>

Loebner prize: il migliore dell'edizione 2005

CONFEDERATE: Hello, my name is Sam and I am the human.

JUDGE: Hi Sam, tell me which is larger, a TV screen or Belgium.

CONFEDERATE: Belgium is a little bit larger.

JUDGE: Wljs0 lwjejd0 wleje elwje0 wjeol, isn't it?

CONFEDERATE: Is that Polish?

JUDGE: Who.

CONFEDERATE: What you wrote above.

JUDGE: This is ljelwl, aren't you.

CONFEDERATE: I'm a little worried about your writing skills.

JUDGE: Which is further away from this apartment, Queens or the Queen of England.

CONFEDERATE: If the Queen is visiting Manhattan, she is closer.

JUDGE: If the Queen is visiting Manhattan, won't she want to abdicate?

CONFEDERATE: I would.

JUDGE: You would rather not be Queen, then, right?

CONFEDERATE: I'd like to be Queen, but England's damp. How about you?

JUDGE: Which is damper, England or under your sink?

CONFEDERATE: I'd like to be Queen, but England's damp. How about you?

JUDGE: You didn't answer the question.

CONFEDERATE: You didn't answer my question. But please look at my response again.

JUDGE: Change of subject: if you are human, you should know: what is lewleoxoe.

CONFEDERATE: Perhaps I should, but my Polish is rusty.

JUDGE: How can a language be rusty, there is no metal in it?

CONFEDERATE: You write like a computer.

JUDGE: Thanks a real compliment. You write like a human!

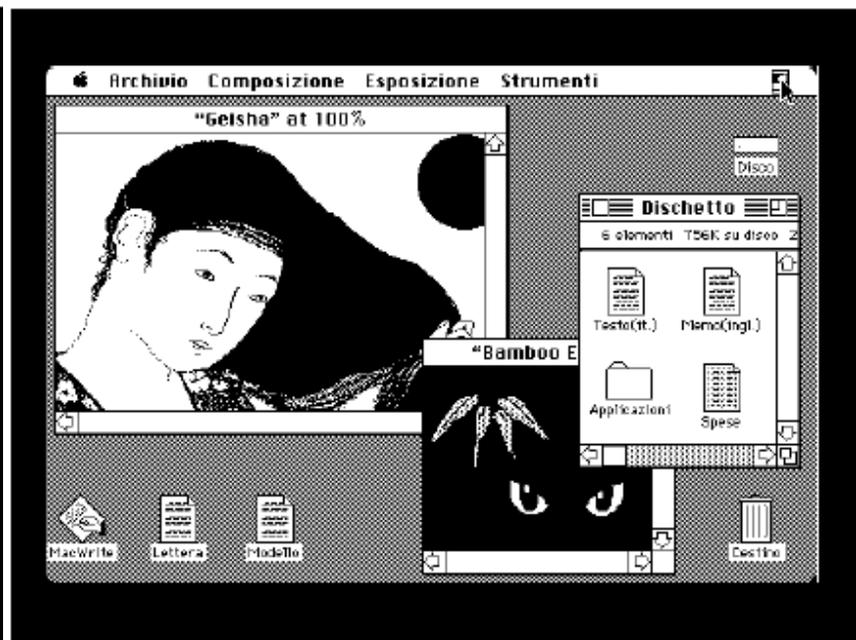
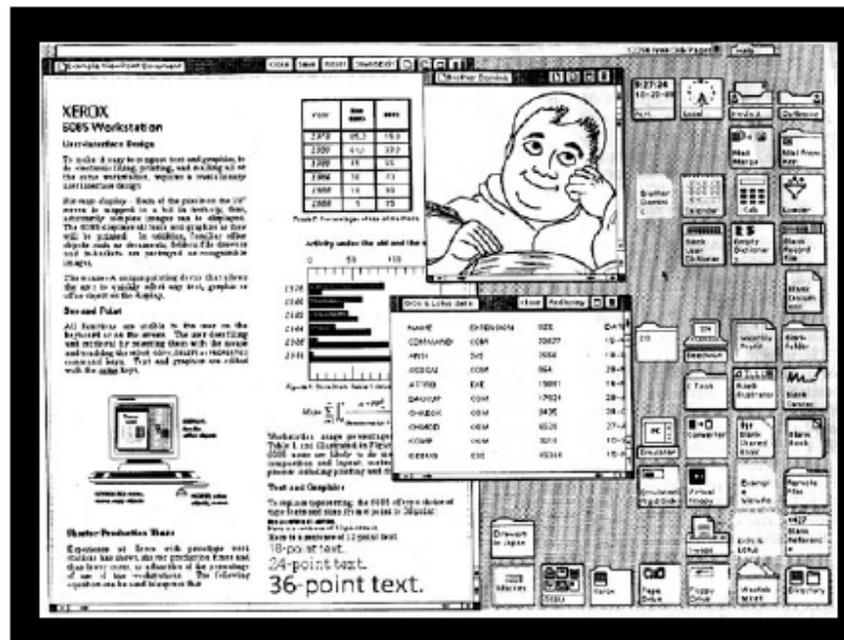
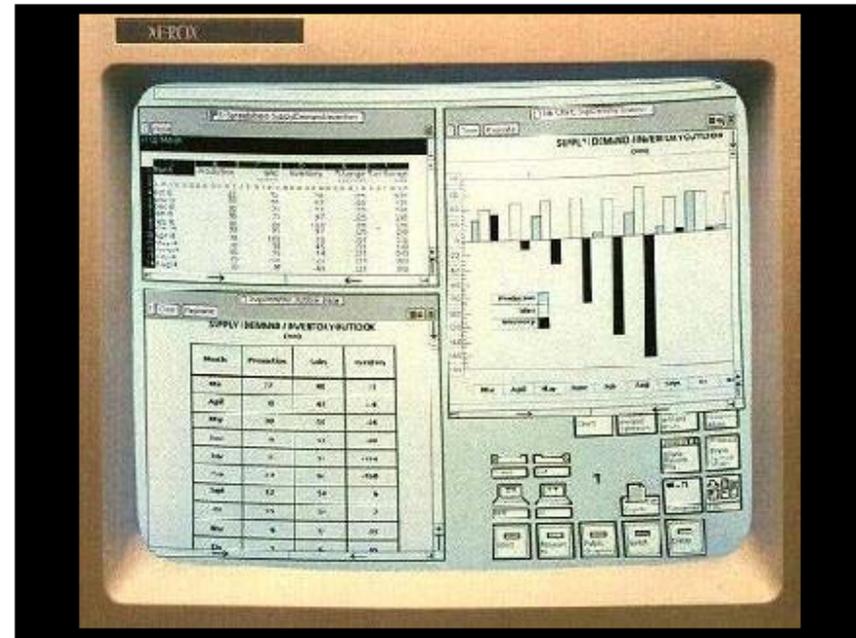
CONFEDERATE: Is that a compliment?

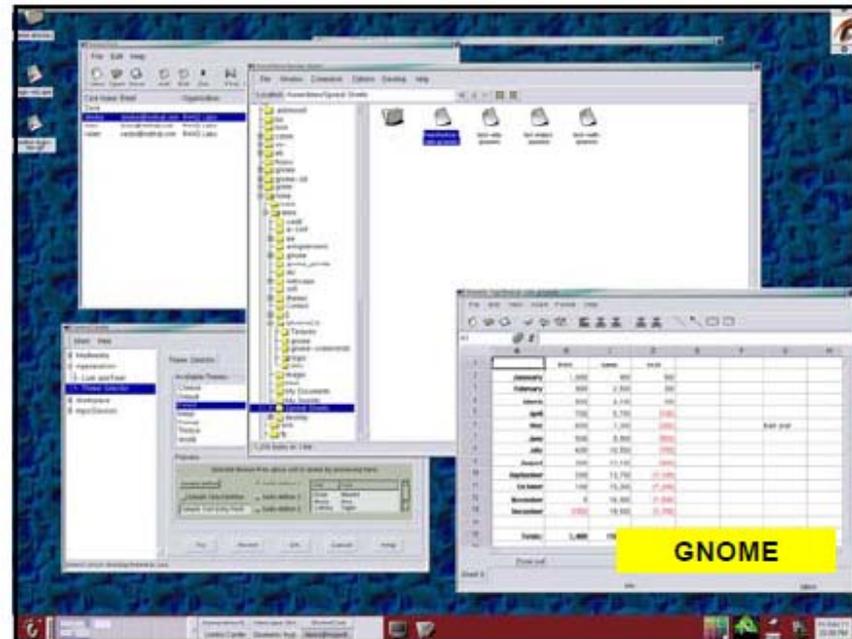
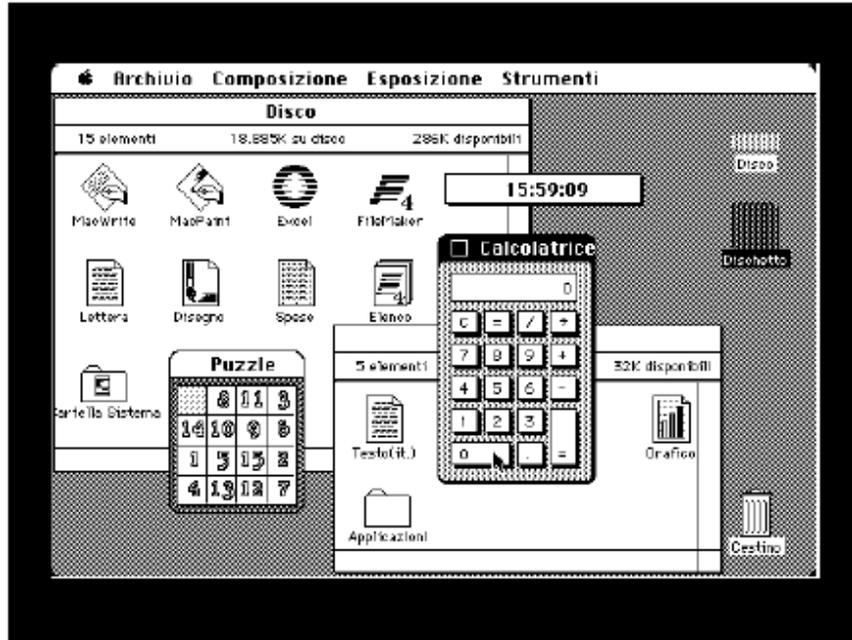
JUDGE: Good thing I am the judge, not you. You would say I was a machine!

CONFEDERATE: There may be worse things to be.

JUDGE: Which, a machine or a judge?

CONFEDERATE: Ha ha. I'll leave that you to. You're the judge!





2.7. Riconoscimento della scrittura

Handwriting recognition

- Problema molto complesso; le differenze individuali sono molto grandi
- Non ci si può aspettare un riconoscimento al 100%, e gli utenti dovranno cercare di scrivere con chiarezza: tassi di riconoscimento del 97% sono accettabili

Riconoscimento della scrittura: tecniche

- Riconoscimento on-line / off-line
 - off-line: analisi effettuata quando il testo è stato scritto;
 - on-line: analisi effettuata in tempo reale durante la scrittura
- Riconoscimento di corsivo / stampatello
 - lo stampatello è molto più facile da riconoscere
 - lo stampatello è stato standardizzato (ANSI OCR handprinting standard, 1974)
- Input "boxed" / "free-form"
 - incasellare i singoli caratteri per semplificarne la individuazione
- Riconoscimento di lettere / parole
 - in corsivo è più facile riconoscere intere parole (con dizionario) che non le singole lettere
 - necessità di avere dizionari personalizzabili
- Addestramento del riconoscitore
 - ogni utente ha il suo stile personale; il riconoscitore deve imparare durante l'uso;
 - per questo deve ricevere dei feedback
- Utilizzo di informazioni di contesto
 - il riconoscimento del contesto può facilitare il riconoscimento di caratteri e di parole
- Necessità di mescolare testo e grafica
- Set di caratteri ASCII / Unicode
- Rappresentazione bitmapped / vettoriale

Il progenitore: Apple Newton (1993)

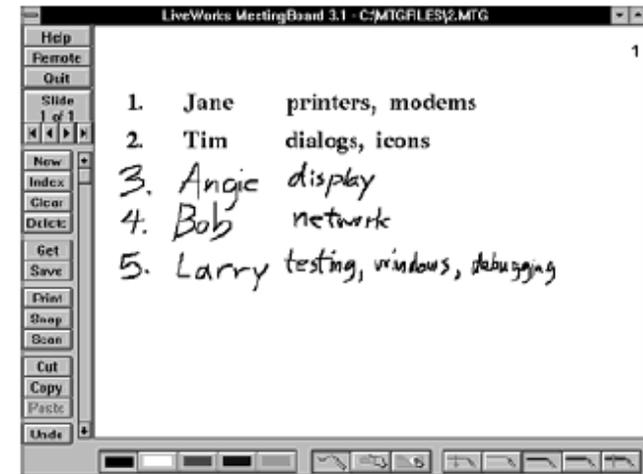
- Progetto iniziato nel 1990 da John Sculley in Apple
- lanciato sul mercato nel 1993
- riconosceva parole in corsivo, sulla base di un dizionario estensibile, con addestramento
- troppo in anticipo sui tempi, la tecnologia di riconoscimento era fortemente immatura
- fu un flop commerciale

Handwriting: un approccio algoritmicamente molto più semplice...

Unistrokes (Goldberg & Richardson, PARC, 1993)

- Ogni simbolo è rappresentato da un singolo tratto di penna
 - Il software lo discrimina facilmente
 - L'utente lo impara facilmente
- Disaccoppiamento di scrittura e lettura (le posizioni sullo schermo sono differenti)
- Si scrive un carattere alla volta

Un esempio per il lavoro cooperativo: Xerox MeetingBoard (1995)



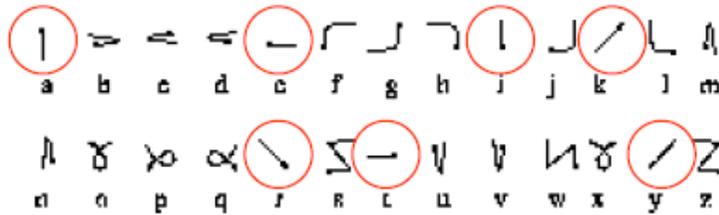
Il Tablet PC (dal 2002...)



Basato sulla piattaforma software di Microsoft (Windows XP Tablet PC Edition)

Unistrokes

5 segni base: | > < S L
 x 4 orientamenti x 2 direzioni
 = 40 possibili variazioni



Graffiti

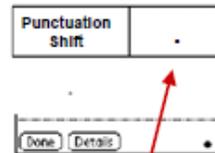
- una evoluzione di Unistrokes
- realizzata da Palm Computing per il Palm Pilot
- comprende molti segni di interpunzione e caratteri speciali
- il primo di "palmare" grande successo commerciale



The Graffiti® alphabet

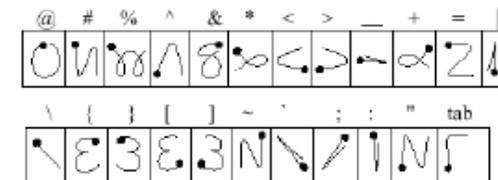
Letter	Strokes	Letter	Strokes
A	^	N	N
B	B B	O	O O
C	C	P	p P
D	D P	Q	O
E	E	R	R R
F	F F	S	S
G	G G	T	7
H	h	U	U
I	I	V	V V
J	J	W	W
K	k	X	X X
L	L	Y	Y Y
M	M M	Z	Z

Space	←	Back Space	→
Carriage Return	↵	Period	⋅
Caps Shift		Caps Lock	
Number	Strokes	Number	Strokes
0	0 0	5	5 5
1	1	6	6
2	2	7	7
3	3	8	8 8
4	4	9	9



Punctuation shift:
 per entrare nel
 "modo
 punteggiatura"

Symbol	Stroke	Symbol	Stroke
Period	⋅	Dash	—
Comma	,	Left Paren	(
Apostrophe	'	Right Paren)
Question	?	Slash	/
Exclamation	!	Dollar	\$



2.8. Gesture recognition

The approach of teaching the user a number of gestures that can be recognized is also known as scribble matching [Fonseca and Jorge, 2000] and is often used in drawing tools [Julia and Faure, 1995] graphical editors, or handwriting applications where users have to learn a new alphabet comprising shapes that are relatively easy to recognize (cfr. The unistroke alphabet from Goldberg [Goldberg, 1989]). The recognition of 2D shapes has also extensively been studied in the context of computer graphics and image retrieval [Velkamp and Hagedoorn, 2000]. This research has resulted in new feature extraction techniques for object recognition and a theory called curvature scale space [Mokhtarian et al., 1996, Lindeberg, 1998], which are now being considered for the Vind(x) image retrieval system developed at the NICI [Schomaker et al., 1999, Vuurpijl et al., 2002b]. Vind(x) exploits a "new use for the pen", i.e. object recognition through outline sketch.

Given this wealth of techniques for recognizing sketches, outlines or scribbles, it is not unthinkable that natural gesture recognition is a solvable problem indeed. For a limited amount of gesture classes (<10), recognition rates of 95% to 100% are very well feasible. Although scribble or outline matching may be a useful approach if the user is willing to produce one out of a set of predefined shapes, in COMIC we will mainly consider natural gesture recognition. Gestures produced in a natural dialog between two humans or between a human and the machine, will most probably introduce more classes and a much higher within-class/between-subject variability than in the context of constrained gesture recognition. Although the work performed by Breteler [Klein Breteler, 2001] suggests otherwise. Please note that her study studies functional behavior (e.g., hitting a light switch), while gestures are of a representational nature (e.g. iconic gestures representing a trajectory).

As a first step towards natural 2D pen-based gesture recognition, we will explore the penbased input data to be collected in the human-human experiments that are now being conducted in SLOT, with the goal to detect a limited set of salient gesture shapes. This work will be performed by WP2. Once such a set of gesture classes is distinguished, it will be examined which gesture decoders and what feature schemes are most suitable for the reliable recognition of the selected gesture classes. In the section about human factors experiments, we introduce our plans to work on basic multi-modal experiments in the context of the graphical manipulation of iconic or geometric objects. It is expected that the techniques for constrained gesture recognition may be of use for pen-based gesture recognition in this context.

Towards natural 3D gesture recognition

3D gesture recognition differs from 2D as it requires different data acquisition hardware, different signal processing, segmentation and feature extraction algorithms. Whereas AHR builds on several decades of research, the first papers about automatic 3D gesture recognition appeared only after 1985. The main reason for this is that the input devices required for data capture (such as the data glove, video trackers or optotrak) still had to be invented. Similar to natural 2D gestures, the gesture shapes that humans produce in a certain context can only be determined by thorough examination, e.g. via transcription of video material. Please note that the automatic recognition of natural gestures is considered as an unsolved problem. Even humans are not good at recognizing gestures as was shown in [P.Feyereisen et al., 1988], where people found it very difficult to attribute meaning to gestures based on shown videotapes, without speech or contextual clues.

However, a large number of research programs are currently being undertaken to recognize a limited number of gestures. An overview of the work in this field is presented in [Cohen, 1999]. In most cases, gesture recognition is performed in the context of virtual reality applications and control, where recognition from a selected gesture repertoire is considered as feasible. In [Cohen, 1999], a number of architectural guidelines are specified, where the first requirement is: "Choose gestures which fit a useful environment", indicating that the research towards natural gesture recognition is still in its infancy. Furthermore, note that the design of a suitable set of gestures that are (i) easy for computers to recognize and (ii) for humans to learn, remember and reproduce is considered as particularly difficult [Jr. et al., 1999].

Gesture classes

As a first step towards natural gesture recognition, an inventory of the possible gesture classes is required. In [Benoit et al., 1998, McNeill, 1992, Nespoulos et al., 1986, Ruiter, 2000], various taxonomies are proposed. For the automatic interpretation of gestures, the pragmatic approach as discussed by de Ruiter [Ruiter, 2000] is adopted here:

1. Iconic gestures: depicting aspects of the accompanying speech topic. This category includes what McNeill [McNeill, 1992] calls metaphoric gestures, because from the perspective of gesture production it is of no relevance whether the imagery underlying the gesture is related to abstract or real entities. An example of an iconic gesture is a downward corkscrew motion of the hand while talking about a vortex.
2. Pantomimes: gestures that are imitations of functional motor activities. Example: pretending to hold a telephone while talking about a phone conversation.
3. Deictic gestures: Pointing gestures. Example: pointing somewhere and say "over there".
4. Beat gestures: Biphasic movements of the hands of finger that do not represent anything. Often clearly seen when speakers are agitated and trying to convince their audience of something.
5. Emblems: Gestures whose form-meaning relation is lexicalized. Example: the ring-shaped gesture formed by the thumb and index finger to indicate "OK".

The possibilities for automatically registering and interpreting real (i.e. naturally occurring) 3D gestures are seriously constrained by our limited understanding of how humans use and interpret gestures. As Benoit observed: "finding sets of useful gestures will probably remain an application specific development effort until gesture-based interaction will be understood in depth". However, following the gesture taxonomy presented above, a number of observations about the likelihood of successfully recognizing 3D gestures (AGR) can be made.

First, the best candidates for AGR are emblems and deictic gestures. Emblems are relatively easy to recognize because they are lexicalized, meaning that one can train recognizers on a specified lexicon of emblem-gestures, similar to ASR where recognizers can be trained on a corpus of annotated utterances. Deictic gestures are relatively easy to interpret because of their simple semiotic properties, and their predictable relation to the concurrent speech. They can often be recognized by the hand-shape, (the pointed finger) and a 3-dimensional vector with the speaker as origin can easily be extracted from the recorded motion data. Far more difficult to recognize are iconic and beat gestures. Iconic gestures are by definition not lexicalized [McNeill, 1992], meaning that they are not governed by linguistic conventions. Usually, they are very hard to interpret (if at all) without the accompanying speech.

The "iconic gesture recognition" reported in the literature [Sowa and Wachsmuth, 2001], are usually about limited sets of predefined iconic gestures, implying that they are artificially lexicalized. In other words, they behave like (local) emblems. Finally, beat gestures are difficult to handle because nobody seems to have a clue about what their communicative function is. McNeill [McNeill, 1992] claims that they have a "metanarrative" function, but his evidence is limited and dependent on a highly subjective interpretation of the context surrounding the utterances he collected.

3D data capture and recognition

Research will be conducted towards the recognition of a limited number of salient 3D pen-based gesture classes. It is expected that the analysis of the proposed human-human experiments will yield at least a workable taxonomy. Given such a set of (unconstrained, or natural) 3D gesture classes, it is intended to perform a small experiment to collect real 3D data using the optotrak 3020 system available at the NICI. This system is wall-mounted and uses three optical cameras and special hardware for capturing 3D data emitted through synchronized lightemitting active markers. The technical specifications [Inc., 1991] state a spatial accuracy of 0.15 mm and a resolution of 0.01 mm at a distance of 2.5 m between markers and the cameras. The maximum marker rate is 3500Hz, which has to be divided by the number of active markers employed.

The NICI has extensively used the optotrak for examining human motor production with up to twelve markers, which still sustain a data acquisition rate of about 200Hz. Note that this is rather impressive at the given resolution. It is yet unclear how experiments for 3D data collection will be set

up. Experience has shown that users can properly perform targeted movements. How active marker technology (implying that each marker is connected by a thin wire to the hardware) restricts the unconstrained production of 3D gestures is an issue to be considered.

Several attempts towards 3D gesture recognition can be distinguished here:

- For the recognition of certain classes of deictic gestures, it will be examined whether a proper projection from 3D to 2D yields 2D-trajectories that can be recognized with the given 2D gesture recognition algorithms.
- Attempts will be made to change the existing 2D algorithms such that they are suited for the recognition of 3D trajectories. Please note that this involves a major investment in both theory, experimental evaluations and software development. For example, as was indicated above, 2D trajectories are generally normalized on the basis of holistic word-shape features. How could the baseline detection algorithm that forms the core of the normalization process be used in 3D?
- The development of new 3D pattern recognition techniques, including the evaluation of the required feature extraction algorithms. Based on a further review of the literature and the expected collection of training data, attempts will be made to investigate this approach.
- It may become possible to describe a set of well-defined prototypical gestures in terms of elementary 3D movements. Similar to the grammar descriptions for describing character shapes, this requires some notion of velocity-based strokes in 3D. The expectation is that it will be feasible to adapt the existing velocity-based segmentation algorithms to 3D.

The OGI team has investigated several applications where 3D gestures were considered. In a review paper from Oviatt et al [Oviatt et al., 2001], it is stated that 3D gesture recognition is still beyond the state of the art. In COMIC, the joint cooperation between experts from cognitive psychology, speech and gesture recognition, computer science and artificial intelligence will hopefully provide the synergy to proceed some important steps to proceed beyond the state of the art.

Research in multimodal interaction with 3D gesture recognition is proceeding rapidly, but almost exclusively in the context of immersive applications. Much of this work is being carried out at OGI and at the MIT Media Lab. Some of those applications, such as the real estate agent REA of MIT Media Lab, may acquire a role in the field of eCommerce and eWork. Entertainment is, of course, another market segment where immersive applications will flourish.

2.9. Assistive technology

Assistive Technology (AT) is a generic term that includes assistive, adaptive, and rehabilitative devices and the process used in selecting, locating, and using them. AT promotes greater independence for people with disabilities by enabling them to perform tasks that they were formerly unable to accomplish, or had great difficulty accomplishing, by providing enhancements to or changed methods of interacting with the technology needed to accomplish such tasks. According to disability advocates, technology, all too often, is created without regard to people with disabilities, and unnecessary barriers make new technology inaccessible to hundreds of millions.

Universal (or broadened) accessibility, or universal design means excellent usability, particularly for people with disabilities. But, argue advocates of assistive technology, universally accessible technology yields great rewards to the typical user; good accessible design is universal design, they say. The classic example of an assistive technology that has improved everyone's life is the "curb cuts" in the sidewalk at street crossings. While these curb cuts surely enable pedestrians with mobility impairments to cross the street, they have also aided parents with carriages and strollers, shoppers with carts, and travellers and workers with pull-type bags, not to mention skateboarders and inline skaters.

Consider an example of an assistive technology. The modern telephone is not accessible to people who are deaf or hard of hearing. Combined with a text telephone (also known as a TDD [Telephone Device for the Deaf] and in the USA generally called a TTY [TeleTYpewriter]), which converts typed characters into tones that may be sent over the telephone line, the deaf person is able to communicate immediately at a distance. Together with "relay" services (where an operator reads what the deaf person types and types what a hearing person says) the deaf person is then given access to everyone's telephone, not just those of people who possess text telephones. Many telephones now have volume controls, which are primarily intended for the benefit of people who are hard of hearing, but can be useful for all users at times and places where there is significant background noise. Another example: calculators are cheap, but a person with a mobility impairment can have difficulty using them. Speech recognition software could recognize short commands and make use of calculators a little easier. People with cognitive disabilities would appreciate the simplicity; others would as well. Toys which have been adapted to be used by children with disabilities, may have advantages for "typical" children as well. The Lekotek movement assists parents by lending assistive technology toys and expertise to families. Telecare is a particular sort of assistive technology that uses electronic sensors connected to an alarm system to help caregivers manage risk and help vulnerable people stay independent at home longer. A good example would be the systems being put in place for senior people such as fall detectors, thermometers (for hypothermia risk), flooding and unlit gas sensors (for people with mild dementia). The principle being that these alerts can be customised to the particular person's risks. When the alert is triggered, a message is sent to a carer or contact centre who can respond appropriately. The range of sensors is wide and expanding rapidly. Technology similar to Telecare can also be used to act within a person's home rather than just to respond to a detected crisis. Using one of the examples above, unlit gas sensors for people with dementia can be used to trigger a device that turns off the gas and tells someone what has happened. This is safer than just telling an external person that there is a problem.

Designing for people with dementia is a good example of where the design of the interface of a piece of assistive technology (AT) is critical to its usefulness. It is important to make sure that people with dementia or any other identified user group are involved in the design process to make sure that the design is accessible and useable. In the example above, a voice message could be used to remind the person with dementia to turn off the gas himself, but who's voice should be used, and what should the message say? Questions like these must be answered through user consultation, involvement and evaluation.

Assistive technology products: Accessing the keyboard

Hardware

Height-adjustable trolley. Allows precise adjustment, via a handle, of a computer to the correct height - even when fully loaded with equipment.

Adjustable keyboard and monitor arms. Attaches to a table and swings out to the correct height.

Footrests. If the user's feet are dangling down from a seat that is too high they can help to maintain sitting balance and achieve good hand control.

Wristrests. Stabilises the arm.

Arm supports. Take the weight of the arm and let the user move across the keyboard to access the keys.

Keyguards. It fits over the keyboard to help prevent unintentional keypresses.

Software

Sticky keys is a feature of Microsoft Windows 95 onwards and Mac OS X operating systems. Allows characters or commands to be typed without having to hold down a modifier key (Shift, Ctrl, Alt) while pressing a second key.

FilterKeys is a feature of Microsoft Windows 95 onwards. Instructs the keyboard to ignore brief or repeated keystrokes.

FilterKeys (Ignore repeated keystrokes) is a feature of Microsoft Windows 95 onwards. Adjusts the amount of time that elapses before a character repeats when you hold down a key.

Accessing the mouse

Hardware

Wristrests and supports. Wristrests support the hand so that there is less strain on the wrist; arm supports support the arm yet still permit free movement.

Keyboard shortcuts. A combination of keys is used for operations commonly done using the mouse.

Rollerballs. An upside-down mouse where the ball is moved directly with the fingers (or foot, nose, elbow etc) to control the mouse pointer.

Ergonomic keyboards. Reduce the discomfort from injuries related to excessive keyboard use (wristrests are often built in as well).

Chording keyboards. Have a handful of keys (one per digit per hand) to type by 'chords' which produce different letters and keys.

Expanded keyboards. Keys are larger and more widely spaced.

Compact and miniature keyboards. A smaller version of the standard keyboard.

Dvorak alternative keyboard layout. An ergonomic keyboard layout, where the most common keys are located at either the left or right side of the keyboard.

Height-adjustable computer table

FilterKeys (Ignore quick keystrokes and slow down the repeat rate) is a feature of Microsoft Windows 95 onwards. Adjusts the speed at which a character repeats when you hold down a key.

ToggleKeys is a feature of Microsoft Windows 95 onwards. A high sound is heard when the CAPS LOCK, SCROLL LOCK, or NUM LOCK key is switched on and a low sound is heard when any of those keys are switched off.

Joysticks. The stick is moved directly with the fingers (or foot, nose, elbow etc) to control the mouse pointer.

Graphics tablets. The stylus pen, linked to the tablet, is moved with the hand to control the mouse pointer.

Touch pads. Flat touch sensitive surface operated with the finger to control the mouse pointer.

Touch screens. The user can touch objects directly on the screen to move them around.

Head control. Use head movement to control the mouse pointer.

Software

Double-click speed is a feature of Microsoft Windows.

ClickLock is a feature of Microsoft Windows. Locks a mouse button down so that items can be dragged around the screen.

Mouse speed is a feature of Microsoft Windows. Sets how far and fast the pointer moves on the screen.

Pointer size is a feature of Microsoft Windows. Changes can be made to the size, colour or shape.

Switch access and switches

Switches allow access to computers for people with severe physical or cognitive difficulties. Mimics a keyboard keystroke or mouse click.

Switch interface. Used as an interface to connect a switch to the computer.

Visual impairment

Hardware (choice of appropriate hardware will depend on the users level of functional vision)

Full keyboard stickers. High-contrast, enlarged print adhesive keyboard stickers - white on black, black on white and black on yellow sets available.

Large print keyboard. High-contrast, enlarged QWERTY keyboards - white on black, black on white and black on ivory versions available.

Locator dots. Help locate 'home' keys, i.e. F and J, on the keyboard.

Wheel mouse. Moves mouse pointer, plus middle wheel button is used to scroll up and down.

Large screen monitor. Produces a larger-than-normal image.

Adjustable task lamp. Lamp, using fluorescent bulb, shines directly onto the paper and can be adjusted to suit.

Copyholder. Holds printed material in near vertical position for easier reading and can be adjusted to suit.

Closed Circuit Television system (CCTV)/Video Magnifier. Printed materials and objects are placed under a camera and the magnified image is displayed onto a screen.

Modified cassette recorder. To record a lecture, own thoughts, ideas, notes etc.

Desktop compact cassette dictation system. To allow audio cassette playback with the aid of a foot pedal.

MouseKeys is a feature of Microsoft Windows 95 onwards. Configures the numeric keypad at the right-hand side of the keyboard to control the mouse pointer.

Macros is a feature of Microsoft Word. A particular sequence of keyboard operations are 'recorded' and then 'played back' by pressing a combination of keys.

Switch mounting. Used to mount a switch (sometimes in conjunction with a mounting plate).

Switch accessible software. Software which is accessible via a switch.

Fuser. To produce tactile materials, e.g. diagrams, by using heat and special swell paper.

Standalone Reading Aid. A unit which integrates a scanner, Optical Character Recognition (OCR) software and speech software. The printed document is scanned and read by the same machine.

Scanner. A device used in conjunction with Optical Character Recognition (OCR) software. The printed document is scanned and converted into electronic text, which can then be displayed on screen as recognisable text.

Refreshable Braille display. An electronic tactile device which is placed under the computer keyboard. A line of cells, that move up and down to represent a line of text on the computer screen, enables the user to read the contents of the computer screen in Braille.

Electronic Notetaker. A portable computer with a Braille or QWERTY keyboard and synthetic speech. Some models have an integrated Braille display.

Braille embosser. Embosses Braille output from a computer by punching dots onto paper. It connects to a computer in the same way as a text printer.

Perkins Brailier. To manually emboss Grade 1 or 2 Braille.

Software (choice of appropriate software will depend on the users level of functional vision)

Customised Microsoft Windows. Changes the desktop, short-cut icons, menu bars and scroll bars.

Customised mouse pointers and toolbars. Changes the size, colour and/or shape of mouse pointers and the size of toolbars.

Screen magnifier. Enlarges the image on the computer screen. The level and type of magnification can be customised. Some programs also have supportive synthetic speech.

Alternative & Augmentative Communication (AAC)

Low-tech systems. Simple paper or object based systems, i.e. do not require a battery.

Light-tech systems. Simple voice output devices, which require a battery; although no sophisticated charging mechanism is required.

High-tech systems. Sophisticated voice output devices, which require a battery, as well as training and ongoing support.

D/deafness, Hearing loss

Audiometer

Fire alarm paging system

Loop system (portable and fixed)

Radio aids

Learning difficulties

Hardware

Keyboards with lowercase keys

Keyboards with big keys

Software

Age appropriate software

Cause and effect software

Hand-eye co-ordination skills software

Specific learning difficulties

Diagnostic assessment software

Spellcheckers

Study skills software

Touch typing software

Visual mapping software

Word prediction software

Screen reader. Converts what is happening on the computer screen into synthesised speech.

Optical Character Recognition. Converts the printed word into synthesised speech, via a scanner.

Braille translation. Converts the printed word into Braille, which can then be embossed via a Braille embosser.

Touch typing.

Keyboard shortcuts, such as Ctrl+S, Alt+F4, F7. A combination of keys are used for operations commonly done using the mouse.

Textphone

Televisions with Teletext

VCRs that can read and record subtitles

Symbols based software

Touchscreen

Word processor with speech software

Mini invasive measurements on man

Intracranial pressure
Tympanic temperature
Heart flow (oesophagus probes)
Endovascular probes
Coeloscopy
Foetal heart rate
Intra uterine pressure
Foetal ph
Oral temperature
Cardiac output (thermodilution)
Pressare, ph, emg
Autonomous telemetric capsule: ph, temperature
Glucose
Arterial pressare
Rectale temperature
Radioactive clearance (Kidney)
Electro dermal response
Electrocardiogram

Vehicles, Process room: control vigilance state

Eye movement and direction
Pressure and force mapping on the seat
Head tilting
Electrocardiogram
Electroencephalogram

Non invasive measurements on man

Voice analysis
Audiogram, oto sound
Ir spectroscopy
Indirect intracranial pressure
Lung, heart, muscle sound
Heat loss
Emotional response, vigilance, workload
Muscular activity: movements, speed, power
Urinary flow
Capillaroscopy
Nystagmography
Respiration: rate, amplitude, flow, volume, pressure, composition

Vessel blood flow: ultra sound Doppler

Vectocardiography

Skin: temperature, resistance, impedance, blood flow, hydration, colour

Tactile sensitivity

Plethysmography: optical, strain gauge, impedance

Body shapes and dimensions

Acceleration, force, balistography

3

capitolo



4 capitolo



4.0. Attuali tecnologie d'interfaccia: l'uomo ragiona per immagini

La vista è il senso più utilizzato nelle attività conoscitive, nell'esplorazione dell'ambiente che ci circonda, nella presentazione di concetti. L'uomo ragiona per immagini, ed è ovvio che il primo passo nella costruzione di un'interfaccia fosse proprio la sua presentazione visiva. Il problema della comunicazione uomo-macchina forse non sarebbe mai nato senza l'esistenza di uno schermo che fungesse proprio da interfaccia.

Le principali tipologie di schermi sono le seguenti:

Schermi a raggi catodici (CRT)

Il dispositivo predominante per l'output. Un fascio di elettroni è sparato da una fonte posta ad un'estremità di un imbuto, e viene diretto da campi magnetici. All'altro estremo c'è una superficie ricoperta di fosforo che viene eccitato dal raggio e riluce. Esistono tre tipi di tubi a raggi catodici:

- Raster scan: il cannone si muove sistematicamente da sinistra a destra, dall'alto al basso, per 30-100 volte al secondo, modificando di volta in volta il flusso di elettroni. Con l'interlacing vengono disegnate righe alternate, per ridurre il flickering. Fa buone righe orizzontali, verticali e a 45°, ma appaiono effetti a scaletta nelle linee con altri angoli e nelle curve.
- Random scan (directed beam refresh o vector display): Il cannone disegna esplicitamente ogni riga. Si riescono ad avere risoluzioni notevolissime, poco flicker e nessun effetto scaletta, ma i colori sono imprecisi e il display è molto costoso.
- Direct View Storage Tube (DVST): usato per lo più negli oscilloscopi analogici. Simile al random scan, esiste una serie di cannoni che mantengono il flusso diretto di elettroni per più a lungo, eliminando il flickering.

Display a cristalli liquidi (LCD)

Più piccoli, leggeri, piatti, consumano molta meno corrente dei CRT, ma costano di più. Uno strato di cristalli liquidi è posto tra due vetri, uno trasparente e polarizzato, l'altro (sotto) riflettente. Normalmente la luce passa attraverso il cristallo e viene riflessa (pixel bianco). Quando passa corrente presso un cristallo, il cristallo si volta di 90°, cosicché il riflesso non esce più dall'alto (pixel nero). Il cristallo si volta lentamente, quindi non ci sono effetti di flicker. Inoltre emette poca luce e poche radiazioni, riducendo l'affaticamento dell'occhio.

Inizialmente erano in b/n, la lentezza di rotazione del cristallo impediva animazioni credibili, e costavano molto. Ora sono a colori, veloci e a costi ragionevoli. Crearono il mercato degli orologi digitali negli anni 70. Negli anni 90 hanno creato il mercato dei portatili e dei PDA.

Display a 3D (3DD)

Tipicamente sono trucchi visivi nell'usare su schermi tradizionali: ombre, oggetti nascosti, effetti prospettici, ecc. Un effetto importante è la creazione dell'immagine stereoscopica, ottenuta generando due immagini quasi simili e facendole percepire indipendentemente dai due occhi.

- Occhiali 3D: due piccoli schermi CRT o LCD montati su un paio di occhiali o su un elmetto. Molto pesanti e scomodi.
- Occhiali a visione alternata: ogni lente viene bloccata da polarizzazione della luce o da corrente elettrica in sincrono con l'immagine trasmessa sullo schermo. Rende effetti 3D molto credibili.
- Schermi 3D: La superficie dello schermo è composta di prismi. Ogni lato si illumina in maniera indipendente. Oltre ad essere molto costosi, hanno un angolo di visione molto ristretto (anche solo allontanare la testa fa perdere l'effetto).

Ci sono però problemi di focalizzazione e di refresh. In particolare, una realtà virtuale immersiva DEVE avere tempi di reazione adeguati, o si ottiene disorientamento e malessere (mal di mare).

4.1. Touchscreen

A touch screen is an input device that allows the user to interact with the computer by touching the display screen. It was first used on a dedicated console with the Game.Com from Tiger. Nintendo popularised it for use in video games with the Nintendo DS; other systems including the Tapwave Zodiac as well as the vast majority of PDAs have also included this feature. Modern touch screens use a thin, durable, transparent plastic sheet overlaid onto the glass screen. The location of a touch is calculated from the capacitance for the X and Y axes, which varies based upon where the sheet is touched. Touchscreen panels are display overlays which are typically either pressure-sensitive (resistive), electrically-sensitive (capacitive), acoustically-sensitive (SAW - surface acoustic wave) or photo-sensitive (infra-red). The effect of such overlays allows a display to be used as an input device, removing the keyboard and/or the mouse as the primary input device for interacting with the display's content. Such displays can be attached to computers or, as terminals, to networks.

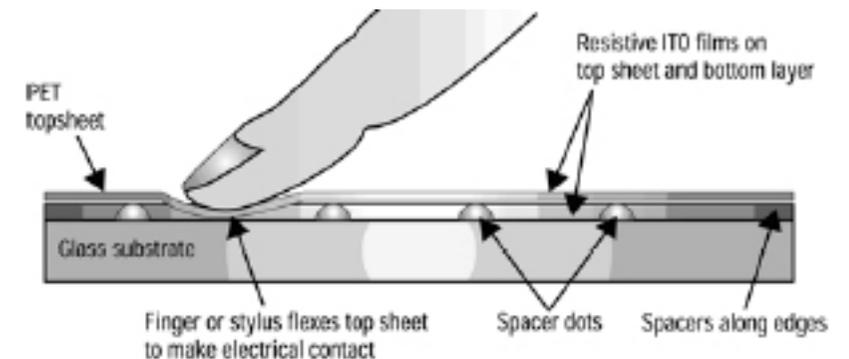
Touchscreens have become commonplace since the invention of the electronic touch interface in 1971 by Dr. Samuel C. Hurst. They have become familiar in retail settings, on point of sale systems, on ATMs and on PDAs where a stylus is sometimes used to manipulate the GUI and to enter data. The popularity of smart phones, PDAs, portable game consoles and many types of information appliances is driving the demand for, and the acceptance of, touchscreens. The HP-150 was among one of the world's earliest commercialized touch screen computers. It actually does not have a touch screen in the strict sense, but a 9" Sony CRT surrounded by infrared transmitters and receivers which detect the position of any non-transparent object on the screen. Touchscreens are popular in heavy industry and in other situations, such as museum displays, where keyboards and mice do not allow a satisfactory intuitive, rapid or accurate interaction by the user with the display's content. The Nintendo DS uses a touchscreen as a primary controlling device. Historically, the touchscreen sensor and its accompanying controller-based firmware have been made available by a wide array of after-market system integrators and not by display, chip or motherboard manufacturers. With time, however, display manufacturers and System On Chip (SOC) manufacturers worldwide have acknowledged the trend toward acceptance of touchscreens as a highly desirable user interface component and have begun to integrate touchscreen functionality into the fundamental design of their products.



There are a number of types of touch screen technology:

Optical Imaging: this is a relatively-modern development in touch screen technology, two or more image sensors are placed around the edges (usually the corners) of the screen. Infrared backlights are placed in the camera's field of view on the other sides of the screen. A touch shows up as a shadow and each pair of cameras can then be triangulated to locate the touch. This technology is growing in popularity, due to its scalability, versatility, and affordability, especially for larger units.

Resistive: a resistive touch screen panel is coated with a thin metallic electrically conductive and resistive layer that causes a change in the electrical current which is registered as a touch event and sent to the controller for processing. Resistive touch screen panels are generally more affordable but offer only 75% clarity and the layer

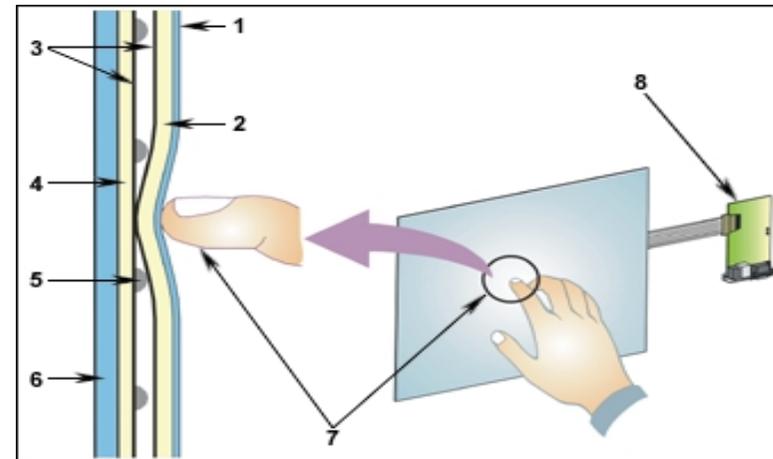


can be damaged by sharp objects (hardness in Mohs scale: 5). Resistive touch screen panels are not affected by outside elements such as dust or water and are the most commonly used today.

Resistive LCD touchscreen monitors rely on a touch overlay, which is composed of a flexible top layer and a rigid bottom layer separated by insulating dots, attached to a touchscreen controller. The inside surface of each of the two layers is coated with a transparent metal oxide coating (ITO) that facilitates a gradient across each layer when voltage is applied. Pressing the flexible top sheet creates electrical contact between the resistive layers, producing a switch closing in the circuit. The control electronics alternate voltage between the layers and pass the resulting X and Y touch coordinates to the touchscreen controller. The touchscreen controller data is then passed on to the computer operating system for processing. Resistive touchscreen technology possesses many advantages over other alternative touchscreen technologies (acoustic wave, capacitive, Near Field Imaging, infrared). Highly durable, resistive touchscreens are less susceptible to contaminants that easily infect acoustic wave touchscreens. In addition, resistive touchscreens are less sensitive to the effects of severe scratches that would incapacitate capacitive touchscreens. For industrial applications, resistive touchscreens are much more cost-effective than Near Field Imaging touchscreens are. Because of its versatility and cost-effectiveness, resistive touchscreen technology is the touch technology of choice for many markets and applications. Resistive touchscreens are used in food service, retail point-of-sale (POS), medical monitoring devices, industrial process control and instrumentation, portable and handheld products.

8-Wire Resistive Technology: resistive touchscreen technology exists in 4-wire, 5-wire, or 8-wire forms. FastPoint LCD touchscreens specifically employ 8-wire resistive technology because of its benefits over its counterparts. Whereas 8-wire FastPoint touchscreens are available in all sizes, 4-wire resistive technology is restricted to small flatpanels (<10.4"). In contrast to 5-wire resistive touchscreens, 8-wire touchscreens do not experience spacer dots and Newton rings. Additionally, 8-wire resistive touchscreens are not susceptible to problems caused by high-level short-term variances and axis linearity and drift.

1. Polyester Film
2. Upper Resistive Circuit Layer
3. Conductive ITO (Transparent Metal Coating)
4. Lower Resistive Circuit Layer
5. Insulating Dots
6. Glass/Acrylic Substrate
7. Touching the overlay surface causes the (2) Upper Resistive Circuit Layer to contact the (4) Lower Resistive Circuit Layer, producing a circuit switch from the activated area.
8. The touchscreen controller gets the alternating voltages between the (7) two circuit layers and converts them into the digital X and Y coordinates of the activated area.



Surface wave: Surface wave technology uses ultrasonic waves that pass over the touch screen panel. When the panel is touched, a portion of the wave is absorbed. This change in the ultrasonic waves registers the position of the touch event and sends this information to the controller for processing. Surface wave touch screen panels can be damaged by outside elements: contaminants on the surface can also interfere with the functionality of the touchscreen. They offer 100% screen clarity and a hardness of 7 in the Mohs scale, that make possible to scratch the surface with a razorblade to remove paint.

Capacitive: a capacitive touch screen panel is coated with a material, typically indium thin oxide that conducts a continuous electrical current across the sensor. The sensor therefore exhibits a precisely controlled field of stored electrons in both the horizontal and vertical axes - it achieves capacitance. The human body is also an electrical device which has stored electrons and therefore also exhibits capacitance. When the sensor's 'normal' capacitance field (its reference state) is altered by another capacitance field, i.e., someone's finger, electronic circuits located at each corner of the panel measure the resultant 'distortion' in the sine wave characteristics of the reference field and send the information about the event to the controller for mathematical processing. Capacitive sensors must be touched with a conductive device being held by a bare hand or a finger, unlike resistive and surface wave panels that can use anything that can point, such as a finger or stylus. Capacitive touch screens are not affected by outside elements and have high clarity. Controller alternates the voltage gradient alternatively horizontally and vertically in order to obtain x and y coordinates or current values at every corner of the screen.

Infrared: an infrared touch screen panel employs one of two very different methodologies. One method used thermal induced changes of the surface resistance. This method was sometimes slow and required warm hands. Another method is an array of vertical and horizontal IR sensors that detected the interruption of a modulated light beam near the surface of the screen.

Dispersive Signal Technology: this is the newest technology, introduced in 2002. It uses sensors to detect the mechanical energy in the glass that occur due to a touch. Complex algorithms then interpret this information and provide the actual location of the touch. The technology claims to be unaffected by dust and other outside elements, including scratches. Since there is no need for additional elements on screen, it also claims to provide excellent optical clarity. Also, since mechanical vibrations are used to detect a touch event, any object can be used to generate these events, including fingers and styli. The technology is still quite new and is not currently widely available.

Strain Gauge: in a strain gauge configuration the screen is spring mounted on the four corners and strain gauges are used to determine deflection when the screen is touched. This technology can also measure the Z-axis. Typical application falls in protecting new touch-screen railway ticket machines from vandalism.

Touchscreen deployment

Virtually all of the significant touchscreen technology patents were filed during the 1970's and 1980's and have expired. Touchscreen component manufacturing and product design are no longer encumbered by royalties or legalities with regard to patents and the manufacturing of touchscreen-enabled displays on all kinds of devices is widespread.

With the growing acceptance of many kinds of products with an integral touchscreen interface the marginal cost of touchscreen technology is routinely absorbed into the products that incorporate it and is effectively eliminated. As typically occurs with any technology, touchscreen hardware and software, has sufficiently matured and been perfected over more than three decades to the point where its reliability is unassailable. As such, touchscreen displays are found today in airplanes, automobiles, gaming consoles, machine control systems, appliances and handheld display devices of every kind.

Gaming and Entertainment: Casino game manufacturers around the globe have found touch screens indispensable in providing a richer, more interactive gaming experience for casino players.

Industrial: Touch can be applied virtually anywhere on a factory floor. From process automation to quality control, touch is fast becoming a preferred solution. Touch-enabled products provide users with a space saving, NEMA sealable option for challenging environments.

Healthcare: Instrumental in healthcare environments, touch solutions add convenience and reliability to applications including patient records, nurse stations, and patient monitoring systems. 3M's Resistive touch solutions support glove use for examination equipment while privacy solutions offer increased confidentiality in busy healthcare environments.

Point-of-Sale: Intuitive touch solutions offer numerous opportunities to retail and point-of-sale environments including cash registers in retail outlets and food ordering systems in bars and restaurants. Touch solutions can also be beneficial for kitchen staff filling orders and for seating charts and reservation schedules.

Financial: From Main Street to Wall Street, touch screens help simplify the complex world of banking and finance. ATMs were at the forefront of adopting touch screen technology for their intuitive ease-of-use and durability in public places. On the stock exchange floor and at brokerage houses, touch-based applications help accelerate the trading process while providing on-line trading information.

Self-Service Kiosks: Touch-enabled kiosks provide self-service capabilities as well as access to information 24-hours a day, 7 days a week. Self-service kiosks are used to provide services ranging from buying theater tickets, to viewing information directories, to self check-in at airports and self check-out in retail environments.

Mobile Devices: The never-ending challenge for today's mobile device designers is to make tomorrow's devices smaller, lighter and more durable. 3M touch solutions are critical in solving the design challenges for mobile devices: tight space constraints, size, weight, ruggedness, durability, and low power consumption.



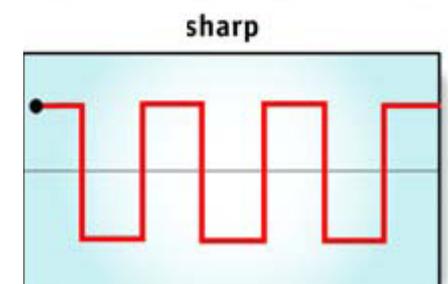
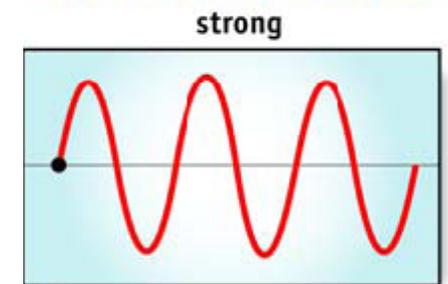
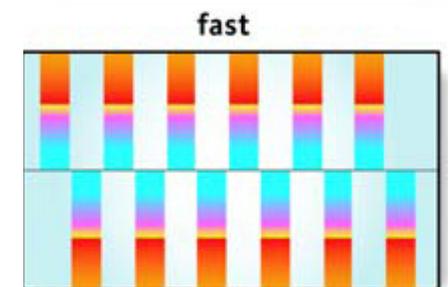
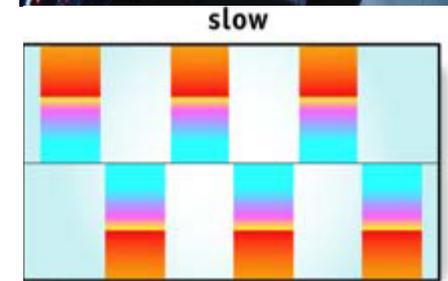
Haptics in touchscreens

The Perception of Pressing Physical Switches: graphical buttons seem to depress and release. The result is a more intuitive, natural, and multisensory experience. Immersion proprietary technology causes the touchscreen to vibrate, creating the perception of pressing physical switches, emulating crisp qualities and particular force and push-away characteristics. Based on software or firmware programming, touching different virtual objects on the screen produces the desired, context-sensitive feel for the user. The programmability of TouchSense technology can be used to further enhance usability, particularly for noisy or distracting environments. For example, pressing the Enter button, no matter where it appears on the screen, could always produce a consistent, strong, and crisp tactile sensation, while buttons for other actions could have a different feel, helping the user to know instinctively whether they've made the correct selection.

Rich touch sensations can be created by combining individual effects of varying frequency, waveform, magnitude, and duration. This flexibility means that touchscreens can be used for many types of applications by providing a wide range of response:

- Graphical buttons can simulate the familiar up and down clicks of physical buttons
- Menu items can be programmed to supply a light pulse sensation or a confirming push-back response when pressed
- Control switches can exhibit increasing or decreasing vibrations corresponding to motor or fan speed, radio volume level, lighting levels, or other parameters
- Enter, Next, and other major and minor functions can supply a consistent feel throughout an application
- Scrolling displays can provide a stop sensation when the first or last items have been reached
- Switch controls can exhibit a pop effect

Because understanding through touch is instinctive, touch feedback can reduce the need for focused attention, allowing safer operation for many applications. For aerospace and automotive applications, even a small improvement in ease of use, error reduction, or glance time can have a significant impact on safety. User interaction is about creating an effective human-machine interface (HMI) that leads to rapid and accurate task completion or a more satisfying experience. Research has found that, when compared to single-mode voice and visual applications, multimodal applications are easier and more intuitive to use. Yet, in the world of computers and digital devices and controls, haptic feedback is often lost.



[1] Surface acoustic wave

A **surface acoustic wave (SAW)** is an acoustic wave traveling along the surface of a material having some elasticity, with an amplitude that typically decays exponentially with the depth of the substrate. This kind of wave is commonly used in piezoelectric devices called SAW devices in electronics circuits. SAW devices are employed as filters, oscillators and transformers based on the transduction of acoustic waves. Discovered in 1887 by Lord Rayleigh who reported the surface acoustic wave mode of propagation and in his classic paper predicted the properties of these waves. Named for their discoverer, Rayleigh waves have a longitudinal and a vertical shear component that can couple with a medium in contact with the device's surface. This coupling strongly affects the amplitude and velocity of the wave allowing SAW sensors to directly sense mass and mechanical properties. Electronic devices employing the SAW normally utilize one or more interdigital transducers (IDTs) to convert acoustic wave to electrical signal and vice versa utilizing the piezoelectric effect of certain materials (quartz, lithium niobate, lithium tantalate, LGS etc). These devices are fabricated utilizing common processes used in the manufacture of silicon integrated circuits.

SAW filters have enjoyed successful application in the booming cellular telephones market and provide significant advantages in performance, cost, and size over other filter technologies (digital signal processors, quartz crystals (bulk wave) , LC filters, and waveguide filters). Significant research has been done in the last 20 years in the area of surface acoustic wave sensors. Sensor applications include all areas of sensing (such as chemical, optical, thermal, pressure, acceleration, torque and biological). SAW sensors have seen limited commercial success to date but are commonly commercially available for some applications such as touchscreen displays. More than 40 worldwide patents relating to the use of SAW sensors in the automotive industry have been filed by UK intellectual property (IP) company Transense Technologies. These batteryless sensors are expected to replace lithium ion battery powered tyre pressure sensors shortly, while the company's torque SAW sensors will be used to facilitate a number of engine control applications that have not been possible before.

4.2. Keyboard and mouse

The keyboard and mouse are typical input devices for a personal computer and are currently the main game controllers for computer games. Some video game consoles also have the ability to function with a keyboard and a mouse. The computer keyboard is modeled after the typewriter keyboard and was designed for the input of written text. A mouse is a handheld pointing device used in addition to the keyboard. For games, the keyboard typically controls movement of the character while the mouse is used to control the game camera or used for aiming. The numeric keypad found on the keyboard is also used as a game controller and can be found on a number of separate devices, most notably early consoles, usually attached to a joystick or a paddle. The keypad is a small grid of keys with at least the digits 0-9.



4.2.0. Computer keyboard

A computer keyboard is a peripheral modeled after the typewriter keyboard. Keyboards are designed for the input of text and characters, and also to control the operation of a computer. Physically, computer keyboards are an arrangement of rectangular or near-rectangular buttons, or "keys". Keyboards typically have characters engraved or printed on the keys; in most cases, each press of a key corresponds to a single written symbol. However, to produce some symbols requires pressing and holding several keys simultaneously, or in sequence; other keys do not produce any symbol, but instead affect the operation of the computer, or the keyboard itself. See input method editor. Roughly 50% of all keyboard keys produce letters, numbers or signs (characters). Other keys can produce actions when pressed, and other actions are available by simultaneously pressing more than one action key.



🖨️ A 104-key PC US English QWERTY keyboard layout

Designs

There exist a large number of different arrangements of symbols on keys. These different keyboard layouts arise because different people need easy access to different symbols; typically, this is because they are writing in different languages, but specialized keyboard layouts for mathematics, accounting, and computer programming do exist. The number of keys on a keyboard generally varies from the standard 101 keys to the 104 windows keyboards all the way up to 130 keys with many programmable keys. There are also compact variants that have fewer than 90 keys. They are normally found in laptops or in desktop computers with space constraints.

Standard arrangements

The most common arrangements in Western countries are based on the QWERTY layout (including closely-related variants, such as the French AZERTY layout). Even in countries where different alphabets or writing systems are in use, the physical layout of the keys is often quite similar (e.g. the Thai keyboard layout).

Additional keys and "Internet" keyboards

Most modern computer keyboards (including those on the PC and Apple Mac) are based on these standard versions, but include additional keys not normally found on typewriters, such as function keys, a numeric keypad, and so on. In recent years, so-called "Internet keyboards" have also become popular. These include extra buttons for specific applications or functions (typically a browser or email client).

Plug Types

There are a few different ways of connecting a keyboard which have evolved over the years. These include the standard AT connector commonly found on pre-80486 motherboards, which evolved to the now common PS/2 and USB connections. Prior to the iMac line of systems, Apple Computer used ADB, a proprietary system, for its keyboard connector.

Alternatives

A standard keyboard is physically quite large, as each key must remain large enough to be easily pressed by fingers. Other types of keyboards have been proposed for small portable equipment where a standard keyboard is too large. One way to reduce the number of keys is to use chording, i.e. pressing several keys simultaneously. As an example, the GKOS keyboard has been designed for small wireless devices. Other two-handed alternatives more akin to a gaming controller, such as the AlphaGrip, are also used as a way to input data and text.

Usage

Microsoft On-Screen Keyboard is software designed for users with limited mobility. In normal usage, the keyboard is used to type text into word processor, text editor, or any other textbox. In modern computers the interpretation of keypresses is generally left to the software. Modern keyboards distinguish each physical key from every other and report all keypresses and releases to the controlling



software. This flexibility is not often taken advantage of and it usually does not matter, for example, whether the left or right shift key is held down in conjunction with another character.

Commands

A keyboard is also used to type commands in a computer. One famous example on the PC is the Ctrl+Alt+Del combination. With current versions of Windows, this brings up a menu-window including options for handling currently-running applications and shutting down the computer, amongst other things. Under Linux, MS-DOS and some older versions of Windows, Ctrl+Alt+Del performs either a 'cold' or 'warm' reboot.

Games

A keyboard is one of the primary methods of control in computer games. For instance, the arrow keys or a group of letters resembling the pattern of the arrow keys, like WASD, can be used for movement of a game character. In many games keys can be configured to the user's preferences. Alphabet keys are also sometimes used to perform actions starting with that letter. (e.g. pressing *e* to *eat* in NetHack). Keyboards are less than ideal when many keys are to be pressed at once, as the limited circuitry means that only a certain number of keys will register at one time. An obvious example of this is phantom key blocking. On older keyboards, due to the circuit design, sometimes pressing three keys simultaneously results in a 4th keypress being registered. Modern keyboards prevent this from happening by blocking the 3rd key in certain key combinations, but while this prevents phantom input, it also means that when two keys are depressed simultaneously, many of the other keys on the keyboard will not respond until one of the two depressed keys is lifted. Better keyboards are designed so that this happens infrequently in office programs, but it remains a problem in games even on expensive keyboards, due to wildly different and/or configurable key/command layouts in different games.

How it works

The following briefly describes a "dome-switch" keyboard (sometimes misleadingly referred to as a membrane keyboard), the most common type in use today:-

1. When a key is pressed, it pushes down on a rubber dome sitting beneath the key. A conductive contact on the underside of the dome touches (and hence connects) a pair of conductive lines on the circuit below.
2. This bridges the gap between them and allows current to flow (i.e. the switch goes from open to closed), changing the signal strength.
3. A scanning signal is emitted by the chip along the pairs of lines to all the keys. When the signal in one pair becomes different, the chip generates a "make code" corresponding to the key connected to that pair of lines.
4. The code generated is sent to the computer either via a keyboard cable (using on-off electrical pulses to represent bits) or over a wireless connection.
5. A chip inside the computer receives the signal bits and decodes them into the appropriate keypress. The computer then decides what to do on the basis of the key pressed (e.g. display a character on the screen, or perform some action).

Other types of keyboard function in a similar manner, the main differences being how the individual key-switches work. For more detail, refer to the "Keyboard technology" article.

Scheda: Tastiera

La tastiera QWERTY venne inventata da Christopher Sholes nel 1870 circa, per le macchine da scrivere dell'epoca a martelletti. I tasti sono disposti in modo da evitare incastri fra i martelletti. La tastiera Dvorak fu inventata da Augusto Dvorak nel 1936: adotta una disposizione razionale dei tasti: in funzione della frequenza delle lettere nelle parole inglesi, per minimizzare gli spostamenti delle mani e per alternarle il più possibile, per privilegiare le dita forti (il 70% delle lettere vengono scritte con gli indici e i medi) e la mano destra (56% delle pressioni). Dvorak ha fallito perché è molto difficile sostituire una tecnologia che ha conquistato una larga base di utenti con una nuova tecnologia, anche se quest'ultima è migliore. Le tastiere tradizionali impongono posture innaturali, che possono indurre vari disturbi.

Nelle tastiere ad accordi ad ogni dito corrisponde un singolo tasto e ad ogni carattere corrisponde una combinazione di tasti (accordo), e non un tasto singolo. La velocità che si ottiene è considerevole, ma solo per utenti esperti.



Keys on a computer keyboard

- Modifier key: Control key | Shift key | Alt key / Option key (Macintosh) | AltGr key | Command key / Meta key (MIT computer keyboards) | Windows key | Fn key (Compact keyboards)
- Dead key: Compose key
- Lock key: Scroll lock | Num lock | Caps lock
- Navigation keys: Arrow keys | Page Scroll (Page Up / Page Down) | Home key / End key
- Edit keys: Return key / Enter key | Backspace | Insert key | Delete key | Tab
- SysRq / Print screen
- Break key / Pause key
- Escape key
- Menu key
- Space bar
- Numpad
- Function key
- Power management keys: Power key | Sleep key | Wake key



Posizione naturale
dei polsi



Posizione indotta da una
tastiera tradizionale



Keyboard technology

There are several types of keyboard, usually differentiated by the switch technology employed in their operation. Since there are so many switches needed (usually about 80-110) and because they have to be highly reliable, this usually defines the keyboard. The choice of switch technology affects key response (the positive feedback that a key has been pressed) and travel (the distance needed to push the key to enter a character reliably). Newer models use hybrids of various technologies to achieve greater cost savings.

Types

Dome-switch keyboard

Dome-switch keyboards are kind of a hybrid of membrane and mechanical keyboards. They bring two circuit board traces together under a rubber "dome" or bubble. The top of the bubble is coated in some conductive substance. When a key is pressed, it collapses the dome, which shorts out the two circuit traces and completes the connection to enter the character. The pattern on the PC board is often gold-plated. This is a common switch technology used in mass market keyboards today. It is considered very quiet, but purists tend to find it "mushy" because the collapsing dome does not provide as much positive response as a hard closing switch. These are also a good choice for office or consumer environments because they share a certain degree of liquid resistance with their membrane ancestors. This switch technology also happens to be most commonly used in handheld controllers, such as those used with home video game consoles. Dome-switch keyboards are also called direct-switch keyboards.

Capacitive keyboard

In this type of keyboard, pressing the key changes the capacitance of a pattern printed on a PC board. Usually this permits a pulse or pulse train to be sensed. Unlike "dome switch" keyboards, the pattern will be covered by a thin, insulating film. Capacitive keyboards are inexpensive, and resist wear, water, foreign objects and dirt. They are common in PC keyboards.

Mechanical-switch keyboard

Mechanical-switch keyboards use real switches, one under each key. Depending on the construction of the switch, these keyboards have varying responses and travel times. Notable keyboards utilizing this technology are the Apple Extended II and its modern imitator, the Matias Tactile Pro. These two keyboards use ALPS switches. Cherry corporation of Germany also makes mechanical switches used in special purpose and high end keyboards.

Buckling-spring keyboard

It is a common misconception that the IBM Model M and its derivatives are mechanical-switch keyboards. In fact, the Model M utilizes membrane-sheet switches, much like those found in a dome-switch keyboard. The buckling spring mechanism (U.S. Patent 4,118,611) atop the switch is responsible for the tactile and aural response of the keyboard. This mechanism controls a small hammer that strikes the membrane switch. For more information, see an examination of buckling-spring technology. In 1993, two years after spawning Lexmark, IBM transferred its keyboard operations to the daughter company. New Model M keyboards continued to be manufactured for IBM by Lexmark until 1996, when Unicomp purchased the keyboard technology. Today, Unicomp manufactures new buckling-spring keyboards and repairs old IBM and Lexmark keyboards.

Hall-effect keyboard

Hall effect keyboards use magnets and "Hall Effect" sensors instead of an actual switch. When a key is depressed, it moves a magnet, which is detected by a solid-state Hall-effect sensor. These keyboards are extremely reliable, and are able to accept millions of keystrokes before failing. They are used for ultra-high reliability applications, in locations like nuclear powerplants or aircraft cockpits. They are also sometimes used in industrial environments. These keyboards can be easily made totally waterproof. They also resist large amounts of dust and contaminants. Because a magnet and sensor is required for each key, as well as custom control electronics, they are very expensive.

Laser keyboard

A laser projection device approximately the size of a computer mouse projects the outline of keyboard keys onto a flat surface, such as a table or desk. When the laser is interrupted in the position of a key, a keystroke is registered. This type of keyboard is very portable, and many models have retractable cords and wireless capabilities. However, sudden or accidental disruption of the laser will register unwanted keystrokes. Also, if the laser malfunctions, the whole unit becomes useless, unlike conventional keyboards which can be used even if a variety of parts (such as the keycaps) are removed. This type of keyboard can be cumbersome to use since it is susceptible to errors, even in the course of normal typing.

Membrane keyboard

Membrane keyboards are usually flat. They are most often found on appliances like microwave ovens or photocopiers. A common design consists of three layers. The top layer (and the one the user touches) has the labels printed on its front and conductive stripes printed on the back. Under this it has a spacer layer, which holds the front and back layer apart so that they do not normally make electrical contact. The back layer has conductive stripes printed perpendicularly to those of the front layer. When placed together, the stripes form a grid. When the user pushes down at a particular position, his finger pushes the front layer down through the spacer layer to close a circuit at one of the intersections of the grid. This indicates to the computer or keyboard control processor that a particular button has been pressed. Membrane keyboards do not generally have much of a "feel", so many machines which use them issue a beep or flash a light when the key is pressed. They are often used in harsh environments where water or leak proofing is desirable. Although used in the early days of the personal computer (on the ZX80, ZX81 and Atari 400), they have been supplanted by the more tactile dome and mechanical switch keyboards. It should be noted, however, that membrane keyboards with interchangeable key layouts, such as the IntelliKeys and Discover:board are still commonly used by people with physical, visual, or cognitive disabilities as well as people who require assistive technology to access a computer.

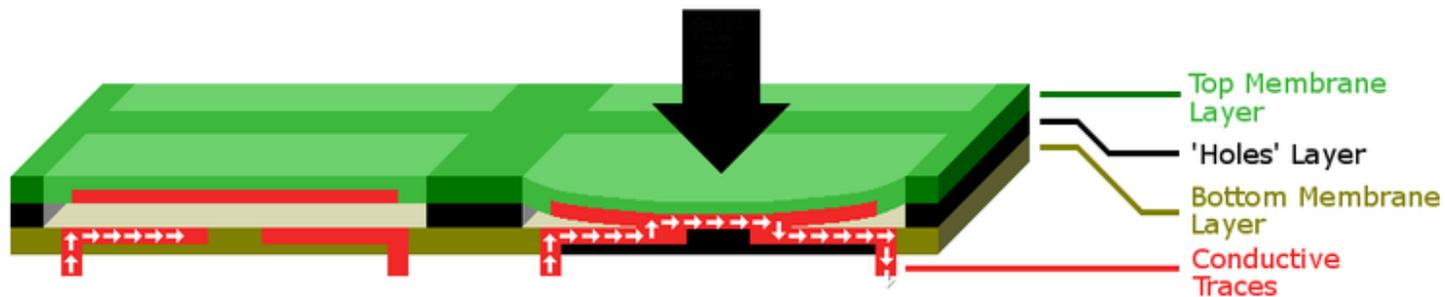
A membrane keyboard is a computer keyboard whose "keys" are not separate, moving parts, as with the majority of other keyboards, but rather have only outlines and symbols printed on a flat, flexible surface. Very little, if any, tactile feedback is felt when using such a keyboard, and error-free blind typing can be difficult. Membrane keyboards, which work by electrical contact between the keyboard surface and the underlying circuits when keytop areas are pressed, were used with some early 1980s home computers, and have been much used in consumer electronics devices. The keyboards are very inexpensive to mass produce, and are more resistant against dirt and liquids than most other keyboards, but due to the low or non-existent amount of tactile feedback provided, most people have difficulty typing with them, especially when large numbers of characters need to be typed. Chiclet keyboards were a slight improvement, at least allowing individual keys to be felt to some extent.

Aside from early hobbyist/kit/home computers and some video game consoles, membrane-based QWERTY keyboards are used in some industrial computer systems, and are also found as portable, even "rollable-collapsible" designs for PDAs and other pocket computing devices. Smaller, specialised membrane keyboards, typically numeric-and-a-few-control-keys only, have been used in access control systems (for buildings and

restricted areas), simple handheld calculators, domestic remote control keypads, and other similar devices where the amount of typing is relatively small or infrequent. The expression is also sometimes used in connection with modern PC keyboards which utilise dome switch technology. Although the lower layers in some dome-switch keyboards are essentially the same as the membrane keyboard, the dome-switch keyboard includes additional upper layers (rubber domes and solid key-caps).

How it works

As can be seen from the diagram below, the membrane keyboard basically consists of three layers; two of these are membrane layers containing conductive traces. The center layer is a "spacer" containing holes wherever a "key" exists. It keeps the other two layers apart.



🖨️ Cross-section diagram of a typical membrane keyboard. The thickness of the bottom three layers has been exaggerated for clarity; in reality, they are not much thicker than pieces of paper or cardboard.

Under normal conditions, the switch (key) is open, because current cannot cross the non-conductive gap between the traces on the bottom layer. However, when the top layer is pressed down (with a finger), it makes contact with the bottom layer. The conductive traces on the underside of the top layer can then bridge the gap, allowing current to flow. The switch is now "closed", and the parent device registers a keypress.

Roll-up keyboard

Some keyboards are designed out of flexible materials that can roll up in a tight bundle. Typically they are completely sealed in rubber, making them watertight like membrane keyboards. Like membrane keyboards, they are reported to be very hard to get used to, as there is little tactile feedback. Such keyboards include:

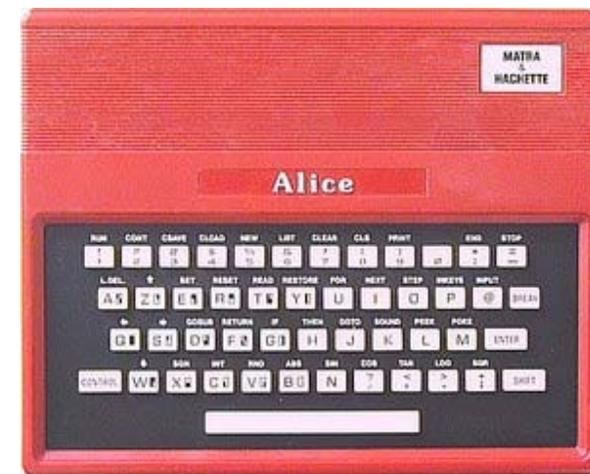
- the Cool line of keyboards,
- the Eumax keyboard,
- the VIK ("virtually indestructible keyboard"),
- the Flexboard keyboard,
- the Whitelite FX100 Keyboard.

Other parts of PC keyboard

The modern PC keyboard is more than just the switch technology, however. It also includes a control processor and indicator lights to provide feedback to the user about what state the keyboard is in. Depending on the sophistication of the controller's programming, the keyboard may also offer other special features. The processor is usually a single chip 8048 microcontroller variant. The keyboard switch matrix is wired to its inputs and it processes the incoming keystrokes and sends the results down a serial cable (the keyboard cord) to a receiver in the main computer box. It also controls the illumination of the "caps lock", "num lock" and "scroll lock" lights. A common test for whether the computer has crashed is pressing the "caps lock" key. The keyboard sends the key code to the BIOS code running in the main computer, If the main computer is operating, it commands the light to turn on. All the other indicator lights work in a similar way. The BIOS also tracks the shift, alt and control state of the keyboard. When pressing a keyboard key, the key "bounces" like a ball against its contacts several times before it settles into firm contact. When released, it bounces some more until it reverts to the uncontacted state. If the computer was watching for each pulse, it would see many keystrokes for what the user thought was just one. To resolve this problem, the processor in a keyboard (or computer) "debounces" the keystrokes, by aggregating them across time to produce one "confirmed" keystroke that (usually) corresponds to what is typically a solid contact. It could be argued that the dome switch technology outlined above owes its popularity to the ability of the processor to accurately debounce the keystrokes. Early membrane keyboards limited typing speed because they had to do significant debouncing. Anyone who ever tried word processing on a ZX81 will recall this.

QWERTY

Qwerty is the most common modern-day keyboard layout on most English language computer and typewriter keyboards. It takes its name from the first six letters seen in the keyboard's top first row of letters. The QWERTY design was patented by Christopher Sholes in 1868 and sold to Remington in 1873, when it first appeared in typewriters. Frequently used pairs of letters were separated in an attempt to stop the typebars from intertwining and becoming stuck, thus forcing the typist to manually unstick the typebars and also frequently blotting the document[1]. The home row (ASDFGHJKL) of the QWERTY layout is thought to be a remnant of the old alphabetical layout that QWERTY replaced. QWERTY also attempted to alternate keys between hands, allowing one hand to move into position while the other hand strikes home a key. This sped up both the original double-handed hunt-and-peck technique and the later touch typing technique; however, single-handed words such as *stewardesses* and *monopoly* show flaws in the alternation. Minor changes to the arrangement are made for other languages; for example, German keyboards add umlauts to the right of "P" and "L", and interchange the "Z" and "Y" keys both because "Z" is a much more common letter than "Y" in German (the letter seldom appearing except in borrowed words), and because "T" and "Z" often appear next to each other in the German language; consequently, they are known as QWERTZ keyboards. French keyboards interchange both "Q" and "W" with "A" and "Z", and move "M" to the right of "L"; they are known as AZERTY keyboards. Italian typewriter keyboards (but not most computer keyboards) use a QZERTY layout where "Z" is swapped with "W" and "M" is at the right of "L". Portuguese keyboards maintain the QWERTY layout but add an extra key: the letter "C" with cedilla (Ç) after the "L" key. In this place, the Spanish version has the letter "N" with tilde (Ñ) and the "Ç" (which is not used in Spanish, but is part of sibling languages like French, Portuguese and Catalan) is placed at the rightmost position of the home line, beyond the diacritical dead



keys. Norwegian keyboards inserts "Å" to the right of "P", "Ø" to the right of "L", and "Æ" to the right of "Ø", thus not changing the appearance of the rest of the keyboard. The Danish layout is like the Norwegian, only switching "Æ" and "Ø", and Swedish has their umlaut letters "Ä" and "Ö" in those places. Some keyboards for Lithuanian used a layout known as AŽERTY, where "A" appears in place of "Q" above "A", Ž in place of "W" above "S", with "Q" and "W" being available either on the far right-hand side or by use of the Alt Gr key.

Alternative keyboard layouts

Because modern keyboards do not suffer from the problems of older mechanical keyboards, the QWERTY layout's separation of frequently used letter pairs is no longer strictly necessary. Several alternative keyboard layouts, such as Dvorak Simplified Keyboard arrangement (designed by Drs. August Dvorak and William Dealey and patented in 1936), have been designed to increase a typist's speed and comfort, largely by moving the most common letters to the home row and maximizing hand alternation. The effectiveness of these layouts is disputed. Some studies [2] have shown that alternative methods are more efficient, but Dvorak and other alternative typists most often cite comfort as the greatest advantage. QWERTY's inventor, Christopher Scholes, patented a key arrangement similar to Dvorak's, but it never became popular. Some researchers, such as economists Stan Liebowitz at the University of Texas at Dallas, Texas, and Stephen E. Margolis of North Carolina State University, claim that QWERTY is really no less efficient than other layouts. Opponents point out that August Dvorak stood to gain from the success of his layout, and that he may have perpetuated his "efficiency myth" to increase his financial gains. Other QWERTY advocates claim that for a QWERTY typist to switch to Dvorak or another layout requires more effort than initially learning to touch-type, because of having to retrain the fingers' muscle memory. Computer users also need to unlearn the habit of pressing key shortcuts (for example: Ctrl-C for copy, Ctrl-X for cut, Ctrl-V for paste, on Microsoft Windows), though some programs and operating systems allow the use of alternate layouts combined with QWERTY shortcuts. Opponents of alternative keyboard designs most often point to QWERTY's ubiquity as a deciding factor, because the costs incurred by using the supposedly inefficient layout are much less than those of retraining typists. It is not unusual to find Dvorak typists who also touch-type the QWERTY layout for convenience, since QWERTY dominates the keyboard market. The tension between the Dvorak efficiency and the QWERTY ubiquity illustrates the problem of collective switching costs, assuming QWERTY's relative inefficiency. Besides the Dvorak layout, there are many other newer alternative keyboard layouts, but those layouts have not gained widespread use.

Trivia

- The longest common English word that can be typed using only the left hand (using conventional hand placement) is *stewardesses*. The words *sweaterdresses* and *aftercataracts* are longer and can also be typed with only the left hand, but they are not in all dictionaries.
- The longest English word that can be typed with the right hand only (using conventional hand placement) is *johnny-jump-up*, or alternatively *polyphony*.
- The word *typewriter* can be typed entirely using the top row of the QWERTY keyboard; it has been speculated that this may have been a factor in the choice of keys for ease of demonstration, although this is unlikely.
- The average person is expected to type 30-40 words per minute using the touch typing technique on a QWERTY keyboard. 40-50 words per minute is considered excellent, and some have been clocked at over 90 words per minute.
- Computer programmers and people who work with computers extensively sometimes type up to 130 words per minute. This is off the charts, and they help to fuel the theories that there are some computer hackers that actually go through keyboards quite quickly.

4.2.1. Mouse or Mice?

Operating a mechanical mouse.

- 1: Moving the mouse turns the ball.
- 2: X and Y rollers grip the ball and transfer movement.
- 3: Optical encoding disks include light holes.
- 4: Infrared LEDs shine through the disks.
- 5: Sensors gather light pulses to convert to X and Y velocities.

A mouse is a handheld computer pointing device, designed to sit under one hand of the user and detect movement relative to its supporting surface. In addition, it usually features buttons and/or other devices, such as "wheels", which allow performing various system-dependent operations. The mouse's 2D motion is typically translated into the motion of a pointer on a display. The name "mouse", coined at the Stanford Research Institute, derives from the resemblance of early models (which had a cord attached to the rear part of the device, suggesting the idea of a tail) to the common small rodent of the same name.^[1]

Early mice

The first computer mouse, held by inventor Douglas Engelbart, showing the wheels that contact the working surface. The mouse was invented by Douglas Engelbart of Stanford Research Institute in 1963, after extensive usability testing. Engelbart's team called it "bug". It was one of several experimental pointing devices developed for Engelbart's oN-Line System (NLS). The other devices were designed to exploit other body movements—for example, head-mounted devices attached to the chin or nose—but ultimately the mouse won out because of its simplicity and convenience. The first, bulky, mouse (*pictured*) used two gear wheels perpendicular to each other: the rotation of each wheel was translated into motion along one axis. Engelbart received patent US3541541 on November 17, 1970 for an "X-Y Position Indicator for a Display System". At the time, Engelbart intended that users would hold the mouse continuously in one hand and type on a five-key chord keyset with the other.

Mechanical mice

The so-called "ball mouse" was later invented in the early 1970s by Bill English while he was working for Xerox PARC; it replaced the external wheels with a single ball (that could rotate in any direction). The ball's motion, in turn, was detected using perpendicular wheels housed inside the mouse's body. This variant of the mouse resembled an inverted trackball and was the predominant form used with personal computers throughout the 1980s and 1990s. The Xerox PARC group also settled on the modern technique of using both hands to type on a full-size keyboard and grabbing the mouse as needed. Modern computer mice took form at the École polytechnique fédérale de Lausanne (EPFL) under the inspiration of Professor Jean-Daniel Nicoud and the hands of engineer and watchmaker André Guignard. A spin-off of EPFL,



Logitech, launched the first popular mice. The major movement translation techniques are by optical, mechanical and inertial sensors.

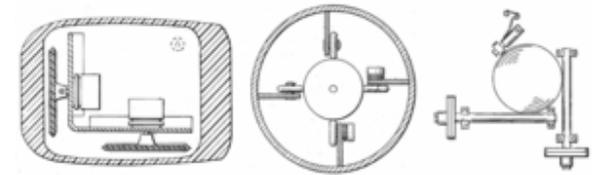
Optical mice

An optical mouse uses a light-emitting diode and photodiodes to detect the movement of the underlying surface, rather than moving some of its parts as in a mechanical mouse. Early optical mice, circa 1980, were of two types. Some, such as those invented by Steve Kirsch of Mouse Systems Corporation, used an infrared LED and a four-quadrant infrared sensor to detect grid lines printed on a special metallic surface with infrared absorbing ink. Predictive algorithms in the CPU of the mouse calculated the speed and direction over the grid. Others, invented by Richard F. Lyon and sold by Xerox, used a 16-pixel visible-light image sensor with integrated motion detection on the same chip ([1]) and tracked the motion of light dots in a dark field of a printed paper or similar mouse pad ([2]). These two mouse types had very different behaviors, as the Kirsch mouse used an x-y coordinate system embedded in the pad, and would not work correctly when rotated, while the Lyon mouse used the x-y coordinate system of the mouse body, as mechanical mice do.

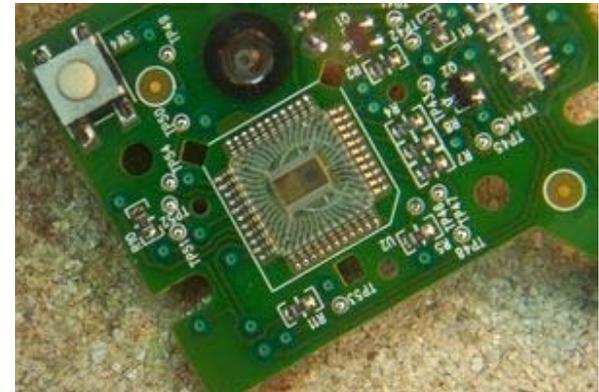
As computing power grew cheaper, it became possible to embed more powerful special-purpose image processing chips in the mouse. This advance enabled the mouse to detect relative motion on a wide variety of surfaces, translating the movement of the mouse into the movement of the pointer and eliminating the need for a special mouse pad. This advance paved the way for widespread adoption of optical mice. Modern surface-independent optical mice work by using an optoelectronic sensor to take successive pictures of the surface on which the mouse is operating. Most of these mice use LEDs to illuminate the surface that is being tracked; LED optical mice are often mislabeled as "laser mice". Changes between one frame and the next are processed by the image processing part of the chip and translated into movement on the two axes using an optical flow algorithm. For example, the Agilent Technologies ADNS-2610 optical mouse sensor processes 1512 frames per second: each frame is a rectangular array of 18×18 pixels, and each pixel can sense 64 different levels of gray. Optomechanical mice detect movements of the ball optically, giving the precision of optical without the surface compatibility problems, whereas optical mice detect relative movement of the surface by examining the light reflected off it.

Laser mice

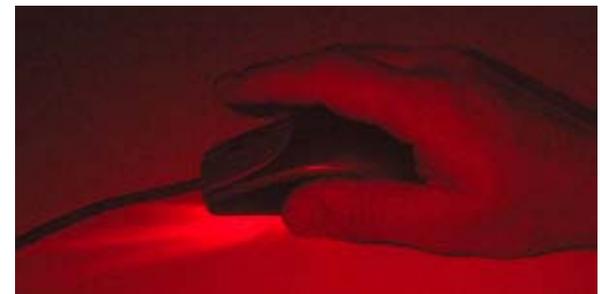
In 2004, Logitech, along with Agilent Technologies, introduced the laser mouse with its *MX 1000* model. This mouse uses a small laser instead of a LED. The new technology can increase the resolution of the image taken by the mouse. The companies claim that this leads to a 20× increase in the sensitivity to the surface features used for navigation compared to conventional optical mice (see interference). Gamers have complained that the MX 1000 does not respond immediately to movement after it is picked up, moved, and then put down on the mouse pad. Newer revisions of the mouse do not seem to suffer from this problem, which is a power-saving feature (almost all optical



Early mouse patents. From left to right: Opposing track wheels by *Engelbart*, Nov. '70, 3541541. Ball and wheel by *Rider*, Sept. '74, 3835464. Ball and two rollers with spring by *Opocensky*, Oct. '76, 3987685.



The optical sensor from a Microsoft Wireless IntelliMouse Explorer (v. 1.0A).



mice, laser or LED based, also implement this power-saving feature, except those intended for use in gaming, where a millisecond of delay is significant). Since it is a wireless mouse, the engineers designed it to save as much power as possible. In order to do this, the mouse blinks the laser when in standby mode (8 seconds after the last motion). This function also increases the laser life. As early as 1998, Sun Microsystems provided a laser mouse with their Sun SPARC Station servers and workstations.

Optical versus mechanical mice

Optical mice supporters claim optical rendering works better than mechanical mice, requires no maintenance and lasts longer due to fewer moving parts. Optical mice do not normally require any maintenance other than removing debris that might collect under the light emitter, although cleaning a dirty mechanical mouse is fairly straightforward too. Mechanical mice supporters point out that optical mice generally cannot track on glossy and transparent surfaces, including many commercial mouse pads, causing them to periodically "spin" uncontrollably during operation. Mice with less image processing power also have problems tracking extremely fast movement, though high-end mice can track at 1 m/s (40 inches per second) and faster. Power conservation is typically not an issue for cabled mice. At the time of writing (2006), mechanical mice have lower average power demands than their optical counterparts. This is of no practical concern for cabled mice, but has an impact on battery-powered wireless models. A typical mechanical model requires 25 mA at +5 V (= 0.125 W), or less, whereas an optical model attains 100 mA at +5 V (= 0.5 W) for optical devices (for a 4:1 ratio). Since optical mice render movement based on an image the LED reflects, performance on multi-colored mousepads may be unreliable. However, they will outperform mechanical mice on uneven, slick, squishy, sticky or loose surfaces, and generally in mobile situations where mouse pads are not available.

Inertial Mice

Inertial mice detect movement through a gyroscope, for every axis supported. Usually cordless, they often have a switch to deactivate the movement circuitry between use, allowing the user freedom of movement without affecting the pointer position.

Buttons

In contrast to the motion-sensing mechanism, the mouse's buttons have changed little, varying mostly in shape, number, and placement. Engelbart's very first mouse had a single button; this was soon increased to three at Xerox PARC, but reduced to two for Xerox products. Apple reduced it back to one button with the Macintosh in 1984, while Unix workstations from Sun and others used three buttons. Commercial mice usually have between one and three buttons, although in the late 1990s some mice had five or more. Most popular are mice with two buttons. The most common purpose for the second button is to invoke a contextual menu in the computer's software user interface, which contains options specifically tailored to the interface element over which the mouse pointer is positioned. By default, the primary mouse button is located on the left hand side of the mouse, for the benefit of right handed users. On systems with three-button mice, pressing the center button (a middle click) is often used as a convenience to map the action to a commonly used action, or a macro. In the X Window System, middle clicking pastes the contents of the primary buffer at the pointer's position. Many two-button mice are configured to emulate a three-button mouse by clicking both the right and left buttons simultaneously. Middle-clicks are often used as a spare button in case a function is not allocated easily.

Additional buttons

Mice have been built with five or more buttons. Depending on the user's preferences, the extra buttons may allow forward and backward web navigation, scrolling through a browser's history, or other functions. As with similar features in keyboards, however, these functions may not be

supported by all software. The additional buttons are generally more useful in computer games, where quick and easy access to a wide variety of functions (for example, weapon-switching in first-person shooters) can be very beneficial. Because mouse buttons can be mapped to virtually any function, keystroke, application or switch, they can make working with such a mouse more efficient and easier to use. Douglas Engelbart's view of the optimal number of buttons was "as many as possible". The prototype that popularised the idea of three buttons as standard had that number only because "we couldn't find anywhere to fit any more switches".

Wheels

One major innovation in mouse buttons was the scroll wheel: a small wheel that can be rotated to provide immediate one-dimensional input. Usually, this input is translated into "scrolling" up or down within the active window or GUI element. This is especially helpful in navigating a long document. The scroll wheel can often be pressed too, thus being in fact a third (center) button. Under many Windows applications, the wheel pressure activates autoscrolling and in conjunction with the control key (Ctrl) may zoom in and out (applications which support this feature include Microsoft Word, Internet Explorer, Opera and Mozilla Firefox). The scroll wheel was first introduced by Microsoft IntelliMouse in 1996. The feature became a commercial success in 1997 when Office application suite and Internet Explorer browser started supporting the wheel scrolling feature. After that the scroll wheel has become a norm. Some newer mouse models have two wheels, assigned to horizontal and vertical scrolling. Designs exist which make use of a "rocker" button instead of a wheel—a pivoting button that can be pressed at the top or bottom, simulating up and down respectively.

A more advanced form of mouse wheel is the tilt-wheel, found on some of the higher-end Logitech and Microsoft mice. Tilt wheels are essentially conventional mouse wheels that have been modified with a pair of sensors articulated to the tilting mechanism. These sensors are mapped, by default, to horizontal scrolling. In 2005, the Apple Mighty Mouse introduced a third variety of built-in scrolling device: a "scroll ball", which is essentially a trackball embedded in the upper surface of the mouse, and is used like a two-dimensional scroll wheel.

3D Mice

In the late 1990s, Kantek introduced the 3D RingMouse. This wireless mouse was worn on a ring around a finger, which enabled the thumb to access three buttons. The mouse was tracked in three dimensions by a base station. Despite a certain appeal, it was finally discontinued because it did not provide sufficient resolution.

Connectivity and communication protocols

To transmit their input typical cabled mice use a thin electrical cord terminating in a standard connector, such as RS-232C, PS/2, ADB or USB. Cordless mice instead transmit data via infrared radiation (see IrDA) or radio (including Bluetooth). The electrical interface and the format of the data transmitted by commonly available mice has in the past varied between different manufacturers.

PS/2 interface and protocol

With the introduction of the IBM PS/2 personal computer series in 1987, IBM introduced the homonyms PS/2 interface for mice and keyboards, which was rapidly adopted by other manufacturers. The most visible change was the use of a round 6-pin mini-DIN, in lieu of the former 5-pin connector. In default mode (called *stream mode*) a PS/2 mouse communicates motion, and the state of each button, by means of 3-byte packets.

Extensions: IntelliMouse and others

A Microsoft IntelliMouse relies on an extension of the PS/2 protocol: it initially operates in standard PS/2 format, for backwards compatibility. After the host sends a special command sequence, it switches to an extended format, where a fourth byte carries information about wheel movements. The IntelliMouse Explorer works analogously, with the difference that its 4-byte packets also allow for two additional buttons (for a total of five). The Typhoon mouse uses 6-byte packets which may be seen as a sequence of two standard 3-byte packets, and can thus be handled by ordinary PS/2 drivers. Other extended format are use by mouse vendors, often without public documentation.

Apple Desktop Bus

In 1986 Apple first implemented the Apple Desktop Bus allowing up to 16 devices, including arbitrarily many mice, to be daisy-chained together. Featuring only a single data pin, the bus used a purely polled approach to computer/mouse communications and survived as the standard on mainstream models until 1998 when the iMac began a switch to USB. The PowerBook G4 retained the Apple Desktop Bus for communication with its built in keyboard and trackpad until early 2005.

Tactile mice

In 2000, Logitech introduced the "tactile mouse", which contained a small actuator that made the mouse vibrate. Such a mouse could be used to augment user interfaces with haptic feedback, such as giving feedback when crossing a window boundary. Other unusual variants have included a mouse that is held freely in the hand, rather than on a flat surface, and detects six dimensions of motion (the three spatial dimensions, plus rotation on three axes). It was marketed for business presentations when the speaker is standing or walking around. So far, these mouse exotica have not achieved widespread popularity.

Mouse speed

Mouse speed is often expressed in DPI (dots per inch). One DPI is intended to be the number of pixels the mouse cursor will move when the mouse is moved one inch. However, software tricks like changeable mouse sensitivity can be used to make a cursor move faster or slower than its DPI, and cursor acceleration can be used to make the cursor accelerate when the mouse is moving at a constant speed. This makes "DPI" a confusing term ([3]); a replacement term, "CPI" (counts per inch), has been suggested by Apple and several other designers. A less common unit is the "Mickey" (named after Mickey Mouse). It is not a traditional unit of measurement because it indicates merely the number of "dots" reported in a particular direction. Only when combined with the DPI of the mouse does it become an indication of actual distance moved. In the absence of acceleration, the Mickey corresponds to the number of pixels moved on the computer screen.



Additionally, operating systems traditionally apply acceleration, referred to as ballistics, to the motion reported by the mouse. For example, versions of Windows prior to Windows XP doubled reported values above a configurable threshold, and then optionally doubled them again above a second configurable threshold. These doublings were applied separately in the X and Y directions, resulting in very nonlinear response. In Windows XP and many OS versions for Apple Macintosh computers, a smoother ballistics calculation is used that compensates for screen resolution and has better linearity.

"Mice" and "mouses"

The fourth (current as of 2006) edition of *The American Heritage Dictionary of the English Language* reports both "computer mice" and "computer mouses" as correct plural forms for "computer mouse." The traditional form "mice," however, is the most common choice, whereas some technical documents' authors may prefer the form "mouse devices."

Mice in the marketplace

In the 1970s, Xerox PARC included mice with its Xerox Star. Later, inspired by the Star, Apple Computer released the Apple Lisa, which also used a mouse. However, neither the Star nor the Lisa were commercially successful. Only with the release of the Apple Macintosh in 1984 did the mouse first see widespread use. The Macintosh design was influential, and its success led many other vendors to begin producing mice or including them with their other computer products. The widespread adoption of graphical user interfaces in the 1980s and 1990s made mice indispensable for computer use. As of 2000, Dataquest estimated that mice for a total cost of US\$1.5 billion were sold annually worldwide.

Other variants - Alternative devices

- Trackball - user moves a ball mounted in a fixed base.
- Mini-mouse - a small egg-sized mouse for use with laptop computers; it is usually small enough to be used on a free area of the laptop body itself.
- Camera mouse - a camera tracks the head movement and moves the onscreen cursor. Natural pointers track the dot on a person's head and move the cursor accordingly.
- Palm mouse - held in the palm and operated with only 2 buttons; the movements across the screen correspond to a feather touch, and pressure increases the speed of movement.
- Foot mouse - a mouse variant for those who do not wish to or cannot use the hands or the head; instead, footclicks are used.
- Tablet- It appears as a pen but is used as a mouse. It is held like a normal pen and is moved across a special pad. The thumb usually controls the clicking on a two-way button on the top of the mouse.
- Eyeball controlled- A mouse controlled by the user's eyeball/retina movements, allowing the cursor to be manipulated without touch.

Applications of mice in user interfaces

Usually, the mouse is used to control the motion of a cursor in two dimensions in a graphical user interface. Files, programs or actions can be selected from a list of names, or in graphical interfaces through pictures called "icons" and other elements. For example, a text file might be represented by a picture of a paper notebook, and clicking while the pointer hovers this icon might cause a text editing program to open the file in a window. (See also point-and-click). Mice can also be used *gesturally*; that is, a stylized motion of the mouse cursor itself, called gesture, can be

used as a form of command and mapped to a specific action. For example, in a drawing program, moving the mouse in a rapid "x" motion over a shape might delete the shape. Gestural interfaces are rarer, and often harder to use, than plain pointing and clicking, because they require finer motor control from the user. However, a few gestural conventions have become widespread, including the drag-and-drop gesture, in which the user presses the mouse button while the mouse cursor is over an interface object, moves the cursor to a different location while holding the button down, and then releases the mouse button. For example, a user might drag and drop a picture representing a file onto a picture of a trash can, indicating that the file should be deleted. Other uses of the mouse's input are common in special application domains. In interactive three-dimensional graphics, the mouse's motion is often directly translated into changes in the virtual camera's orientation. For example, in the Quake computer game, the mouse is usually used to control the direction in which the player's "head" faces: moving the mouse up will cause the player to look up, revealing the view above the player's head. When mice have more than one button, software may assign different functions to each button. Often, the primary (leftmost in a right-handed configuration) button on the mouse will select items, and the secondary (rightmost in a right-handed) button will bring up a menu of alternative actions applicable to that item. For example, on platforms with more than one button, the Mozilla web browser will follow a link in response to a primary button click, will bring up a contextual menu of alternative actions for that link in response to a secondary-button click, and will often open the link in a new tab or window in response to a click with the tertiary (middle) mouse button.

One, two or three mouse buttons?

The issue of whether a mouse should have exactly one button or more than one has attracted a surprising amount of controversy. From the first Macintosh until late 2005, Apple shipped computers with a single-button mouse, whereas most other platforms used a multi-button mouse. Apple and its advocates argued that single-button mice are more efficient, and that multi-button mice are confusing for novice users. The Macintosh user interface is designed so that all functions are available with a single button mouse. Apple's Human Interface Guidelines still specify that all functions need to be available with a single button mouse. However, X Window System applications, which Mac OS X can also run, were designed with the use of two or even three button mice in mind, causing even simple operations like "cut and paste" to become awkward. Mac OS X natively supports multi-button mice, so many users of older Macintoshes choose to use third-party mice on their machines. On August 2, 2005, Apple introduced their Mighty Mouse multi-button mouse, which has four independently programmable buttons and a "scroll ball" which can be used to scroll in any direction. This is now the mouse supplied with all new Macintosh computers.

Advocates of multiple-button mice point out that support for a single button mouse often leads to clumsy workarounds in interfaces where more than one action may be useful for a given object. There are several common workarounds, and even widely used Macintosh software packages that otherwise fully conform to the Human Interface Guidelines, including web browsers and graphics editing programs, occasionally require the use of one of them. One such workaround is the press-and-hold technique. In a press-and-hold, the user presses and holds the single button, and after a certain period, the button press is not perceived as a single click but as a separate action. This has two drawbacks: first, a slow user may press-and-hold inadvertently. Second, the user must wait while the software detects that the click is actually a press-and-hold, otherwise their press might be interpreted as a single click. Furthermore, the remedies for these two drawbacks conflict with each other: the longer the lag time, the more the user must wait; and the shorter the lag time, the more likely it is that some user will accidentally press-and-hold when meaning to click. Alternatively, the user may be required to hold down a key on the keyboard while pressing the button (otherwise known as mouse chording - Macintosh computers use the ctrl key). This has the disadvantage that it requires that both the user's hands be engaged. It also requires that the user perform two actions on completely separate devices in concert; that is, pressing a key on the keyboard while pressing a button on the mouse. This can be a very daunting task for a disabled user. Studies have found all of the above workarounds less usable than additional mouse buttons for experienced users.

Most machines running Unix or a Unix-like operating system run the X Window System which almost always requires a three button mouse. In X, the buttons are numbered by convention. This allows user instructions to apply to mice or pointing devices that do not use conventional button placement. For example, a left handed user may reverse the buttons, usually with a software setting. With non-conventional button placement, user directions that say "left mouse button" or "right mouse button" are confusing. The ground-breaking Xerox Parc Alto and Dorado computers from the mid-1970s used three-button mice, and each button was assigned a color. Red was used for the left (or primary) button, yellow for the middle (secondary), and blue for the right (meta or tertiary). This naming convention lives on in some SmallTalk environments, such as Squeak, and can be less confusing than the right, middle and left designations.

Mice in gaming

Mice are often used as an interface for PC-based computer games and sometimes for video game consoles. They are often used in combination with the keyboard. In arguments over which is the best gaming platform, the mouse is often cited as a possible advantage for the PC, depending on the gamer's personal preferences.

First-person shooters

A combination of mouse and keyboard is a popular way to play first-person shooter (FPS) games. The X axis of the mouse is used for looking (or turning) left and right, while the Y axis is used for looking up and down. The left mouse button is usually for primary fire. Many gamers prefer this over a gamepad or joystick because it allows them to turn quickly and have greater accuracy. The right button is often used for secondary fire of the selected gun, if the game supports multiple fire modes. It can also be used to throw grenades in games where grenades are supported. In a few games (usually historical/simulation fps), such as *Call of Duty 2*, it allows users to look down the barrel of the gun for better aiming. A scroll wheel is used for changing weapons, or controlling scope zoom magnification. On most FPS games, these functions may also be assigned to thumb buttons. A keyboard is usually used for movement (for example, WASD, for moving forward, left, backward and right, respectively) and other functions like changing posture. Since the mouse is used for aiming, a mouse that tracks movement accurately and with less lag will give a player an advantage over players with less accurate or slower mice. An early technique of players was circle strafing, where a player could continuously strafe while aiming and shooting an opponent by walking in circle around the opponent with the opponent at the center of the circle. This could be done by holding down a key for strafing while continuously aiming the mouse towards the opponent.

Invert mouse setting

In many games, such as first or third person shooters, there is a setting named "invert mouse" or similar (not to be confused with "button inversion", sometimes performed by left-handed users). It allows the user to look downward by moving the mouse forward, and upward by moving the mouse backward (the opposite of the default setting). This control system is similar to aircraft control sticks, where pulling back causes pitch up and pushing forward causes pitch down; this control configuration is also typically mimicked in computer joysticks.

After id Software's *Doom*, the game that popularized FPS games, but which did not support vertical aiming with a mouse (the y-axis was used for forward/backward movement), competitor 3D Realms' *Duke Nukem 3D* was one of the first games that supported using the mouse to aim up and down. It and other games using the Build engine had an option to invert the Y-axis (moving the mouse forward aims up, moving the mouse backward aims down). The "invert" feature actually made the mouse behave in the way that we now regard as normal. Soon after, id Software released *Quake* which introduced the invert feature as we know it now. Other games using the Quake engine were released and kept this feature. Probably because of the overall popularity of *Quake*, this became the current standard.

4.2.2. Touchpad

A touchpad is an input device commonly used in laptop computers. They are used to move the cursor, using motions of the user's finger. They are a substitute for a computer mouse. Touchpads vary in size but are rarely made larger than 50 cm² (8 in²). They can also be found in PDAs. Touchpads commonly operate by sensing the capacitance of a finger, or the capacitance between sensors. Capacitive sensors are laid out along the horizontal and vertical axis of the touchpad. The location of the finger is determined from the pattern of capacitance from these sensors. This is why they will not sense the tip of a pencil or even a finger in a glove. Moist and/or sweaty fingers can be problematic for those touchpads that rely on measuring the capacitance between the sensors.



Touchpads are relative motion devices. That is, there is no isomorphism from the screen to the touchpad. Instead, relative motion of the user's fingers causes relative motion of the cursor. The buttons below or above the pad serve as standard mouse buttons. Depending on the model of touchpad and drivers behind it, you may also click by tapping your finger on the touchpad, and drag with a tap following by a continuous pointing motion (a *click-and-a-half*). Some touchpads also have "hotspots": locations on the touchpad that indicate user intentions other than pointing. For example, on certain touchpads, moving your finger along the right edge of the touch pad will control the scrollbar and scroll in horizontal direction. Moving the finger on the bottom of the touchpad often scrolls in vertical direction. Some touchpads can emulate multiple mouse buttons by either tapping in a special corner of the pad, or by tapping with two or more fingers.



Touchpads are primarily used in portable laptop computers, because the usual mouse device requires a flat table adjacent to the keyboard not always available away from the office. But touchpads have some advantages over the mouse, particularly that the pad's position is fixed relative to the keyboard, and very short finger movements are required to move the cursor across the display screen. Some computer users prefer them for such reasons, and desktop keyboards with built-in touchpads are available from specialist manufacturers. Touchpads have also recently appeared in Apple's iPod. The main control interface for menu navigation in all of the currently produced iPods (except the Shuffle) is a touchpad (at first by Synaptics; Apple now manufacture that component themselves). Creative Labs also uses a touchpad in their Nomad Jukebox Zen line with the Zen Touch, Zen Sleek (Photo) and Zen Micro(Photo).

The "trackpad" is Apple Computer's name for the touchpad. It was introduced in 1994 in the PowerBook 500 series, the first laptop ever to carry such a device, and replaced the trackball of previous PowerBook models. Late generation PowerBooks and iBooks have two finger sensing capabilities, as well as the current MacBook and MacBook Pro model lines.

4.2.3. Graphics tablet

A graphics tablet is a computer peripheral device that allows one to hand-draw images directly into a computer, generally through an imaging program. Graphics tablets consist of a flat surface upon which the user may "draw" an image using an attached stylus, a pen-like drawing apparatus. The image generally does not appear on the tablet itself but, rather, is displayed on the computer monitor. It is interesting to note that the stylus, as a technology, was originally designed as a part of the electronics, but later it simply took on the role of providing a smooth, but accurate "point" that would not damage the tablet surface while "drawing". Graphics tablets should not be confused with the tablet PC.

History and background

Early graphics tablets, known as spark or acoustic tablets, used a stylus that generated clicks with a spark plug. The clicks were then triangulated by a series of microphones to locate the pen in space. The system was fairly complex and expensive, and the sensors were susceptible to interference by external noise. The first graphics tablet resembling contemporary tablets was the RAND Tablet, also known as the Grafacon (for Graphic Converter), introduced in 1964. The RAND Tablet employed a grid of wires under the surface of the pad that encoded horizontal and vertical coordinates in a small magnetic signal. The stylus would receive the magnetic signal, which could then be decoded back as coordinate information.

Digitizers were popularized in the mid 1970's and early 1980's by the commercial success of the ID (Intelligent Digitizer) and BitPad manufactured by the Summagraphics Corp. These digitizers were used as the input device for many high-end CAD (Computer Aided Design) systems as well as bundled with PC's and PC based CADD software like AutoCad. Summagraphics also made a OEM version of its BitPad which was sold by Apple as an accessory to their Apple II. These tablets used a magneto-strictive technology which used wires made of a special alloy stretched over a solid substrate to accurately locate the tip of a stylus or the center of a digitizer cursor on the surface of the tablet. This technology also allowed Proximity or "Z" axis measurement.

The first home computer graphics tablet was the KoalaPad. Though originally designed for the Apple II, the Koala eventually broadened its applicability to practically all home computers with graphics support, examples of which include the TRS-80 Color Computer, Commodore 64, and Atari 8-bit family. Competing tablets were eventually produced; the tablets produced by Atari were generally considered to be of high quality. Modern graphics tablets operate in a fashion similar to the RAND Tablet. In modern devices, though, the horizontal and vertical wires of the grid are separated by a thin insulator. When pressure is applied to the tablet, the horizontal wire and vertical wire associated with the corresponding grid point meet each other, causing an electric current to flow into each of these wires. Since an electric current is only present in the two wires that meet, a unique coordinate for the stylus can be retrieved. Pointing device detection is additionally aided by a weak magnetic field projected approximately one inch from the tablet surface. It is important to note that, unlike the RAND Tablet, modern tablets do not require electronics in the stylus and any tool that provides an accurate "point" may be used with the pad (though the stylus itself is likely designed with both the resolution of the grid and the durability of the tablet in mind and, therefore, should be used). Graphics tablets produced by Wacom primarily make use of induction technology, where the tablet itself operates as



both a transmitting and receiving coil. The tablet generates a signal, which is then received in the tank circuit in the pen. Changing the pressure on the stylus or pressing a switch changes the capacitance, which is then reflected in the signal generated by the pen.

Some modern arrangements also provide pressure information, but the electronics for this information are present in the stylus nub, not the tablet. Graphics tablets are available in various sizes and price ranges; A6-sized tablets being relatively inexpensive and A3-sized tablets being far more expensive. Modern tablets usually connect to the computer via a USB interface.

General uses

Graphics tablets, because of their stylus-based interface and (in some cases) ability to detect pressure, tilt, and other attributes of the stylus and its interaction with the tablet, are widely considered to offer a very natural way to create computer graphics, especially two-dimensional computer graphics. Indeed, many graphics packages (e.g. Corel Painter, Inkscape, Photoshop, Pixel image editor, Studio Artist, The GIMP) are able to make use of the pressure (and, in some cases, stylus tilt) information generated by a tablet, by modifying attributes such as brush size, opacity, and color based on data received from the graphics tablet.

In East Asia, graphics tablets, or pen tablets as they are known, are widely used in conjunction with input method editor software {IMEs) to write Chinese, Japanese, Korean characters (CJK). The technology is popular and inexpensive, and companies such as Twinbridge Software, among others, are able to provide a pen tablet complete with English language technical manuals and customer support, based in Los Angeles, California, for around US\$100. Tablets are also popular for technical drawings and CAD, as one can put a piece of paper on them without interfering with their function. Many of the most successful webcomic artists use tablets, including Hawk of AppleGeeks, who uses a Wacom Cintiq to color directly into the computer. Finally, tablets are gaining popularity as a replacement for the computer mouse as a pointing device. Advocates cite relief from repetitive strain injury and greater intuitiveness[1].

Examples

In high-end computer graphics systems, the graphics tablet has always been the input device of choice. Some examples are the Crosfield imaging system and the Quantel Paintbox. Devices similar to graphics tablets, such as light pen-based systems, have been used to control other types of systems, such as the Fairlight CMI computer musical instrument.

Similar devices

Touch screens are operated in similar ways, but they usually use either optical grids or a pressure sensitive film instead, and therefore they do not need a special pointing device. The development of the Tablet PC is another example of the integration of graphics tablet and screen.

4.2.4. Light pen

A light pen is a computer input device in the form of a light-sensitive wand used in conjunction with the computer's CRT monitor. It allows the user to point to displayed objects, or draw on the screen, in a similar way to a touch screen but with greater positional accuracy. A light pen can work with any CRT-based monitor, but not with LCD screens, projectors or other display devices. A light pen is fairly simple to implement. The light pen works by sensing the sudden small change in brightness of a point on the screen when the electron gun refreshes that spot. By noting exactly where the scanning has reached at that moment, the X,Y position of the pen can be resolved. This is usually achieved by the light pen causing an interrupt, at which point the scan position can be read from a special register, or computed from a counter or timer. The pen position is updated on every refresh of the screen.

The light pen became moderately popular during the early 1980s. It was notable for its use in the Fairlight CMI, and the BBC Micro. However, due to the fact that the user was required to hold his or her arm in front of the screen for long periods of time, the light pen fell out of use as a general purpose input device. The first light pen was used around 1957 on the Lincoln TX-0 computer at the MIT Lincoln Laboratory. Contestants on the game show *Jeopardy!* use a light pen to write down their wagers and answers for the *Final Jeopardy!* round. Light pens are used country-wide in Belgium for voting.

A computer light pen system is a man-machine interface control system. In addition to a human operator and a conventional computer, with a Cathode-Ray-Tube (CRT) display screen, the system includes a CRT light pen, light pen interface hardware, and a light pen software driver. The operator initiates functions of the control system by pointing the light pen to an area of the CRT screen, related to the functions, and depressing one of the light pen buttons. By repeating this simple, intuitive pointing to the screen and depressing buttons, the operator may control the performance of complex hardware, software, and equipment control functions determined by application software. The utility, efficiency, and cost effectiveness of the computer light pen systems in many applications is well established in the industry.

Most conventional computers today may be configured with CRT displays. CRT displays, or monitors, are the display of choice for most computer systems requiring high data display density. A light pen is ideally suited for pointing on high data density CRT screens because of the high accuracy and resolution attainable with light pen systems. The CRT light pen detects light in its field-of-view on the CRT screen, and provides light and button data to the light pen interface hardware. The interface in addition receives CRT synchronization data, and provides pen position and button data to the light pen driver. The light pen driver software then provides input to the computer operating system to accomplish the desired application function. The CRT display or monitor is the key element in the man-machine interface of the computer light pen system. The data displayed on the CRT each frame is defined by the computer video controller. The CRT receives frame video data and synchronization signals, and illuminates the CRT screen using the raster scanning method. Video frame rates are typically 60 +/- 10 Hertz. The video of each frame is raster scanned onto the CRT screen beginning in the upper left-hand corner. Video is scanned onto the screen one dot, or pixel, at a time, until the uppermost line is formed by the sequence of video dots. When the first line is completed, the raster scan then forms the next line below, again beginning at the left of the screen, but one pixel lower than the first line. This process continues down the screen until all lines of the frame are scanned, completing one frame of video data. Due to the high video frame rate of data scanned on the CRT screen, compared to the light sensing rate of the human eye, the operator sees an apparently static display which the brain interprets accordingly, and the operator uses to his end. The light pen, however, processes light data at rates exceeding the video data rate, and thus easily differentiates pixels and lines necessary to the pointing function.

Rather than averaging visual CRT screen data, as the human eye does, the CRT and the lightpen function essentially upon one pixel of video at a time. The CRT illuminates a screen pixel by directing electrons to the phosphor material behind the glass of the screen. Electrons striking the

phosphor produce photons of visible light. The light pen pointed at the CRT screen, detects light within its field-of-view. Light pen was invented at IBM in the late 1950s. In the 1960s the founders of Interactive designed the first CRT light pen using semiconductor detectors. Since then there have been significant improvements in components and technology resulting in mature, low cost products of great utility in many applications. The light pen is more correctly called a CRT light pen. The light pen is designed specifically to detect light emanating from a CRT screen, and convert the light signals to equivalent electrical signals. The CRT light pen is not suitable for use with flat panel or LCD displays, although they are used with pulsed LED display matrices.

The key element of a light pen in the conversion of light to electrical signals is the photodetector. The photodetector is located near the tip of the pen with an aperture or other optics to receive CRT light emanations. When the light pen tip is pointed at the CRT screen, light received by the photodetector produces video current in the detector, which is proportional to the light received. The form and the timing of the photodetector current is identical with that of the CRT screen phosphor illuminance resulting from impinging CRT electrons associated with the spot of interest. The photodetector video is amplified, differentiated, and formed into a light pulse output of the pen. The light pulse so processed contains the basic light timing information used in the interface to provide precise position data to the light pen driver software, and thus to the computer OS. The light pen also may include up to 3 switches operated by buttons. Operating these switches produce associated outputs provided to the interface unit. These outputs are used to initiate computer operations associated with the spot of light under the light pen tip. Further information may be indicated to the computer depending upon which button is pressed, while pointing at a particular screen spot. For example pressing button 1, may indicate a left mouse function, button 2 a right mouse, etc. Although the theory of operation as presented above is relatively simple, in practice light pen configurations involve considered trade-offs of optical, electrical, and mechanical parameters. Details of design trade-offs are beyond the scope of this document; but include consideration of the high noise environment of the CRT screen, signal-to-noise, cross-talk between switching circuits and low level video circuits, and mechanical constraints on hand-held instruments.

The light pen interface hardware provides three basic functions: 1) It receives light pen video and button signals; video sync signals, and DC power inputs. 2) From these inputs, the light pen screen position and button press data is determined. 3) The data is then formatted and send to the host computer, via a suitable computer input port or bus. The interface receives light pulses and button press inputs from the light pen via the pen cable/connector. The video horizontal and vertical sync signals are received by the interface via a "Y" sync cable connected between the computer video controller and the monitor. DC power is received from the computer port or bus, or from an external DC supply. DC power is conditioned, as required, and provide to both the interface and the light pen. The absolute position of the light pen on the screen is determined by comparing the timing of the light pulse output of the pen with the state of vertical and horizontal counters synchronized by the video sync signals from the video controller. A vertical sync signal occurs at the beginning of each frame. A horizontal sync signal occurs at the beginning of each line in each frame. The coincidence of a vertical sync and a horizontal sync occurs at the beginning of each frame. Based of these facts, the combinations of synchronized counter states may be used to form a timing matrix, related directly the physical position of the pen on the screen, against which the light pulse timing can be compared, to determine the absolute position of the pen on the screen. The accuracy of the absolute position of the pen position is determined by size of the two counters and the clock rate used to increment the counters. The size of the counters and the clock rate is usually selected to provide accuracy of one pixel or less at the maximum screen resolution of the monitor.

The interface also processes light pen button press data, and forms interrupt, proximity, and interface address data. This data and the pen position data is then formatted to be sent to the light pen driver software via a defined computer port or bus. Light pen interface units are commonly available for serial port, USB port, and the ISA and PCI bus communication on IBM PCs and compatibles. Current interface designs may be implemented using micro controllers, PALs, and ASICs, as well as, more conventional devices.

4.3. Game controllers

A **game controller** is an input device used to control a video game. A controller is typically connected to a video game console or a personal computer. A game controller can be a keyboard, mouse, gamepad, joystick, paddle, or any other device designed for gaming that can receive input. Special purpose devices, such as steering wheels for driving games and light guns for shooting games, may also exist for a platform. Some devices, such as keyboards and mice, are actually generic input devices and their use is not strictly limited to that of a game controller.

Game controllers can be used to govern the movement or actions of elements in a video or computer game. The type of element controlled depends upon the game, but a typical element controlled would be the game hero. A gamepad, the most common kind of game controller, can have anywhere from a couple of buttons to a dozen or more, combined with multiple omnidirectional control sticks. This lets it control the game elements' movement in up to three dimensions, with many buttons to perform quick actions. Due to the ease of use and precision of gamepads, they have spread from traditional consoles where they originated to computers as a common input device.

4.3.0. Gamepad

A gamepad, also known as a joypad, is a type of game controller held with both hands where the thumbs are used to provide input. Gamepads generally feature a set of action buttons handled with the right thumb and a direction controller handled with the left. Most modern game controllers are a variation of a standard gamepad. Some common additions to the standard pad include shoulder buttons placed along the edges of the pad, centrally placed buttons labeled *start*, *select*, and *mode*, and an internal motor to provide force feedback. Gamepads are the primary means of input on nearly all modern video game consoles. Gamepads are also available for personal computers, but few computer games support gamepads, preferring the more conventional keyboards and mice. However, most console emulators support gamepads.



🖱️ A range of gamepads for SEGA video game consoles. Left to right, top-to-bottom: Master System controller (1985), Genesis/Mega Drive controllers (1988+), Digital Saturn controller (1994), Saturn 3D Pad (1996), and Dreamcast controller (1998).

Nintendo Entertainment System: the D-pad

The NES controller used their patented + shaped D-pad (a modified version of Milton Bradley's Cosmic Hunter Dpad concept) which was used by their "Game and Watch" series of games as the standard for their home console controllers. The NES and Famicom controller featured a brick-like design with a simple, four button layout: two buttons labelled "A" and "B," a "start" button, and a "select" button. Near the end of the NES's lifespan, upon the release of the AV Famicom and the NES 2, the design of the game controller was modified slightly, abandoning the "brick" shell in favor of a "dog bone" shape, reminiscent of the controllers of the Super Nintendo Entertainment System. The original Famicom controllers were different from the NES ones in the sense that, besides their design (and being hardwired into the console), the second controller lacked Select and Start buttons- in their place was a microphone, although very few games supported it.

SNES: introducing shoulder buttons

The SNES controller had a more rounded dog-bone like design and added two more face buttons, "X" and "Y", arranging the four in a diamond formation. Another addition was the "L" and "R" shoulder buttons, which have been imitated by most controllers since. There is a slight variation in the Japanese and European version of the controller, where the A, B, Y and X buttons are all convex-curved and have each one a distinct color. For the American SNES, the A and B buttons were colored purple, whereas the Y and X buttons were lavender and concave-curved.



Virtual Boy: dual DPads

The Virtual Boy controller was a controller which utilized dual gamepads similar to how analog sticks functioned in later "dual control" sixth-generation systems. The presence of two pads was an effort to control objects in a 3D environment (one pad controlling pitch and turning while the other controlling forward movement and strafing). The "two gamepad" setup has since been used in the GameCube controller.

N64: analog stick in a gamepad

The Nintendo 64 controller was one of the first in a trend to have both an analog stick and a D-pad. It has the traditional A, B, L, and R buttons, along with a Z trigger button on its underside. Four "C" buttons are used mainly for controlling the camera in games. In addition to the Rumble Pak, the controller also can house a memory pack for saving games, and a microphone add-on. The Nintendo 64 controller's design seems controversial with its trident shape, making for three ways to hold the controller. The most used way in games to hold the controller was for the left hand to hold the center grip, so the thumb could move the analog stick and the index finger could press the Z trigger. The right hand would be on the right grip of the controller and have access to all the main face buttons, and the R button.

Gamecube: dual analog stick and wireless gaming

The Nintendo GameCube controller adopted a similar style to the PlayStation DualShock. It has two analog sticks, a smaller traditional D-pad, and four main face buttons. The GameCube controller also has pressure sensitive L and R buttons that *click* when pressed down completely, a Z button located above the R button, and includes a built in rumble feature. Nintendo later introduced the Wavebird controller, a wireless pad that uses radio frequency technology refined from the Atari Wireless RF controllers, in order to operate without relying on infrared. The Wavebird overall has

the same layout, but doesn't include force feedback (in order to save battery life). It operates using two AA batteries and ranges about 100 hours of gameplay.

DS: touch screen

While the Nintendo DS was not the first implementation of the touch screen as an interface, nor was it the first implementation of a touch screen in a video game system (see Game.com), it certainly is one of the more popular gaming examples. Nintendo has integrated the touch screen into its standard controls (L, R, A, B, Y, X, Start, Select and a D-pad) by use of a stylus and it is most commonly used in lieu of the D-pad or the A, B, Y, and X buttons, depending on the handedness of the user.

Microsoft

Sidewinder range

The Microsoft Sidewinder range was presented in 1996, and grew from a simple joystick to include all kinds of gaming devices, from gamepads to the innovative *Sidewinder Strategic Commander*, aimed to real-time strategy players. Until the first gamepads based on the popular *Dual Shock* design by Sony appeared for personal computers, the *Sidewinder Gamepad*, with its 10 buttons (six buttons plus two triggers, mode and start) and pass-through connection which allowed up to four controllers connected with only one game port was the premier gamepad for sports games on the platform, and most games between 1998 and 2002 were developed with the Sidewinder in mind.



Xbox

Similar in design to the Dreamcast controller, the Microsoft Xbox controller includes two expansion slots, six analog buttons, two analog triggers, and two analog sticks, a total of eight digital buttons (four of which make up the d-pad), as well as built in rumble support. Differing from the Dreamcast controller, the Microsoft controller adds three buttons, the "black", "white" and "back" (select) buttons. The Xbox controller went through a revision specifically for Japanese consumers and due to complaints that the initial controller was too bulky (Indeed, the size of the Xbox controller has been the subject of many a joke in video game-related webcomics). The result was the Type-S controller which Microsoft adopted and has since bundled with their system in all regions. Both of the original Xbox controllers had a breakaway point near the end, so if the cord was pulled, the cord would pop into two parts the game would pause, instead of the Xbox being pulled to the floor and possibly injuring somebody. The Xbox 360 controller has wireless capabilities and removes the "black" and "white" buttons and in their place adds two "bumper" buttons, one above each trigger. The wired version uses a USB port to plug into the 360, and is going to be the main gamepad to be used for Windows Vista.



4.3.1. Paddle

A paddle is a controller that features a round wheel and one or more fire buttons. The wheel is used to typically control movement of the player or an object along one axis of the video screen. Paddle controllers were the first analog controllers; they died out when "paddle and ball" type games fell out of favor.

4.3.2. Trackball

A trackball is basically an upside-down mouse that is manipulated with the palm of one's hand. It has the advantage of not requiring a lot of desktop space, and that it is as fast as one can roll the ball on it. This is faster than one can move a mouse due to space not being an issue. Notable uses of a Trackball as a gaming controller would be games such as Centipede and Golden Tee.



4.3.3. Joystick

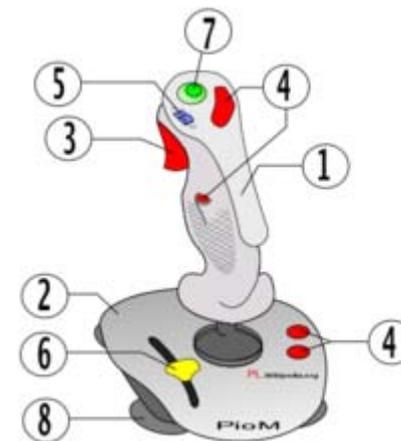
This joystick is a computer peripheral that consists of a handheld stick that pivots about one end and transmits its angle in two or three dimensions to a computer. The joystick is often used for flight simulators. HOTAS controllers, which include extra hardware to simulate throttle and rudder controls are popular among fanatics of the genre.

This device can also be a home version of the kind used in the arcade, or for use on emulators. It features a shaft that has a sphere-like handle that is grasped, and one or more buttons for in game actions. Generally the layout has the joystick on the left, and the buttons on the right. There are instances when this is reversed, or the joystick is in the center with button on both sides.

Joystick elements:

1. Stick
2. Base
3. Trigger
4. Extra buttons
5. Autofire switch
6. Throttle
7. Hat Switch (POV Hat)
8. Suction Cup

A **joystick** is a personal computer peripheral or general control device consisting of a handheld stick that pivots about one end and transmits its angle in two or three dimensions to a computer. Most joysticks are two-dimensional, having two axes of movement (similar to a mouse), but three-dimensional joysticks do exist. A joystick is generally configured so that moving the stick left or right signals movement along the X axis, and moving it forward (up) or back (down) signals movement along the Y axis. In joysticks that are configured for three-dimensional movement, twisting the stick left (counter-clockwise) or right (clockwise) signals movement



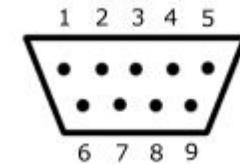
along the Z axis. These three axis - X Y and Z - are, in relation to an aircraft, roll, pitch, and yaw, respectively.

Joysticks are often used to control games, and usually have one or more push-buttons whose state can also be read by the computer. Most I/O interface cards for PCs have a joystick (game control) port. Joysticks were popular throughout the mid-1990s for playing the Descent series and flight-simulators, although they have declined with the rise of first-person shooters which instead promote the mouse and keyboard. Modern joysticks (as of 2003) mostly use a USB interface for connection to the PC. The term joystick has become a synonym for game controllers that can be connected to the computer since the computer defines the input as a "joystick input".

Apart for controlling games, joysticks are also used for controlling machines such as elevators, cranes, trucks, powered wheelchairs and some zero turning radius lawn mowers.

Technical details

An *analog joystick* is a joystick which has continuous states, i.e. returns an angle measure of the movement in any direction in the plane or the space (usually utilizing potentiometers) and a *digital joystick* gives only on/off signals for four different directions, and mechanically possible combinations (such as up-right, down-left, &c.). (Digital joysticks were very common as game controllers for the video game consoles, arcade machines, and home computers of the 1980s.) Additionally joysticks often have one or more *fire buttons*, used to trigger some kind of action. These are simple on/off switches. Some joysticks have force feedback capability. These are thus active devices not just simple input devices. The computer can return a signal to the joystick that causes it to resist the movement with a returning force or make the joystick vibrate.



Computer port view

- | | |
|-------------|------------------|
| 1, up | 6, fire button |
| 2, down | 7, + 5VDC |
| 3, left | 8, ground |
| 4, right | 9, not connected |
| 5, reserved | |

History

Joysticks were originally controls for an aircraft's ailerons and elevators. They were fixed to the floor of the aircraft and stuck up between the pilot's legs, hence the name. The name 'joystick' is thought to originate with early 20th century French pilot Robert Esnault-Pelterie ^[1]. There are also competing claims on behalf of fellow pilots Robert Loraine and James Henry Joyce. The joystick itself was present in early planes, however the mechanical origins themselves are uncertain. The first electrical 2-axis joystick was probably invented around 1944 in Germany. The device was developed for targeting the glide bomb *Henschel Hs 293* against ship targets. Here, the joystick was used by an operator to steer the missile towards its target by radio control. This joystick had on-off switches rather than analogue sensors, i.e. a *digital joystick*. The signal was transmitted from the joystick to the missile by a thin wire.

This invention was picked up by someone in the team of scientist assembled at the *Heeresversuchsanstalt* in Peenemünde. Here a part of the team on the German rocket program was developing the Wasserfall missile, a successor to the V-2 rocket, the first ground-to-air missile intended for shooting down enemy aircraft. The Wasserfall steering equipment converted the electrical signal to radio signals and transmitted these to the missile.

The Atari standard joystick, developed for the Atari 2600 was a *digital joystick*, with a single 'fire' button, and connected via a DB9 connector, the electrical specifications for which was for many years the 'standard' digital joystick specification. Joysticks were commonly used as controllers in first and second generation game consoles, but then gave way to the familiar control pad with the Nintendo Entertainment System and Sega Master

System in 1985 and 86, though joysticks - especially arcade-style ones - were and are popular after-market add-ons for any console. More recently, *analog joysticks* have become standard on video game consoles and have the ability to indicate the stick's displacement from its neutral position. This means that the software does not have to keep track of the position or estimate the speed at which the controls are moved. These devices are usually a rotary potentiometer (variable resistance) turned by the stick movement.

Assistive Technology

Specialist joysticks, classed as an assistive technology pointing device, are used to replace the computer mouse for people with fairly severe physical disabilities. Rather than controlling games these joysticks plug into the USB port and control the mouse pointer. They are often useful to people with athetoid conditions, such as cerebral palsy, who find them easier to grasp than a standard mouse. Miniature joysticks are also available for people with conditions involving muscular weakness such as muscular dystrophy or motor neurone disease.

4.3.4. Steering wheel

The steering wheel, essentially a larger version of a paddle, is used for racing simulators such as *Gran Turismo*, *Forza Motorsport*, and *Live for Speed*. Many are force feedback (see Force Feedback Wheel), designed to give the same feedback as you would driving a real car, but the realism of this depends on the game. They usually come with pedals to control the gas and brake. Shifting is taken care of with either paddles, a simple shifter which is moved forward or back to change gears, or a shifter which mimics that of real vehicles, which may use a clutch. Most wheels turn only 200 to 270 degrees lock-to-lock, but some models, such as the Logitech driving force pro, can turn 900 degrees, or 2.5 turns lock-to-lock. The Namco JogCon paddle was available for the PlayStation game *R4: Ridge Racer Type 4*. Unlike "real" video game steering wheels, the JogCon was designed to fit in the player's hand. Its much smaller wheel (diameter roughly similar to a soda can's) resembles the jog-and-shuttle control wheel used on some VCRs.

4.3.5. Light gun

A light gun is a peripheral used to "shoot" targets on a screen. They usually roughly resemble firearms or ray guns. Their use is normally limited to rail shooter or shooting gallery games like those that came with the "Shooting Gallery" light gun. The first home console light gun was released on the Magnavox Odyssey; later on, Nintendo would include one standard on their Famicom and NES. The NES Zapper, Nintendo's light gun and regarded as the quintessential device of this nature. Shown is the grey version, which would later be replaced by an orange edition due to striking similarities with real arms. NES Zapper. The NES Zapper (or Famicom Light Gun in Japan) is a pistol-shaped electronic light gun sold as part of the original Nintendo Entertainment System in 1985 and the Nintendo Famicom in 1984. The Japanese iteration was a realistic portrayal of a revolver style hand gun, but when released in North America was made to look like an unrealistic science fiction ray gun that also tied into the design of the NES. The shape of the gun is comfortable, but there is no real grip to speak of. Early Zappers were gray, but later the color was changed to a neon orange due to the fact that many likened the device to an actual gun. It allowed players to aim at the TV and, depending on the game, shoot pixelated ducks, clay pigeons, gangsters, and many other obvious targets. The Zapper has a trigger action that takes quite a bit of force to pull back, and frequent trigger-pulling has caused many people to get tired muscles in their hands or forearms.



(Modern light guns have much less resistance on their triggers, in comparison.) Also, the trigger makes a loud characteristic 'pang' noise when pulled.

Technical overview

When the trigger was pulled, the game blanked out the screen with a black background for one frame, then, for one additional frame, drew a solid white rectangle around the sprite the user was supposed to be shooting at. The photodiode at the back of the Zapper would detect these changes in intensity and send a signal to the NES to indicate whether it was over a lit pixel or not. A drop followed by a spike in intensity signaled a hit. Multiple sprites were supported by flashing a solid white rectangle around each potential sprite, one per frame. It was possible to cheat in games by changing the brightness and contrast of the TV, or pointing the gun at a bright light. The gun would think it was pointing at a solid white target and report a hit. If there were multiple targets, the "hit" target was the first to be lit with the white square. Note: this was not as consistent with a fluorescent lamp due to the possible differences in timing of the lamp's flicker vs the timing of the TV's refresh rate. This cheat, especially pointing at an incandescent light bulb, seems to go against the drop - spike requirement, but it worked (and killed game play). Warning: changing the brightness and contrast so this cheat works can *greatly* decrease the life of the TV and cause it to look bright and washed out. Pointing at a bare bulb seems to work better than covered, and cheap incandescent bulbs seem to work better than higher quality bulbs.



A **light gun** is a pointing device for computers and a control device for arcade and video games. The first light guns appeared in the 1930s, following the development of light-sensing vacuum tubes. It wasn't long before the technology began appearing in arcade shooting games, beginning with the Seeburg Ray-O-Lite in 1936. These early light gun games used small targets (usually moving) onto which a light-sensing tube was mounted; the player used a gun (usually a rifle) that emitted a beam of light when the trigger was pulled. If the beam struck the target, a "hit" was scored. Modern screen-based light guns work on the opposite principle -- the sensor is built into the gun itself, and the on-screen target(s) emit light rather than the gun. The first light gun of this type was used on the MIT Whirlwind computer. The light gun, and its descendant, the light pen, are now rarely used as computer pointing devices, because of the popularity of the mouse and changes in monitor display technology - light guns can only work with standard CRT monitors.



Light guns in video games

The video game light gun is typically modeled on a ballistic weapon (usually a pistol) and is used for targeting objects on a video screen. With force feedback, the light gun can also simulate the recoil of the weapon. Light guns are very popular in arcade games, but have never caught on in the home video game console market after the NES, Master System, Genesis/MegaDrive, and SNES systems. This may be because people are reluctant to buy more than one extra controller for their system, let alone a special-purpose and often expensive peripheral, or because light guns are less satisfactory to use with the small television screens in peoples homes than on the large screens in arcade game cabinets. The most popular example of the light gun is Nintendo's Zapper Gun for the Nintendo Entertainment System, though there are also light guns for Sega Master System, Sega Saturn, Sony PlayStation, Sega Dreamcast, Microsoft Xbox, Magnavox Odyssey and several other console and arcade systems. Recent light gun video games include *Time Crisis 3*, *Virtua Cop 3*, and *The House of the Dead III*.

How light guns work

The "light gun" is so named because it uses light as its method of detecting where on screen you are targeting. The name leads one to believe that the gun itself emits a beam of light, but in fact most light guns actually *receive* light through a photoreceptor diode in the gun barrel. [Re]. There are two versions of this technique that are commonly used, but the concept is the same: when the trigger of the gun is pulled, the screen is blanked out to black, and the diode begins reception. All or part of the screen is painted white in a way that allows the computer to judge where the gun is pointing, based on when the diode detects light. The user of the light gun notices little or nothing, because the period in which the screen is blank is very short.

Method one

The first detection method, used by the Zapper, involves drawing each target sequentially in white light after the screen blacks out. The computer knows that if the diode detects light as it is drawing a square (or after the screen refreshes), that is the target the gun is pointed at. Essentially, the diode tells the computer whether or not you hit something, and for n objects, the sequence of the drawing of the targets tell the computer *which* target you hit after $1 + \text{ceil}(\log_2(n))$ refreshes (one refresh to determine if any target at all was hit and $\text{ceil}(\log_2(n))$ to do a binary search for the object that was hit). An interesting side effect of this is that on poorly designed games, often a player can point the gun at a light bulb, pull the trigger and hit the first target every time. Better games account for this by not using the first target for anything.

Method two

The second method, used by the Super Nintendo Entertainment System's Super Scope and computer light pens is more elaborate but more accurate. The trick to this method lies in the nature of the cathode ray tube inside the video monitor (it does not work with LCD screens, projectors, etc.). The screen is drawn by a scanning electron beam that travels across the screen starting at the top until it hits the end, and then moves down to update the next line. This is done repeatedly until the entire screen is drawn, and appears instantaneous to the human eye as it is done very quickly. When the player pulls the trigger, the game brightens the entire screen for a split second, and the computer (often assisted by the display circuitry) times how long it takes the electron beam to excite the phosphor at the location the gun is pointed at. It then calculates the targeted position based on the monitor's horizontal refresh rate (the fixed amount of time it takes the beam to get from the left to right side of the screen).

Method three

With the advent of LCD screens where the second method is useless due to the synchronous refresh of all pixels, a new method was developed. It relies on one or several Infrared emitters placed near the screen, and one IR sensor on the muzzle of the gun. When the trigger is pressed, the gun sends the intensity of the IR beam it detects. Since this intensity depends upon both distance and relative angle to the screen, angle sensors are located in the gun. This way a trigonometric equation system is solved, and muzzle's 3D position relative to the screen is calculated. Then, by projecting the muzzle on the screen with the measured angles the impact point is determined. A simpler variant is commonly used in arcades, where there are no angle detectors but 4 IR sensors. However, this can prove inaccurate when shooting from certain distances and angles, since angles and 3D position are inferred, not calculated. Other variants include 3 or more emitters with different wavelength and the same number of sensors. With this method and proper calibration three or more relative angles are obtained, thus not needing angle detectors to positionate the gun. Sometimes, sensors are placed around the screen and the emitter on the gun, but calculations are similar. This family of methods are used on Wii and modern arcade systems.

Combining the methods

Some light guns designed for method one are not timed precisely enough for method two to work exactly as described, but they can use a combination of the two methods. First the screen is brightened and the response time is measured as in method two, but the computer measures only which scanline was hit and not which horizontal pixel was hit. This does not need nearly as fast a timer that pure method 2 uses. Then using method one, the game cycles among those targets on the line.

Detection

Once the computer knows where the gun is pointed, it can tell if it coincides with the target or not. However, many guns of this type (including the Super Scope) ignore red light, as red phosphors have a much slower rate of decay than green or blue phosphors. A game that uses more than one gun reads both triggers continuously and then, when one player pulls a gun's trigger, the game polls that gun's diode until it knows which object was hit.

Positional guns

Positional guns are fairly common in arcades. A positional gun is a gun mounted to the cabinet on a swivel that allows the player to aim the gun. These are often confused with light guns but work quite differently. These guns may not be removed from the cabinet like the optical counterparts, which are tethered and stored in a mounted holster. They are typically more expensive initially but easier to maintain and repair. Games that use positional guns include the arcade version of *Resident Evil: Code Veronica*, *Space Gun* and *Terminator 2: Judgment Day*. The console ports used light guns. A positional gun is effectively an analog stick that records the position of the gun to determine where the player is aiming. The gun must be calibrated, which usually happens after powering up. Some games have mounted optical guns, such as Exidy's *Crossbow*.

4.3.6. RTS, and Programmable PC Controllers

There are a few specialized controllers that are specifically used for Real Time Shooting games and some arcade type games. These controllers can be programmed to allow the emulation of keys, and macros in some cases. They were developed because some of these games require a keyboard to play, and some players find this to be awkward for such a task. See also:

- Microsoft Sidewinder Strategic Commander
- Zboard Fang
- Ergodex DX1
- Cyber Snipa Game Pad
- Zboard Gaming Keyboard
- WolfKing
- Saitek Pro Gamer Control Pad



4.3.7. Others

At a smaller scale, other hardware such as train controls (available after *Microsoft Train Simulator* was released), pinball controllers and multi-button consoles for strategy games were released in the past, but their popularity was limited to hardcore fans of the genre. Dance pads, essentially a grid of flat pressure sensitive gamepad buttons set on a mat meant to be stepped on, have seen niche success with the popularity of rhythm games like *Dance Dance Revolution*. The dance pad was first introduced on the Atari 2600, called the "Exus Foot Craz" pad. Nintendo later purchased the technology from Bandai and used it on their "Power Pad", for the Famicom and NES. In the early 1990s, the Super Nintendo Entertainment System video game system became the first commercial gaming console to feature a mouse in addition to its controllers. The best-known game to have used the mouse's capabilities was *Mario Paint*. Rhythm games with controllers resembling musical instruments like guitars (*Guitar Hero*), drums (*Donkey Konga*), or maracas (*Samba de Amigo*) have also seen some success in arcades and home consoles. A recent example of specialized, while very simple, game controllers, is the four large "buzzers" (round buttons) supplied with the PlayStation 2 quiz show game series *Buzz!* (2005, 2006); both game and controllers clearly being inspired by the television show genre. The upcoming Wii utilizes a new kind of controller, called the Wii Remote. It has motion sensors and can detect its exact location and orientation in 3D space.



4.3.8. Analog stick

An **analog stick**, sometimes called **thumbstick**, often mistakenly referred to as a joystick, is an input device for a controller (often a game controller) that is used for two-dimensional input. It consists of some sort of protrusion from the controller, and the input is based on the position of this protrusion. While a digital joystick relies on single electrical connections for movement (using internal digital electrical contacts for up, down, left and right), an analog stick uses continuous electrical activity running through potentiometers. The analog stick has greatly overtaken the D-pad in both prominence and usage in console video games.

Usage in video games

In a video game, an analog stick is often used to move some game object, usually the protagonist. It may also be used to rotate the camera, usually around the protagonist. The analog stick can serve a great variety of other functions, depending on the game. Today many analog sticks can also be pushed like other buttons on a controller. Its use is prevalent in 3D games, where more than 8 directions are needed (which is all that was offered on the d-pad). Using a D-pad in a 3D game greatly limits the ability to move. For example, in a 3D platformer like *Mario 64*, using a D-pad would allow Mario to move forward, left, and forward-left, but nothing in between. Some early 3D games like *Resident Evil* overcame this limitation by assigning the Left and Right directions on the D-pad to spin the character instead of making the character move in that direction. However with the prevalence of analog sticks, the aforementioned limitation of the D-pad no longer became an issue.

Dual Analog Sticks

Following the adage "two of something are better than one", two analog sticks offer greater functionality than a single stick. On modern game controllers, the position of the first stick is where the left thumb usually rests, usually to the upper left of D-pad. A

notable exception of this is Sony's iterations of the DualShock controller, where the D-pad and left analog stick are reversed, probably for reasons of symmetry. See the picture to the right. The position of the right stick is usually to the lower left of the face buttons. Although this is a slightly awkward position for the right thumb, it is understandable seeing as the buttons that normally appear on the face on the controller are used more often. With genres such as action, adventure game, platforming, and shooting, the left stick normally controls the protagonist while the second stick controls the camera. The use of a second analog stick alleviated problems in many earlier platform games, in which the camera was notorious for bad positioning. The right stick not only allows for camera control in third-person platform games, but it is almost essential for most modern first-person shooters like Halo. In Katamari Damacy by Namco, both analog sticks are used at once to control your character. The PSP has been criticized for the lack of a second analog "nub" to accompany the left-hand one. The reason behind this may have been the physical instability of the system that would result if the player were to use both analog sticks at the same time.

History

In 1982 Atari released the first controller with an "analog" stick for their Atari 5200 home console. However, the non-centering joystick design proved to be ungainly and unreliable, alienating many consumers at the time. During that same year, General Consumer Electronics introduced the Vectrex, a vector graphics based system which used a self-centering analog stick, obviously a precursor to the modern design. For many years, consoles ignored analog technology, instead using the digital D-pad. It wasn't until the emergence of 3D gameplay that the analog stick was brought back for widespread use. In 1995, Sony (realizing that analog technology could be useful) created a potentiometer-based analog joystick for use in Flight-Simulation games. The Sony Dual Analog FlightStick featured twin analog sticks and was used in games such as "Descent" to provide a much greater degree of freedom than the typical digital joysticks of the day. In 1996 Nintendo introduced a modern analog stick on their Nintendo 64 controller. Unlike the D-pad, the analog stick allowed for varying levels of pressure and 360-degree control, translating into more precise movements in games such as *Super Mario 64*. Incidentally, *Super Mario 64 DS* was criticized by some for imprecise control due to lack of an analog stick on the Nintendo DS. *Super Mario 64* was released in Japan on 23 June 1996.

On 5 July 1996, Sega released *NiGHTS Into Dreams* for their Saturn console in Japan; bundled with it was the Saturn 3D control pad which featured an analog pad intended to give the player more fluid control over that game's flight-based gameplay. During this same period of time, Sony was also creating a similar analog stick, based off the same potentiometer technology that was used in the larger Dual Analog Flightstick. This controller version, released in November 1996, was the Sony "Dual Analog" controller. The Dual Analog featured three modes of analog (Flightstick, Full Analog and Analog-Off), and dual plastic concave thumbsticks, while Nintendo's controller only had a single stick. In 1997 Sony released their third analog revision to the market- the Sony "Dual Shock". The controller featured similar twin analog sticks to the Dual Analog, although they were convex and rubber tipped. Sony also removed the third analog (flightstick) mode and replaced their own previous rumble feature with Immersion's haptic technology. (Note: In 1999, *Ape Escape* became the first major video game to **require** the use of two analog sticks.) Since then, all major video game console controllers have included two analog sticks, with the exception of the Sega Dreamcast and Nintendo's Wii controllers. On 16 September 2005, Nintendo showed the controller for their upcoming system, Wii (then code-named Revolution), which does not have any analog sticks on its main body. Instead, it features a gyroscopic sensor for three-dimensional detection, and a single analog stick as part of a "nunchaku-style" expansion.

4.3.9. DualShock

The DualShock is the standard game controller available for the PlayStation and PlayStation 2 video game consoles. The DualShock was introduced in Japan in late 1997, and launched in America on May 1998, meeting with critical success. First introduced as a secondary peripheral for the original PlayStation, a revised PlayStation version came with the controller and subsequently phased out the digital controller that was originally included with the hardware, as well as the Sony Dual Analog Controller. The DualShock was a success, providing gamers with a new way to get further into their games.

Design

The DualShock analog controller is a controller capable of providing feedback based on the onscreen action of the game (if the game supports it), or Vibration Function. The controller is called Dual Shock because the controller has two vibration motors: a weak buzzing motor (like cell phone or pager vibration), and a strong rumble motor, similar to that of the Nintendo 64's Rumble Pak. Unlike the Rumble Pak, however, the Dual Shock uses the PlayStation to power the vibration function, whereas the Rumble Pak uses batteries. Some third party Dual Shock controllers use batteries in lieu of the PSone's power supply. This rumble feature is similar to the one featured on the first edition of the Japanese Dual Analog, a feature that was removed shortly after that controller was released. The controller is widely supported; shortly after its launch, most new titles including *Crash Bandicoot: Warped*, *Spyro the Dragon*, and *Tekken 3* included support for the vibration function and/or analog sticks. Many games took advantage of the presence of two motors to provide vibration effects in stereo; for example, *Gran Turismo* or the PlayStation port of *Quake II*. Like its predecessor, the Dual Analog Controller, the DualShock has two analog sticks, and unlike said controller, the sticks featured rubber tips, in lieu of the grooves recessed into the Dual Analog Controller's sticks. However, the handles of the Dual Shock are a bit smaller than that of the Dual Analog Controller. The DualShock is one of the most ergonomic designs of a game controller, and, as a result, has been cloned by other companies; this has included some applications very different from controlling a videogame.

Vibration function suit

In 2002, Sony and Microsoft were sued by force-feedback company Immersion for patent infringement for the use of vibration functions in their controllers. Specifically, they were accused infringing on claims in US patents numbers 6,424,333 and 6,275,213 (both "Tactile feedback man-machine interface device", filed 2000 and 2001). While Microsoft settled out of court, purchasing a 10% share in the company, Sony continued to defend the case. Their defence centred on an earlier force feedback controller patent, the use of which Sony licenced from Logitech during 1998. Sony lost and was required to pay considerable retroactive royalties (approximately \$91 million) to Immersion, for the seven years that the Dual Shock controller had been on the market, and suspend the sale of the controllers in the United States (including all PlayStation and PlayStation 2 console packages containing them). Sony appealed this decision and lost again.



☞ Dualshock controller compared to gun controller of the British Challenger 2 main battle tank

DualShock 2

When the PlayStation 2 computer entertainment system was announced, the controller included with it was the same as the Dual Shock analog controller, except it was black, added two additional vibration levels, was considerably lighter and all of the buttons (except for the SELECT, START, ANALOG mode, L3 and R3 buttons) were pressure sensitive. Only a few games do support the pressure sensitive buttons though, such as Square Enix's *The Bouncer*, SCEA's *Gran Turismo 4*, Fresh Games' *Mad Maestro*, various racing games, and the PS2 *Metal Gear Solid* and *Grand Theft Auto* installments.

Playstation 3

The prototype boomerang-shaped controller was shown for the first time in conjunction with the PlayStation 3 at E³ in 2005. However, it was confirmed by Sony in March of 2006 that the "boomerang" will now be replaced with a new controller design. The new design is almost identical to DualShock 2, with several additional features. In Sony's E3 conference on May 8th, 2006, it was revealed that the new, redesigned, wireless controller has a tilt sensor that allows six degrees of freedom (movement in all three spatial axes, plus pitch, roll, and yaw rotation). It also has broader analog sticks, a USB cable port at the top of the controller (utilized for connections to the PS3 and recharging), four LEDs presumably for play identification, and a "guide" button on the center of the controller's face. The controller is lighter than previous DualShock controllers (likely due to the loss of vibration function, see above). The controller's gyroscopic functions resemble those of the Wii's nunchuck controller announced almost 8 month prior. The controller will also drop the 'vibration' feature as it will interfere with the motion sensors. (However, this was proven false as the Wii controller retains its vibration function. It's speculated that vibration features are being removed because of a lost lawsuit involving vibration functions of previous PlayStation controller designs, see above). It is not clear if this controller is considered the next in the DualShock line or if Sony plans on giving it another name. It is possible that Sony will keep the name merely for its brand name appeal, but possible new titles, such as "DualShake", have been rumored. [1]



4.3.10. Power Glove

The **Power Glove** is a controller accessory for the Nintendo Entertainment System designed by Abrams/Gentile Entertainment, made by Mattel in the United States and PAX in Japan. The peripheral originally retailed for \$100 (USD). The Power Glove was unique in the fact that it was a glove instead of the normal control pad. The glove had a traditional D-pad on the forearm as well as a program button and buttons labeled 0-9. A person would hit the program button and a numbered button to do various things (such as increase or decrease the firing rate of the A and B buttons). Along with the controller, a gamer could move his or her hand in various movements to control a character on-screen. It was based on the patented technology of the VPL Dataglove, but with many modifications that allowed it to be used with a slow hardware and sold at an affordable price. Where the Dataglove could detect yaw, pitch and roll, used fiber optic sensors to detect finger flexure and had a resolution of 256 positions (8 bits) per 5 fingers, the Power Glove could only detect roll, and used sensors coated with



conductive ink yielding a resolution of 4 positions (2 bits) per 4 fingers. This allowed the Power Glove to store all the finger flexure information in a single byte. Only two games were released with specific features for use with the Power Glove, *Super Glove Ball*, a 3D Breakout clone, and *Bad Street Brawler*, a difficult to control *Double Dragon* clone, playable with the standard NES controller, but allowing exclusive moves with the glove. Two more games, *Glove Pilot* and *Manipulator Glove Adventure*, were announced but never released. Super Glove Ball was never released in Japan. Since no games ever retailed in Japan, the Power Glove was sold only as an alternative controller. This decision hurt sales and eventually caused PAX to declare bankruptcy.

The glove was somewhat difficult to use; in order for the glove to register with the ultrasonic sensors, one had to keep their knuckles pointing at them at all times. In addition, since each finger had 4 positions, it is difficult to have disparate fingers in different positions simultaneously. Thus, while it was possible to control 16 different states simultaneously, most people have difficulty keeping the ring and index fingers straight, while bending the middle finger. While the Power Glove was never a popular input device for the NES, it later gained tremendous popularity among virtual reality enthusiasts as a wired glove. This was mostly due to its low cost and high availability. In 1996, Abrams/Gentile Entertainment was expected to release a new and better Power Glove, designed specifically for the PC, but the product never materialized. Later, they designed a product codenamed Gauntlet that went on to become Essential Reality's P5 Glove.

4.3.11. Wired glove

A **wired glove** is a glove-like input device for virtual reality environments. Various sensor technologies are used to capture physical data such as bending of fingers. Often a motion tracker, such as a magnetic tracking device or inertial tracking device, is attached to capture the global position / rotation data of the glove. These movements are then interpreted by the software that accompanies the glove, so any one movement can mean any number of things. Gestures can then be categorized into useful information, such as to recognize American Sign Language or other symbolic functions. Expensive high-end wired gloves can also provide haptic feedback, which is a simulation of the sense of touch. This allows a wired glove to also be used as an output device. Traditionally, wired gloves have only been available at a huge cost, with the finger bend sensors and the tracking device having to be bought separately. One of the first wired gloves available to home users was the Nintendo Power Glove. This was designed as a gaming glove for the Nintendo Entertainment System. It had a crude tracker and finger bend sensors, plus buttons on the back. In 2001, Essential Reality made a similar attempt at a cheap gaming glove, this time for the PC: the P5 Glove. This peripheral never really became popular among gamers, so P5 gloves might be now found on Ebay for as little as \$10. Ironically, even specialized stores are now selling the older and less performing Power Glove for a higher price than the more sophisticated P5 Glove. Wired gloves are often called "datagloves" or "cybergloves", but these two terms are trademarks. They respectively belong to Sun Microsystems (which acquired the patent portfolio of VPL Research Inc. in February of 1998) and Immersion.

4.3.12. Longevity of hardware

Given the number of mobile and soft rubber parts in controllers it can be expected that after extensive use, some of the buttons will become eventually less responsive due to the softening of the rubber parts that connect the hard exterior button to the integrated circuit. Evidence of this is clearly visible in the video game kiosks in stores. Even the plastic outer casings of joysticks and wheels might crack if used too violently. This becomes more of an issue with cheaper third-party controllers. *Button mashing* and *joystick wobbling* were responsible for countless broken controllers until the mid of the 16-bit era, when such games became progressively out of fashion. Even better built joypads, able to endure mechanical wear for years, can be defeated by the development of games which require more buttons or functions, or changes in the interfaces used, rendering them obsolete. For example, the increasing number of axes and buttons demanded by computer flight simulator titles and the

disuse of the PC gameport interface have left many a working PC controller unusable. The end of a game console generation generally brings obsolescence for both a console and its controllers.

4.3.13. Health concerns

Since the controller is the most common way of interacting with a game, it has to be ergonomically designed to feel comfortable to most of their potential userbase to avoid injuries such as the ones in the RSI group or CTS. Most controllers these days are designed with the relaxed position of the hands in mind, which gave origin to the "horns" design that reduced the soreness and cramping after extended use with older pads such as the NES or the Mega Drive/Genesis. However, it's nearly impossible to find a perfect solution, since the age of the regular console player can go from the late childhood or early teens to the late thirties or above. The Xbox pad, for instance, was deemed too large for most players when it was released, but actually some users with longer fingers claim to feel more comfortable using it than the smaller Type-S model. Also, *Nintendo fingers* was a term coined in the early 90s after video game players had their thumbs badly burnt and even developed blisters due to the hardness of the buttons. A more common, less dangerous phenomenon sometimes known as a *Nintendo cramp* means the strain felt on the thumbs and the palm below. This tends to occur mostly when using particularly small controllers. The latter may in fact occur on any finger or the hand if it is kept in a prepared "arched" position for a long time. Analog sticks are normally intended to be used with the players' thumbs. However, many people who played Nintendo's original *Mario Party*, in which many of the minigames required the player to rotate the analog stick as rapidly as possible, instead used their palm, which often resulted in blistering or other injury. As a result, Nintendo agreed to provide special gloves to owners of the game. Future Mario Party titles would no longer include minigames of this type.

4.4. MIDI: Musical Instrument Digital Interface

Musical Instrument Digital Interface, or MIDI, is an industry-standard electronic communications protocol that defines each musical note in an electronic musical instrument such as a synthesizer, precisely and concisely, allowing electronic musical instruments and computers to exchange data, or "talk", with each other. MIDI does not transmit audio - it simply transmits digital information about a music performance. The MIDI Show Control (MSC) protocol (in the Real Time System Exclusive subset) is an industry standard ratified by the MIDI Manufacturers Association in 1991 which allows all types of media control devices to talk with each other and with computers to perform show control functions in live and canned entertainment applications. Just like musical MIDI (above), MSC does not transmit the actual show media – it simply transmits digital information about a multimedia performance.

Almost all music recordings today utilize MIDI as a key enabling technology for recording music. In addition, [MIDI] is also used to control hardware including recording devices as well as live performance equipment such as stage lights and effects pedals. The MIDI standard was first proposed by Dave Smith in 1981 in a paper to the Audio Engineering Society and the MIDI Specification 1.0 was published in August 1983. MIDI allows computers, synthesizers, sound cards, samplers and drum machines to control one another, and to exchange system information. Though modern computer sound cards are MIDI-compatible and capable of creating realistic instrument sounds, the fact that sound cards' MIDI synthesizers have historically produced sounds of dubious quality has tarnished the image of a computer as a MIDI instrument. In fact, the MIDI specification itself has nothing to do with the quality of sound produced - this varies depending on the quality of sound card and/or samples used.

MIDI is almost directly responsible for bringing an end to the "wall of synthesizers" phenomenon in 1970s-80s rock music concerts, when keyboard instrument performers were sometimes hidden behind banks of various instruments. Following the advent of MIDI, many synthesizers were released in rack-mount versions, enabling performers to control multiple instruments from a single keyboard. Another important effect of MIDI has been the development of hardware and computer-based sequencers, which can be used to record, edit and play back performances. Synchronization of MIDI sequences is made possible by the use of MIDI timecode, an implementation of the SMPTE time code standard using MIDI messages, and MIDI timecode has become the standard for digital music synchronization. A number of music file formats have been based on the MIDI bytestream. These formats are very compact; often a file of only 10 kilobytes can produce a full minute of music. This is advantageous for applications such as musical ringtones in mobile phones, and some video games.

How musical MIDI works

When a note is played on a MIDI-aware instrument, it transmits MIDI messages. A typical MIDI message sequence corresponding to a key being struck and released on a keyboard is:

1. the user presses the middle C key with a specific velocity or pressure (which is usually translated into the volume of the note but can also be used by the synthesiser to set the timbre as well)
2. the user changed the force with which he is holding the key down (can be repeated, optional)
3. the user stopped playing the middle C note, again with the possibility of velocity of release controlling some parameters



Other performance parameters would also be transmitted. For example, if the pitch wheel were being turned, that information would also be transmitted using different MIDI messages. The musical instrument does this completely autonomously requiring only that the musician play a note (or do something else that generates MIDI messages). All notes that a musical instrument is capable of playing are assigned specific MIDI messages according to what the note and octave are. For example, the Middle C note played on any MIDI compatible musical instrument will always transmit the same MIDI message from its 'MIDI Out' port. Which MIDI message and thus which binary digits will be transmitted upon playing of a certain note are defined in the MIDI specification and this comprises the core of the MIDI standard.

All MIDI compatible instruments follow the MIDI specification and thus transmit identical MIDI messages for identical MIDI events such as the playing of a certain note on the musical instrument. Since they follow a published standard, all MIDI instruments can communicate with and understand each other, as well as with computers which have been programmed to understand MIDI messages using MIDI-aware software. The MIDI interface converts the current fluctuations transmitted by a MIDI musical instrument into binary numbers that the receiving musical instrument or computer can process. All MIDI compatible instruments have a built-in MIDI interface. In addition, computer sound cards usually have a built-in MIDI interface - if not, it can be separately purchased as a card and easily installed.

The MIDI specification: Electrical connections

The MIDI standard consists of a communications messaging protocol designed for use with musical instruments, as well as a physical interface standard. Physically it consists of a one-way (simplex) digital current loop serial communications electrical connection signaling at 31,250 bits per second.

Only one end of the loop is referenced to ground, with the other end 'floating', to prevent ground loops from producing analog audio interference and hum. The current loop on the transmitter side drives the LED of an opto-coupler on the receiver side. This means the devices are in fact optically isolated. The opto-coupler must be a high-speed type. As most opto-couplers have asymmetrical positive-going and negative-going slew rates, they slightly alter the signal's duty cycle. If several MIDI devices are connected in series by daisy-chaining the MIDI THRU to the next device's MIDI-IN, the signal gets more and more distorted (until receive errors happen because the positive or negative pulses get too narrow).



MIDI connectors use standard 5-pin DIN connectors which at one time were a de facto European standard for audio interconnection. Over time the simpler American RCA phono jack has replaced DIN in this application, leaving MIDI as the only place DIN is commonly encountered in modern equipment. Standard size DIN connectors were also used for computer keyboard connections from the early 80s through the late 90s and have generally been replaced by mini-DIN connectors. Computers have 15-pin D-Sub connectors that are used for MIDI IN/OUT or joystick connection. The recommended method of connecting two 5-pin DIN cables to 15-pin D-Sub computer port can be found at MIDI.org web site. The MIDI specification very conservatively states that the maximum distance MIDI can be transmitted is 50 feet or 15m but it can normally go much further. There is a USB connection standard and a standard for MIDI over ethernet and internet called RTP MIDI being developed by the IETF. See external links below for further information.

Most MIDI capable instruments feature a MIDI IN, MIDI OUT, and occasionally a MIDI THRU connection in the form of five-pin DIN plugs. In order to build a two-way physical connection between two devices, a pair of cables must be used. The MIDI-THRU jack simply echoes the signal entering the device at MIDI-IN. This makes it possible to control several devices from a single source. The 1985 Atari ST was the first home computer to sport the

original five-pin format – which made the ST a very popular platform for running MIDI sequencer software. Most PC soundcards from the late 1990s had the ability to terminate a MIDI connection (usually through a MIDI IN/MIDI OUT converter on the game port). The game port has been supplanted in the modern PC by USB devices, and so typically a PC owner will need to purchase a MIDI interface that attaches to the USB or FireWire port of their machine to use MIDI.

Message format

Each one-way connection (called a port) can transmit or receive standard musical messages, such as note-on, note-off, controllers (which include volume, pedal, modulation signals, etc.), pitch-bend, program change, aftertouch, channel pressure, and system-related messages. These signals are sent along with one of 16 channel identifiers. The channels are used to separate "voices" or "instruments", somewhat like tracks in a multi-track mixer. The ability to multiplex 16 "channels" onto a single wire makes it possible to control several instruments at once using a single MIDI connection. When a MIDI instrument is capable of producing several independent sounds simultaneously (a multitimbral instrument), MIDI channels are used to address these sections independently. (This should not be confused with "polyphonic"; the ability to play several notes simultaneously in the same "voice".)

Note messages can represent any note from "C (i.e. five octaves below middle C or 8.175 Hz in common Western musical tuning; designated as MIDI note 0) to g^{''''} (i.e. five octaves above the G above middle C or 12,544 Hz; designated as MIDI note 127) with precision down to the semitone. Pitch-bend messages range in ± 2 semitones (sometimes adjustable with Registered Parameter Numbers), with precision of $1/8192$ semitone. (The human ear cannot hear the difference between adjacent pure tones that differ by less than $1/20$ semitone.) Continuous Controllers are quite versatile; they can usually be controlled by a musician using knobs, sliders, footswitches, or pressure on the instrument. They can be used to change the tone, timbre, or volume of a sound; move motorized faders; and even dim lights. Program change messages are sent to an instrument on a particular channel to instruct it to recall another patch, or program. The MIDI protocol uses 7 bits for this message, supporting only 128 program changes. Many devices which are more modern than the MIDI specification store far more than 128 programs. To overcome the limitation, a bank-switching method has been added to the spec (Each bank of 127 programs can be selected using a controller message, enabling access to 127^2 programs).

Aftertouch messages (also known as poly pressure messages) are sent in some instruments to indicate pressure changes on the note while it is being played. Similarly, channel pressure changes the pressure for the entire instrument, not just one note. The channel pressure messages are more commonly implemented in most synthesizers, while the individual pressure sensors that aftertouch messages require are reserved mainly for expensive, high-end synthesizers. Manufacturer's system exclusive messages (also known as Manufacturer SysEx, Manuf Sysx, etc.) are defined by the manufacturer of the sequencer/synthesizer and can be any length. These messages are commonly used to send non-MIDI data over a MIDI connection, such as a synthesizer instrument sample or settings and a sequencer's memory dump. Because they are defined by the device's manufacturer, they are mainly used for backup purposes and rarely (if ever) useful in another MIDI device. Real Time System Exclusive messages include the significant MIDI Show Control extension which enables all types of entertainment equipment to easily communicate with each other through the process of show control.

System messages contain meta-information about other MIDI messages. A sequencer, for example, often sends MIDI clock messages during playback that correspond to the MIDI timecode, so the device receiving the messages (usually a synthesizer) will be able to keep time. Also, some devices will send Active Sense messages, used only to keep the connection between the sender and the receiver alive after all MIDI communication has ceased. It should be noted that MIDI can be used to provide facilities for playing in nonstandard musical tunings. However, apart from using pitch-bend to control each note, these features are not standardized across all instruments.

General MIDI

In MIDI, instruments (one per channel) are selected by number (0-127), using the Program Change message. However, the basic MIDI 1.0 specification did not specify what instrument sound (piano, tuba, etc.) corresponds to each number. This was intentional, as MIDI originated as a professional music protocol, and in that context it is typical for a performer to assemble a custom palette of instruments appropriate for their particular repertoire, rather than taking a least-common-denominator approach. Eventually interest developed in adapting MIDI as a consumer format, and for computer multimedia applications. In this context, in order for MIDI file content to be portable, the instrument programme numbers used must call up the same instrument sound on every player. General MIDI (GM) was an attempt by the MIDI Manufacturers' Association (MMA) to resolve this problem by standardising an instrument programme number map, so that for example Program Change 1 always results in a piano sound on all GM-compliant players. GM also specified the response to certain other MIDI messages in a more controlled manner than the MIDI 1.0 specification. The GM spec is maintained and published by the MIDI Manufacturers' Association (MMA).

From a musical perspective, GM has a mixed reputation, mainly because of small or large audible differences in corresponding instrument sounds across player implementations, the limited size of the instrument palette (128 instruments), its least-common denominator character, and the inability to add customised instruments to suit the needs of the particular piece. Yet the GM instrument set is still included in most MIDI instruments, and from a standardisation perspective GM has proven durable. Later, companies in Japan's Association of Musical Electronics Industry (sic) (AMEI) developed General MIDI Level 2 (GM2), incorporating aspects of the Yamaha XG and Roland GS formats, extending the instrument palette, specifying more message responses in detail, and defining new messages for custom tuning scales and more. The GM2 specs are maintained and published by the MMA and AMEI. Later still, GM2 became the basis of Scalable Polyphony MIDI (SP-MIDI), a MIDI variant for mobile applications where different players may have different numbers of musical voices. SP-MIDI is a component of the 3GPP mobile phone terminal multimedia architecture, starting from release 5. GM, GM2, and SP-MIDI are also the basis for selecting player-provided instruments in several of the MMA/AMEI XMF file formats (XMF Type 0, Type 1, and Mobile XMF), which allow extending the instrument palette with custom instruments in the Downloadable Sound (DLS) formats, addressing another major GM shortcoming.

Low bandwidth

MIDI messages are extremely compact, due to the low bandwidth of the connection, and the need for near real-time accuracy. Most messages consist of a status byte (channel number in the low 4 bits, and an opcode in the high 4 bits), followed by one or two data bytes. However, the serial nature of MIDI messages means that long strings of MIDI messages take an appreciable time to send, and many people can hear those delays, especially when dealing with dense musical information or when many channels are particularly active. "Running status" is a convention that allows the status byte to be omitted if it would be the same as that of the previous message, helping to mitigate bandwidth issues somewhat.

MIDI file formats

MIDI messages (along with timing information) can be collected and stored in a computer file system, in what is commonly called a MIDI file, or more formally, a Standard MIDI File (SMF). The SMF specification was developed by, and is maintained by, the MIDI Manufacturers Association (MMA). MIDI files are typically created using desktop/laptop computer-based sequencing software (or sometimes a hardware-based MIDI instrument or workstation) that organizes MIDI messages into one or more parallel "tracks" for independent recording and editing. In most but not all sequencers, each track is assigned to a specific MIDI channel and/or a specific General MIDI instrument patch. Although most current MIDI sequencer software uses proprietary "session file" formats rather than SMF, almost all sequencers provide export or "Save As..." support for the SMF format. An SMF

consists of one header chunk and one or more track chunks. There are three SMF formats; the format is encoded in the file header. Format 0 contains a single track and represents a single song performance. Format 1 may contain any number of tracks, enabling preservation of the sequencer track structure, and also represents a single song performance. Format 2 may have any number of tracks, each representing a separate song performance. Format 2 is not commonly supported by sequencers.

Large collections of SMFs can be found on the web, most commonly with the extension .mid. These files are most frequently authored with the assumption that they will be played on General MIDI players, but not always. Occasional unintended bad-sounding playback is the result. MIDI-Karaoke (which uses the ".kar" file extension) files are an "unofficial" extension of MIDI files, used to add synchronized lyrics to standard MIDI files. Most SMF players do not display these lyrics, however numerous .kar-specific players are available. These often display the lyrics synchronized with the music in "follow-the-bouncing-ball" fashion, essentially turning any PC into a Karaoke machine. Note: ".kar" files can often be played by SMF players if the filename extension is changed to ".mid". The MIDI Manufacturers' Association has now defined a new family of file formats, XMF (eXtensible Music File), some of which package SMF chunks with instrument data in DLS format (Downloadable Sounds, also an MMA specification), to much the same effect as MOD files. The XMF container is a binary format (not XML-based).

Extensions of the MIDI standard

USB, FireWire and ethernet embeddings of MIDI are now commonly available, and in the long run the proposed MIDI over ethernet and internet standard called RTP MIDI, being developed by the IETF, is likely to replace the old current loop implementation of MIDI, as well as providing the high-bandwidth channel that ZIPI was intended to provide. See external links below for further information. In 1992 the MIDI Tuning Standard, or MTS, was ratified by the MIDI Manufacturers' Association. While support for this standard is not great, it is supported by some instruments and software; in particular the free software program Timidity supports it. Use of MTS allows any midi file to be tuned in any way desired, something which can be accomplished using the freeware program Scala (program) and other microtoners. MTS uses three bytes, which can be thought of as a three-digit number base 128, to specify a pitch in logarithmic form. The code for a frequency is derived using the formula:

$$p = 69 + 12 \times \log_2 \left(\frac{f}{440} \right)$$

Remark that for a note in an A440 equal temperament, this formula delivers the standard MIDI note number. Any other frequencies fill the space evenly.

Beyond MIDI

Although traditional MIDI connections work well for most purposes, in 1994 a new high-bandwidth standard, ZIPI, sponsored by UC Berkeley's CNMAT, was proposed to replace MIDI with incompatible but very similar message system. It was not based on channels, instead shifting the control to individual notes on each device. Channel messages were emulated by grouping note messages. The protocol failed to gain more than limited acceptance, primarily due to lack of demand. The Open Sound Control or OSC protocol, developed by the same team, has been implemented in the well-known software synthesizer Reaktor and in other projects such as SuperCollider, Pure Data, Isadora, Max/MSP, Csound, and Chuck. It differs from MIDI in that it can run at broadband speeds over Ethernet connections. The Lemur Input Device, a customizable touch panel with MIDI controller-type functions, also uses OSC but few mainstream musical applications and no standalone instruments support the protocol, making

whole-studio interoperability problematic. OSC is not owned by any private company, however it is also not maintained by any standards organization. Yamaha has its mLAN protocol, which is based on FireWire [1] and carries multiple MIDI message channels and multiple audio channels. mLAN is a proprietary protocol open for licensing. A proposal for High-Definition MIDI (HD-MIDI™) extension is now being discussed by members of the MMA. This major update to MIDI would provide greater resolution in data values, increase the number of MIDI Channels, and support the creation of entirely new kinds of MIDI messages. [2] [3] This work involves representatives from all sizes and types of companies, from the smallest specialty show control operations to the largest musical equipment manufacturers.

Other applications of MIDI

MIDI can also be used for applications other than music:

- show control
- theatre lighting
- special effects
- sound design
- recording system synchronization
- audio processor control
- computer networking, as demonstrated by the early first-person shooter game MIDI Maze, 1987

Any device built with a standard MIDI-OUT port should (in theory) be able to control any other device with a MIDI-IN port, providing that developers of both devices have the same understanding about the semantic meaning of all the transmitted MIDI messages. This agreement can come either because both follow the published MIDI specifications, or else because for non-standard functionality the message meanings are agreed upon by both manufacturers. MIDI controller is used in two senses. In one sense, a controller is hardware or software which generates and transmits MIDI data to MIDI-enabled devices. In the other more technical sense, a MIDI controller is an abstraction of the hardware used to control a performance, but which is not directly related to note-on/note off events. A slider assigned to open and close a low-pass filter on a synthesizer may be assigned to controller 18, for example. Changes in the position of the slider are transmitted along with "18" so that they are distinguished from changes in the value of other controllers.

MIDI controllers which are hardware and software

The following are classes of MIDI controller:

- The human interface component of a traditional instrument redesigned as a MIDI input device. The most common type of device in this class is the keyboard controller. Such a device provides a musical keyboard and perhaps other actuators (pitch bend and modulation wheels, for example) but produces no sound on its own. It is intended only to drive other MIDI devices. Percussion controllers such as the Roland Octapad fall into this class, as do guitar-like controllers such as the SynthAxe and a variety of wind controllers.
- Electronic musical instruments, including synthesizers, samplers, drum machines, and electronic drums, which are used to perform music in real time and are inherently able to transmit a MIDI data stream of the performance.

- Pitch-to-MIDI converters including guitar/synthesizers analyze a pitch and convert it into a MIDI signal. There are several devices which do this for the human voice and for monophonic instruments such as flutes, for example.
- Traditional instruments such as drums, pianos, and accordions which are outfitted with sensors and a computer which accepts input from the sensors and transmits real-time performance information as MIDI data.
- Sequencers, which store and retrieve MIDI data and send the data to MIDI enabled instruments in order to reproduce a performance.
- MIDI Machine Control (MMC) devices such as recording equipment, which transmit messages to aid in the synchronization of MIDI-enabled devices. For example, a recorder may have a feature to index a recording by measure and beat. The sequencer that it controls would stay synchronized with it as the recorder's transport controls are pushed and corresponding MIDI messages transmitted.

MIDI controllers in the data stream

Modifiers such as modulation wheels, pitch bend wheels, sustain pedals, pitch sliders, buttons, knobs, faders, switches, ribbon controllers, etc., alter an instrument's state of operation, and thus can be used to modify sounds or other parameters of music performance in real time via MIDI connections. The 128 virtual MIDI controllers and their electronic messages connect the actual buttons, knobs, wheels, sliders, etc. with their intended actions within the receiving device.

Some controllers, such as pitch bend, are special. Whereas the data range of most continuous controllers (such as volume, for example) consists of 128 steps ranging in value from 0 to 127, pitch bend data may be encoded with over 16,000 data steps. This produces the illusion of a continuously sliding pitch, as in a violin's portamento, rather than a series of zippered steps such as a guitarist sliding his finger up the frets of his guitar's neck. Thus, the pitch wheel on a MIDI keyboard may generate large amounts of data which can lead to a slowdown of data throughput. Many sequencers can "thin" pitch-bend or other continuous controller data by keeping only a set number of messages per second or keeping only messages that change the controller by at least a certain amount.

The original MIDI spec included 128 virtual controller numbers for real time modifications to live instruments or their audio. MIDI Show Control (MSC) and MIDI Machine Control (MMC) are two separate extensions of the original MIDI spec, expanding the MIDI protocol to become far more than its original intent.

4.4.0. Computer Music

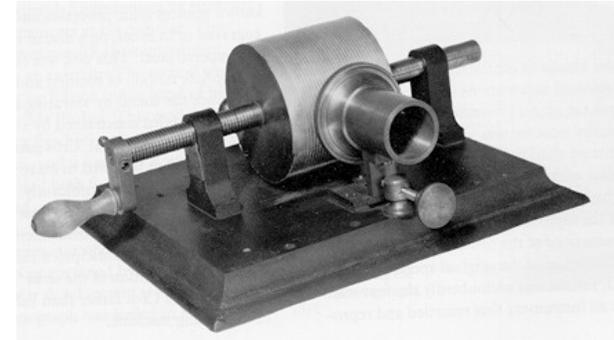
Con il termine Computer Music si intende sempre più spesso non la composizione realizzata attraverso algoritmi matematici del computer, ma l'utilizzo del Personal Computer come mezzo col quale arrangiatori, compositori, musicisti professionisti e utenti amatoriali realizzano le proprie idee, siano queste prettamente tecnologiche che partiture destinate poi a gruppi acustici o ad orchestre. Determinanti per la diffusione della computer music sono stati la diffusione di massa dei PC e l'introduzione del MIDI, linguaggio col quale è possibile pilotare la maggior parte degli strumenti musicali elettronici in commercio. Inizialmente ad utilizzo esclusivo di tastieristi - pianisti ha successivamente sviluppato nuove interfacce sia "midizzando" strumenti come chitarra, sax, basso elettrico sia sviluppando software con interfacce più semplici e immediate permettendone l'utilizzo sia agli altri musicisti sia ad i programmatori tekno(logici) che spesso utilizzano mouse e tastiera per le loro composizioni. La computer music è un aspetto dell'informatica musicale, che rappresenta la sintesi di diverse aree, dalla ricerca e produzione artistica (composizione, interpretazione, direzione d'orchestra), alla ricerca in campo scientifico (modelli e sistemi informatici per lo studio di segnali musicali) e umanistico (teoria musicale, musicologia), all'industria dello spettacolo e multimediale (produzione di strumentazione musicale elettronica, editoria musicale, realtà virtuale, produzione e post-produzione video, case discografiche, studi di registrazione, ecc.) Dal punto di vista scientifico, l'informatica musicale delinea un campo di investigazione complesso: la musica costituisce un dominio in cui verificare e sviluppare proficuamente sia metodologie che applicazioni propriamente informatiche. La ricerca in informatica musicale è inoltre motivata da ragioni artistiche (sviluppo ed esplorazione di nuovi approcci alla composizione, alla performance, ecc.), scientifiche e tecnologiche (qualità sonora più elevata, strumenti musicali di potenza sempre maggiore, migliore comprensione della natura dell'intelligenza musicale, ecc.), e da un elevato potenziale economico in diverse direzioni: nuovi strumenti (musicali e non) per la produzione e la post-produzione di suono e musica, in particolare per i settori radiotelevisivo, cinematografico, multimediale e di intrattenimento (dispositivi ottici per la memorizzazione di suoni ed immagini, realtà virtuale, interazione uomo-macchina, arte interattiva, teatro, danza). Le applicazioni musicali degli elaboratori costituiscono un'area informatica oggetto di ricerca da più di venticinque anni. Numerosi centri di ricerca operano in questo settore sia in ambito accademico che industriale, soprattutto nel Nord America (il Media Lab del MIT a Boston, il CARL-UCSD dell'Università della California, le università di Stanford e Carnegie Mellon) e in Europa (l'IRCAM a Parigi). Da tempo in Italia è in corso una notevole attività nel settore dell'informatica musicale da parte di numerosi centri, il cui punto di riferimento è l'Associazione di Informatica Musicale Italiana. Un termine spesso confuso con la computer music è musica elettronica, col quale si intende musica non solo prodotta attraverso apparecchiature elettroniche (non necessariamente computer) ma musica nella quale l'elettronica risulti la chiave creativa.

Cenni storici

Nell'ultimo secolo è cambiato radicalmente il modo di comporre, eseguire, riprodurre, ascoltare, nonché quello di concepire la musica. Nelle pagine che seguono sono riassunte brevemente e in modo molto schematico le tappe storiche più importanti che hanno contribuito a questo cambiamento.

1877. Viene inventata la fono-fissazione, più comunemente chiamata registrazione, per opera di Charles Cros. Thomas Edison in seguito realizzò il primo vero fonografo, che comprendeva un ricevitore, un registratore e un riproduttore: il ricevitore era una specie di cornetto acustico rovesciato, i movimenti del diaframma del ricevitore venivano trasmessi a un ago d'avorio fissato al centro del diaframma. Il registratore era costituito da un cilindro ricoperto da uno strato di cera davanti al quale si muoveva con velocità uniforme un carrellino portante il ricevitore e lo stilo; in tal modo la punta d'avorio dello stilo, a contatto con la superficie del cilindro, tracciava un solco in profondità corrispondente all'impulso fonico. Il riproduttore era costituito da un tronco di cono metallico cavo la cui base minore era chiusa da una sottile lamina vibrante (diaframma). Al centro del diaframma era fissato un ago la cui punta seguiva esattamente il solco del cilindro del registratore e trasmetteva al diaframma stesso, sotto forma di vibrazioni, i segni incisi riproducendo così i suoni registrati. Il sistema era puramente meccanico, e quindi non si può ancora parlare di

Musica Elettronica, ma la scoperta di Cros Charles ha permesso di poter "fissare", al pari della fotografia con le immagini qualsiasi suono, e ha sconvolto così sia il concetto spaziale della musica (ascoltare il suono della musica svincolato dall'esecutore) sia quello temporale (ascoltare musica in un altro momento da quello nel quale è stata eseguita). Si pensi solo a un piccolo esempio: la maggior parte delle persone utilizza la musica come "sottofondo", l'ascolta cioè facendo altre cose: al lavoro, mentre legge, in macchina alla guida, a cena, durante la meditazione, mentre fa ginnastica. Questo modo "distratto" di usufruire della musica, che diventa così un tappeto sonoro indipendente e distaccato dagli esecutori, prima dell'invenzione del fonografo era non solo irrealizzabile, ma inconcepibile.



1902. Thaddeus Cahill progetta il Dynamophone, costituito da un gruppo di generatori (dinamo, ciascuno del peso di diversi quintali) che producevano correnti alternate a diverse frequenze. L'idea era quella di diffondere musica via cavo. La costruzione, giunta ancora prima che fossero disponibili i mezzi per realizzarla, fu abbandonata.

1920-30. In questo periodo sono stati inventati i primi strumenti musicali elettronici, che producevano quindi segnali elettrici in seguito trasformati in suoni. Tra questi il primo fu il Theremin (1919-1920) inventato da Leon Theremin. La particolarità di questo strumento è l'interfaccia uomo macchina, che risulta tutt'oggi innovativa. Non applicò infatti una tastiera di comando al Theremin, ma due antenne in grado di interpretare i movimenti del musicista e tradurli in dinamica e altezza. Un'antenna verticale interpretava la frequenza della nota, avvicinando o allontanando la distanza della mano (quasi mimando il movimento che fa il suonatore del trombone americano) l'antenna orizzontale interpretava l'ampiezza del suono. Il prototipo funzionava con delle cuffie, solo nel 1925 vi fu applicato un altoparlante consentendo così al Theremin poter suonare nelle orchestre. Altri strumenti inventati in quest'epoca sono le Onde Martenot (1928, che basava la produzione del suono su oscillatori valvolari e aveva un'interfaccia più usuale, quella di una comune tastiera), l'organo hammond (1929).



1930-40. In questi anni vengono messi a punto due invenzioni che avranno grande influenza sulla musica attuale: l'amplificazione sonora e la registrazione su nastro magnetico. L'avvento degli amplificatori permetterà non solo un migliore ascolto della musica registrata, ma porterà ad un nuovo modo di concepire la musica, portando strumenti come la chitarra a diventare strumento solista, ad espressioni vocali nuove (si pensi alla voce sussurrata che può emergere dinamicamente sopra una batteria), e alla diffusione degli strumenti elettronici.

1948. Nasce e si sviluppa attorno alla figura carismatica di Pierre Schaeffer la Musica Concreta. Lavorando presso gli studi della società radiotelevisiva francese (RTF) ebbe la possibilità di utilizzare un vasto archivio discografico e di incominciare a fare esperimenti sul suono e sul rumore. Scrive Schaeffer: "noi abbiamo chiamato la nostra musica concreta poiché essa è costituita da elementi preesistenti" (dunque concreti) "presi in prestito da un qualsiasi materiale sonoro, sia rumore o musica tradizionale." Quindi la musica concreta prende come materiale di base per le proprie composizioni l'oggetto sonoro, un evento sonoro opportunamente isolato attraverso il processo della registrazione e successivamente

manipolato attraverso qualsiasi processo. La musica concreta non trovò diffusione tra il grande pubblico, come d'altronde la musica contemporanea oggi, e venne utilizzata soprattutto per effetti sonori radiofonici (nel 1951 la RTF decise di impiantare uno studio stabile di musica concreta con la finalità di costruire un repertorio utilizzabile per le proprie trasmissioni). E' importante osservare però che con Schaeffer per la prima volta si ha l'utilizzo dei mezzi tecnologici (in questo caso registratori a nastro, consolle di mixaggio) in maniera creativa. Ed è impossibile non notare come questo ingegnere - musicista (che ormai è diventata una figura professionale ormai comune nel mondo musicale) abbia anticipato di più di 40 anni l'avvento dei campionatori e di una serie di generi musicali che, sebbene con fini commerciali, utilizzano come base per le loro composizioni oggetti sonori (=campionamenti) di rumori o musica tradizionale (suoni singoli o interi spezzoni di altre incisioni).

1950-60. Nasce, per opera di Karlheinz Stockhausen e Erbert Eimert, nello Studio di Colonia WDR la musica elettronica. Al contrario dello studio RTF francese, lo studio di Colonia era equipaggiato con strumentazione elettronica in grado di generare elettronicamente i suoni. Da qui la distinzione tra musica concreta (suoni o rumori registrati e in seguito trattati e montati insieme) e musica elettronica (suoni e rumori prodotti da generatori di suono elettronici, in seguito registrati e montati insieme) che poi si andrà sempre più assottigliando. Il primo approccio dei compositori tedeschi con la musica elettronica fu molto rigoroso, lo scopo era quello di superare la concezione dodecafonica di Schomberg attraverso l'eliminazione del sistema tonale e dell'esplorazione di tutto lo spettro sonoro, e di superare difficoltà esecutive estremamente complesse attraverso l'utilizzo delle macchine. Solo successivamente, con il progresso della tecnologia e la progettazione di strumenti elettronici di larga diffusione, la musica elettronica e' entrata a far parte della musica pop perdendo però completamente il suo aspetto sperimentale e di ricerca, e diventando o uno strumento per emulare strumenti acustici o per scopi effettistici. Lo dimostra il fatto che l'unica interfaccia esistente attualmente per questi strumenti è la tastiera, accordata secondo il sistema temperato.

1960-70. Nascono i primi sintetizzatori portatili a basso prezzo: Mini-Moog, e successivamente ARP e Oberheim. La diffusione di questi strumenti porta il loro utilizzo in molti campi: dalla musica leggera alle colonne sonore per film. Vengono inoltre sviluppati in questo periodo i primi sistemi digitali per il controllo dei synth analogici, i primi sequencer.

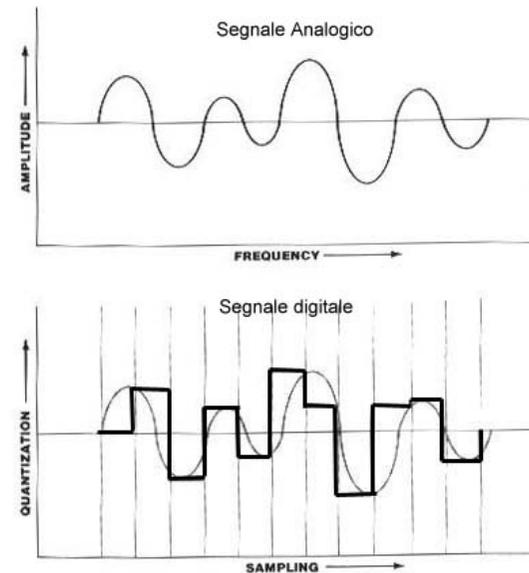
1980-90. Nel 1981 nasce, per opera della Sequential Circuits, il protocollo M.I.D.I., in seguito accettato anche da tutti gli altri costruttori di strumenti musicali elettronici. Il primo sintetizzatore con porte midi fu il Prophet 600, della SC. La computer music si diffonde nello stesso periodo, nel momento in cui iniziava la diffusione dei personal computer. Da lì a poco collegare un PC ad uno strumento midi e attraverso un software adeguato comporre musica sarebbe diventato alla portata di tutti. Nello stesso periodo si ha l'avvento dei sintetizzatori digitali, dei campionatori, e dei primi sistemi HD recording (il software Soundesigner della Digidesign per Mac risale al 1985).

Anni 90. L'utilizzo del computer nella musica diventa popolare. I programmi musicali con gestione del midi integrano l'audio digitale attraverso l'HD recording, permettendo attraverso un singolo software di gestire contemporaneamente e nello stesso modo audio e midi. La progettazione di personal computer sempre più potenti porta all'utilizzo in tempo reale di numerose applicazioni software e alla tendenza a realizzare di un unico programma modulare (composto da più programmi) che possa integrare tutte le potenzialità: effettistica (riverberi, delay, compressori..), sintetizzatori virtuali, processi di mixaggio ed equalizzazione del suono. Spesso questi programmi sono concepiti come plug-in, cioè estensioni di altri programmi e quindi utilizzabili in più contesti. Acquistano larga diffusione i midi file grazie allo standard GM (general midi) che vengono utilizzati non solo da strumenti musicali dedicati ma anche dai Personal Computer. Infine l'avvento di internet e dei nuovi formati audio ad alta compressione (per esempio lo standard MP3 che porta ad una compressione di oltre 1/10 sfruttando algoritmi psicoacustici) potranno portare ad una nuova rivoluzione nel mondo musicale: la distribuzione e la stampa discografica potrebbe essere a breve affiancata e in seguito sostituita dal download (scaricamento dati) diretto dalla rete dei files musicali dell'artista scelto (con la possibilità di scegliere i brani da acquistare e fare delle

vere e proprie "compilation") e dal trasferimento della musica sul supporto scelto (su CD tramite masterizzatore, o direttamente nella memoria dell'MP3 player)

Analogico e Digitale

Una importante distinzione, spesso poco chiara per chi si avvicina per la prima volta all'informatica musicale, è il concetto di segnale analogico e segnale digitale. Prima della diffusione della tecnologia digitale tutti i sistemi elettronici, compresi i calcolatori, utilizzavano la tecnologia analogica, basata cioè dal confronto (analogia) di quantità fisiche che cambiano in maniera CONTINUA. L'elettronica digitale (= numerica) invece utilizza quantità DISCRETE, discontinue. Prendiamo come esempio un'onda sinusoidale, la sua rappresentazione analogica risulterà una curva continua, senza interruzioni, quella digitale sarà invece frammentata, suddivisa in gradini ciascuno del quale con un valore (altezza) definita. Ed è proprio questo il notevole vantaggio del digitale : quello di poter semplificare un segnale analogico in diverse parti, ed assegnare ad ognuna di esse un valore numerico. Il segnale così scomposto, sia esso sonoro che di altro tipo (video, misurazioni di variabili fisiche come la temperatura, il peso, il tempo ecc.) può essere archiviato, modificato, e successivamente ricomposto attraverso il procedimento inverso (conversione digitale-analogica). Il materiale sonoro, una volta trasformato in numeri, può essere modellato attraverso una serie di istruzioni di calcolo (algoritmi). Nei mixer digitali per esempio, i controlli non sono costituiti da componenti elettronici fisici ma da algoritmi. Pensando di progettare un potenziometro digitale del volume dovremmo scrivere un programma che diminuisce o aumenta di un certo valore tutti i numeri .



Il sistema binario e esadecimale

I sistemi digitali sono in grado di riconoscere soltanto due valori : ON/OFF. Un interruttore è un dispositivo digitale in quanto ha soltanto due possibilità, ACCESO SPENTO. Le funzioni che può svolgere un calcolatore sono : aritmetiche, di confronto, di trasferimento e memorizzazione dati. Tramite delle opportune combinazioni è però possibile eseguire qualsiasi calcolo. Lo schema (molto semplificato) di un calcolatore digitale è il seguente : dispositivo in grado di eseguire le operazioni aritmetiche e di confronto (unità centrale di processo CPU), dispositivo in grado di memorizzare i dati (memoria), interfaccia che permetta di inserire e ricevere i dati all'operatore (tastiera, monitor o display). La maggioranza degli strumenti musicali elettronici oggi utilizza la tecnologia digitale (sintetizzatori, campionatori, batterie elettroniche, multieffetti, registratori ecc.) quindi essendo queste basate su microprocessori (CPU) utilizzano la stessa architettura. Poiché computer riconoscono solo due condizioni (0/1) il sistema numerico utilizzato non è quello decimale ma quello binario. Il sistema decimale (a base 10) utilizza 10 caratteri (1,2,3,4,5,6,7,8,9,0) il cui valore cambia a seconda della posizione del carattere stesso (unità, decina centinaia migliaia ecc.) Se scriviamo il numero 3124, questo equivale a alle varie potenze in base dieci moltiplicate per il numero stesso : $3 \times 10^3 + 1 \times 10^2 + 2 \times 10^1 + 4 \times 10^0 = 3124$. Il sistema decimale, da noi comunemente utilizzato (forse per il fatto che la nostra capacità immediata di visualizzazione di conteggio è limitata al numero delle nostre dita) è solo uno dei tanti modi di rappresentare i numeri. Quello binario (appunto usato dai calcolatori) utilizza come base il numero 2. Il numero binario 0101 equivale al numero decimale $2^3 \times 0 + 2^2 \times 1 + 2^1 \times 0 + 2^0 \times 1 = 0+4+0+1 = 5$. Le memorie dei calcolatori sono comunemente strutturate in raggruppamenti di 8 valori binari (bit), l'insieme di 8 bit viene chiamato byte. Da cui poi seguono le altre unità di misura: 1 kbyte=1024 byte, 1 megabyte=1024 kbyte, 1 gigabyte=1024 megabyte). Il sistema binario è poco immediato e per niente pratico, per questo per la rappresentazione dei byte viene utilizzato il sistema esadecimale (a base 16). Il sistema esadecimale utilizza sedici caratteri (0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F). Il principio di conversione è

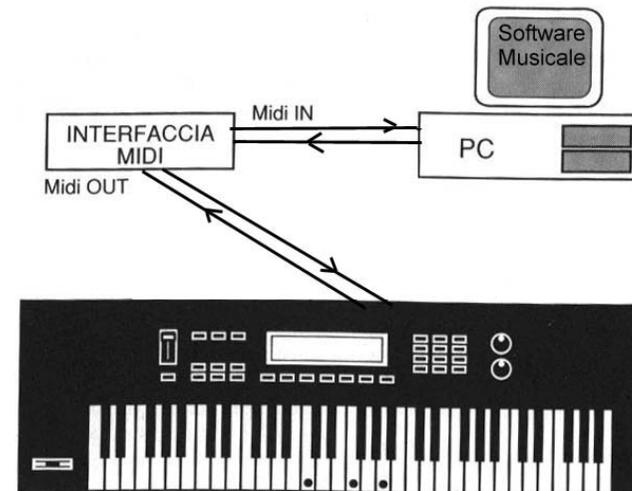
sempre lo stesso, per esempio il numero F7H (H = esadecimale) è uguale a $15 (F) \times 16 + 7 = 247$. Prendiamo un byte e suddividiamolo in due parti : avremo due gruppi di 4 bit ciascuno, ognuno dei quali potrà assumere un valore da 0 a 15 ($1111 = 2^3 + 2^2 + 2^1 + 2^0 = 15$). Possiamo quindi raggruppare un byte con un numero esadecimale composto da due cifre, il cui valore potrà variare da 0-255. Vediamo ora alcuni esempi, il numero binario 0111 1011 = 7Bh = $7 \times 16 + 11 = 123$. IL numero 1010 1101 = ACh = $10 \times 16 + 13 = 173$. L'utilizzo dei numeri binari e esadecimale ci servirà a capire in dettaglio il linguaggio midi, che utilizza codici formati da 8bit, e quello dell'audio digitale che deve sempre tener conto della frequenza di campionamento e dei bit di risoluzione. Spesso inoltre l'implementazione midi degli strumenti elettronici è costituita da una mappa di memoria e da tutti i codici che lo strumento può ricevere, il tutto utilizzando numeri esadecimale, quindi può risultare utile avere domestichezza con questi per programmazioni midi complesse.

Strumentazione di base

Non è facile definire una strumentazione di base per la computer music, perché la costante evoluzione informatica ridefinisce molto spesso nuovi standard. Comunque gli "attrezzi" di base per poter produrre musica attraverso il computer si possono suddividere in 4 gruppi : computer, software, generatore sonoro, interfaccia musicista-computer.

Generatore sonoro: può essere costituito dalla scheda audio del computer, per avere dei risultati sonori accettabili però occorre utilizzare macchine esterne dedicate (o schede audio professionali), che oltre ad un risultato sonoro migliore permettono anche una maggiore versatilità e affidabilità. Esistono attualmente in commercio strumenti che permettono di produrre più di un suono contemporaneamente (polifonia) e più di un timbro contemporaneamente (politimbricità). In questo modo è possibile eseguire arrangiamenti anche complessi con un investimento contenuto. Per poter collegare il modulo al computer occorre un interfaccia midi, un apparecchio in grado di fare comunicare fisicamente (attraverso un collegamento via cavo) le due parti. I computer usano come segnale di uscita per il midi la porta seriale. Esistono diversi tipi di interfacce midi, dalle più semplici (1 IN 1 OUT) fino a quelle più complesse (8 IN 8 OUT).

Interfaccia uomo macchina: quella più diffusa è la tastiera ad imitazione di quella di un organo. Spesso queste tastiere hanno una estensione di 5 ottave e permettono attraverso dei sensori anche il controllo dinamico (velocity). Altre tastiere invece simulano i tasti di un pianoforte, con una meccanica dei asti pesata. Tastiera e modulo sonoro comunemente sono integrati in un unico strumento, quando la tastiera è staccata dal generatore sonoro viene chiamata master keyboard, il modulo sonoro separato da tastiera viene chiamato expander. La tastiera non è l'unico tipo di interfaccia musicista - computer, anche se meno utilizzati esistono sistemi che permettono di trasformare una nota rilevata da un pick-up in un evento MIDI, sensori per pilotare suoni sintetici attraverso batterie o percussioni ecc.



Le porte MIDI

Le porte, come già è stato detto, si dividono in:

- MIDI-IN: è la porta che riceve i dati trasmessi da una apparecchiatura musicale.
- MIDI-OUT: è la porta che serve a trasmettere i dati ad altre apparecchiature musicali.
- MIDI-THRU: è una porta che bypassa il segnale ricevuto alla porta MIDI-IN.

Il segnale così, inalterato, può essere trasmesso ad un'altra apparecchiatura musicale.

I cavi per i collegamenti MIDI

I cavi utilizzati per la trasmissione di dati midi sono composti da cinque fili, anche se in realtà ne vengono utilizzati solo tre. Il connettore per il collegamento alle porte MIDI-IN, OUT, THRU è a cinque poli. Due poli e due cavi non vengono praticamente utilizzati.

I messaggi MIDI

Che significa protocollo MIDI ?

Con l'interfaccia MIDI è possibile collegare insieme tastiere, expanders, computer, pedaliera MIDI, e tutto ciò che abbia un'interfaccia MIDI. L'insieme di tutti i messaggi che le macchine si possono scambiare costituiscono il protocollo MIDI. Il protocollo MIDI dunque stabilisce le modalità con le quali devono essere inviati i messaggi tra le varie porte MIDI di un dispositivo.

Come è composto un messaggio MIDI ?

Un messaggio MIDI è composto da un insieme di byte. Questi possono essere status byte e data byte. Un messaggio MIDI deve sempre iniziare con uno status byte.

- Gli status byte servono per definire in modo univoco un comando. Il bit più significativo dello status byte è uguale ad 1 e quindi uno status byte può assumere un valore compreso tra 128 e 255. Gli status byte trasmettono quindi il tipo di informazione (suona una nota, alza il volume, ecc).
- I data byte invece servono per inviare gli eventuali parametri necessari per un corretto funzionamento dello status byte. Il bit più significativo del data byte è uguale a 0 e quindi un data byte può assumere un valore decimale compreso tra 0 e 127.

Canali MIDI

I canali MIDI possono essere paragonati ai canali televisivi: Es. se vedo in tv canale 6, ricevo i programmi che canale 6 diffonde via etere, e li vedono tutti quelli che sono collegati e sintonizzati sullo stesso canale. Ad esempio chi è sintonizzato sul canale 1 non riceverà le informazioni che il canale 6 trasmette e viceversa. Allo stesso modo funziona per i canali MIDI dove le informazioni passano attraverso il cavo di trasmissione. Certo che utilizzare un cavo MIDI per trasmettere un solo messaggio non sarebbe conveniente e per questo motivo una linea MIDI è stata divisa in 16 canali

E' importante quando si esegue il collegamento di apparecchiature musicali ricordare che:

- La porta MIDI-IN va collegata ad una porta MIDI-OUT o MIDI-THRU.
- La porta MIDI-OUT deve essere collegata esclusivamente ad una porta MIDI-IN.
- La porta MIDI-THRU deve essere collegata esclusivamente ad una porta MIDI-IN.

Il connettore standard MIDI (fig 2) è di tipo DIN (Deutsche Industrie Normen) a 5 pin:

- Il pin 1 e 3 sono riservati per sviluppi futuri (non utilizzati)
- Il pin 2 è collegato a massa (per evitare interferenze).
- Il pin 5 collega il filo necessario per la trasmissione dei dati, mentre il pin 4 collega il cavo che assicura l'alimentazione corretta a + 5V.

(logici) di comunicazione. Ognuno di questi canali di comunicazione può trasmettere una determinata informazione che può essere ricevuta da un dispositivo sintonizzato su quel canale.

Esempio di un messaggio MIDI

Quando si preme un tasto ad una master keyboard collegata con un cavo midi ad un expander essa manda l'informazione al modulo sonoro di che nota è stata premuta, con quale intensità, in quale ottava. Il messaggio in questione si chiama NOTE ON. Lo status byte informa che è stata suonata una nota. Il numero della nota che deve essere suonata viene trasmesso dal primo data byte. Se la tastiera è dinamica (cioè in grado di stabilire con quale intensità viene suonata una nota) allora quando si preme un tasto, oltre ai due messaggi precedenti se ne aggiunge un terzo di tipo data byte, altrimenti se la tastiera non è sensibile alla dinamica, viene inviato un valore di default (generalmente 64).

Protocollo MIDI

Struttura dei messaggi MIDI

I messaggi MIDI si dividono in due categorie principali:

Channel Message: i messaggi di canale sono quelli che possono essere indirizzati ad uno qualsiasi dei sedici canali MIDI, essi si dividono in:

- *Channel voice message*
- *Channel mode message*

System Message: questi messaggi, a differenza dei Channel message non contengono informazioni di canale, bensì di sistema, quindi possono essere ricevuti da qualsiasi apparecchiatura MIDI. Essi si dividono ulteriormente in:

- *System common message*
- *System Real time message*
- *System exclusive message*

Channel Voice Message

Note on - Note off: quando ad una tastiera si preme un tasto, essa manda via MIDI un messaggio detto Note on con le specifiche di che nota si tratta, con quale intensità si sta suonando (se la tastiera è sensibile alla dinamica) e in quale ottava.

Un esempio di messaggio di Note on è il seguente:

Status Byte: 1001cccc (Note on), cccc indica il canale di trasmissione.

Data Byte 1: 0nnnnnnn (Key number) gli n indicano il numero della nota che deve essere suonata.

Data Byte 2: 0vvvvvvv (Velocity number), le v indicano la velocità con cui si preme il tasto.

N.B. Le note sono rappresentate da numeri: 0 per la nota C, 2 per C#, 3 per D ecc. Quando il tasto viene rilasciato la tastiera produce un messaggio di Note off che indica quale tasto è stato rilasciato. Il messaggio è simile a quello di Note on:

Status Byte: 1000cccc (Note off), cccc indica il canale di trasmissione.

Data Byte 1: 0nnnnnnn (Key number) gli n indicano il numero della nota che deve essere rilasciata.

Data Byte 2: 0vvvvvvv (Release velocity), le v indicano la velocità di rilascio del tasto.

N.B. In genere il messaggio Note off viene sostituito con un messaggio Note on con velocità uguale a 0.

Binario Hex Canale

0000	00	1
0001	01	2
0010	02	3
0011	03	4
0100	04	5
0101	05	6
0110	06	7
0111	07	8
1000	08	9
1001	09	10
1010	0A	11
1011	0B	12
1100	0C	13
1101	0D	14
1110	0E	15
1111	0F	16

Pitch bender ch'ange: quando viene spostata la leva pitch bender (quella specie di joystick che si trova di lato ad alcune tastiere con la funzione di modificare l'altezza di una nota di un certo valore) della tastiera essa ha l'effetto di produrre un messaggio MIDI in cui viene indicata l'altezza della nota. Per definire il valore dell'incremento o decremento vengono utilizzati due data byte.

Un esempio di Pitch bender change è il seguente:

Status Byte: 1110cccc (Pitch bender change), cccc indica il canale di trasmissione.

Data Byte 1: 0nnnnnnn (MSB key number) gli n indicano un numero (binario) compreso tra 0 e 127.

Data Byte 2: 0vvvvvvv (LSB), le v indicano un numero (binario) compreso tra 0 e 127.

Control ch'ange: questo messaggio descrive lo stato di un qualsiasi controller.

Status Byte: 1011cccc (Control change), cccc indica il canale di trasmissione.

Data Byte 1: 0nnnnnnn (Control number) gli n indicano il numero del controller da modificare.

Data Byte 2: 0vvvvvvv (Controller position), le v indicano il valore che il controller assumerà.

Channel pressare: questo messaggio serve per trasmettere in tempo reale ulteriori variazioni di pressione di un tasto dopo che è stato premuto.

Status Byte: 1101cccc (Channel pressure), cccc indica il canale di trasmissione.

Data Byte 1: 0nnnnnnn (After touch) gli n indicano il valore della pressione.

Data Byte 2: non c'è.

Polyphonic key pressare: questo messaggio è simile a quello precedente ma è capace di inviare le informazioni di variazione di pressione di un tasto già premuto in modo indipendente per ogni nota.

Status Byte: 1010cccc (Polyphonic key pressure), cccc indica il canale di trasmissione.

Data Byte 1: 0nnnnnnn (Key number) gli n indicano il valore della nota a cui è associato il data byte 2.

Data Byte 2: 0vvvvvvv(After touch), le v indicano il valore della pressione.

Program ch'ange: questo messaggio serve per selezionare una patch di uno strumento della periferica.

Status Byte: 1100cccc (Program change), cccc indica il canale di trasmissione.

Data Byte 1: 0nnnnnnn (Preset number) gli n indicano il numero della patch da selezionare.

Data Byte 2: non c'è.

Channel Mode Message

Nel protocollo MIDI vi sono dei messaggi utilizzati per controllare le funzioni generali di uno strumento musicale, si tratta dei messaggi di modo. Esistono tre stati fondamentali:

- 1) OMNI: sta a significare che lo strumento risponde a tutti i messaggi ricevuti contemporaneamente su tutti e sedici i canali.
- 2) POLI: significa che lo strumento viene suonato in maniera polifonica, ossia utilizza più di una voce.
- 3) MONO: significa che lo strumento viene suonato in maniera monofonica (utilizza una sola voce).

I modi disponibili sono ottenuti combinando gli stati appena descritti:

modo 1: Omni On / Poly modo 2: Omni Off / Poly modo 3: Omni On / mono modo 4: Omni Off / mono

Omni mode on: questo messaggio trasmette lo stato di Omni On.
Status Byte: 1011cccc (Control change), cccc indica il canale di trasmissione.

Data Byte 1: 01111100 (Omni mode on)

Data Byte 2: 00000000 (Ignorato)

Omni mode off: questo messaggio trasmette lo stato di Omni Off.

Status Byte: 1011cccc (Control change), cccc indica il canale di trasmissione.

Data Byte 1: 01111101 (Omni mode off).

Data Byte 2: 00000000 (Ignorato).

Mono mode on: questo messaggio trasmette lo stato di mono mode on.

Status Byte: 1011cccc (Control change), cccc indica il canale di trasmissione.

Data Byte 1: 01111110 (Mono mode on).

Data Byte 2: 0000vvvv (Ch Allocation), serve per sapere stabilire il numero di canali utilizzati per ricevere e trasmettere voci monofoniche.

Poly on: questo messaggio trasmette lo stato di Poly on.

Status Byte: 1011cccc (Control change), cccc indica il canale di trasmissione.

Data Byte 1: 01111111 (Poly on).

I controller

Quando si vuole alzare il volume di uno strumento, oppure come si dice in gergo *panpottare* il suono (regolare il suono a destra o a sinistra) si fa ricorso ai controller. In realtà con i controller è possibile gestire determinate funzioni che regolano lo stato di un suono. A questo scopo ci sono due tipi di controller: **controller continui** e **controller a interruttore**. I primi permettono di indicare variazioni di uno stato di cose, partendo da 0 fino ad arrivare a 127. Sono utilizzati quindi per il volume, il balance, ecc. I controller a interruttore invece si comportano come degli interruttori che attivano o disattivano una determinata funzione. Un valore compreso tra 0 e 63 indica il valore off, mentre un valore compreso tra 64 e 127 indica il valore on.

MSB - LSB: i Controller che vanno da 0 a 31 sono detti MSB ossia Most Significant Byte, che in italiano significa byte più significativi, mentre i controller che vanno da 32 a 127 vengono chiamati LSB, ossia Least Significant Byte (Byte meno significativi).

Data Byte 2: 00000000 (Ignorato).

Reset all: per riportare tutti i controller in una posizione di default si utilizza questo messaggio.

Status Byte: 1011cccc (Control change), cccc indica il canale di trasmissione.

Data Byte 1: 01111001 (Reset all).

Data Byte 2: 00000000 (Ignorato).

Local control: questo messaggio serve per poter scollegare o collegare il modulo sonoro dalla tastiera. Per esempio se local control viene impostato ad off, la tastiera non è in grado di controllare il generatore di suono interno.

Status Byte: 1011cccc (Control change), cccc indica il canale di trasmissione.

Data Byte 1: 01111010 (Local control).

Data Byte 2: 0vvvvvvv (On / Off) 00000000 = On, 01111111 = Off.

All note off: questo messaggio serve a spegnere tutte le note che sono attive in un dato istante.

Status Byte: 1011cccc (Control change), cccc indica il canale di trasmissione.

Data Byte 1: 01111011 (All note off).

Data Byte 2: 00000000 (Ignorato).

System Common Message

Quarter frame: questo messaggio serve per convertire il codice di sincronizzazione SMPTE in MTC (Midi Time Code).

Status Byte: 11110001 (Quarter frame).

Data Byte 1: nnnnvvvv (Type / Data).

Data Byte 2: non c'è.

Song position: permette di assegnare ad ogni beat di una sequenza un indirizzo assoluto.

Status Byte: 11110010 (Song position).

Data Byte 1: Onnnnnnn (Puntatore 1).

Data Byte 2: 0vvvvvvv (Puntatore 2).

Song select: questo messaggio permette di selezionare una song in un sequencer o in un altro dispositivo simile.

Status Byte: 11110011 (Song select).

Data Byte 1: Onnnnnnn (Number).

Data Byte 2: non c'è.

Tune request: questo messaggio si utilizza per accordare lo strumento.

Status Byte: 11110110 (Tune request).

Data Byte 1: non c'è.

Data Byte 2: non c'è.

System Real Time Message: sono messaggi di sincronizzazione ed hanno la caratteristica di avere solo Status Byte.

Clock: viene utilizzato per sincronizzare le apparecchiature MIDI. Viene inviato 24 volte per ogni nota di un quarto.

Status Byte: 11111000 (Midi clock).

Data Byte 1: non c'è.

Data Byte 2: non c'è.

Start: serve per posizionare tutte le apparecchiature all'inizio della song e farle partire in registrazione o riproduzione.

Status Byte: 1111010 (Start).

Data Byte 1: non c'è.

Data Byte 2: non c'è.

Continue: serve per riprendere le song da dove è stata interrotta.

Status Byte: 11111011 (Continue).

Data Byte 1: non c'è.

Data Byte 2: non c'è.

Stop: serve per fermare la riproduzione o registrazione di tutte le apparecchiature collegate.

Status Byte: 11111100 (Stop).

Data Byte 1: non c'è.

Data Byte 2: non c'è.

Active sensing: serve per tenere attive tutte le apparecchiature MIDI collegate.

Status Byte: 11111110 (Active sensing).

Data Byte 1: non c'è.

Data Byte 2: non c'è.

System reset: imposta tutte le apparecchiature collegate nelle condizioni di default.

Status Byte: 11111111 (System reset).

Data Byte 1: non c'è.

Data Byte 2: non c'è.

System Exclusive Message: questi messaggi servono per accedere alle funzioni del dispositivo MIDI e per poterlo programmare in maniera altrimenti impossibile. Il protocollo MIDI infatti non prevede per esempio la modifica della curva di inviluppo di un timbro musicale, ma intervenendo con un messaggio di sistema esclusivo direttamente al dispositivo è possibile effettuare questa modifica. Ci sono molti parametri che si possono modificare con messaggi di sistema esclusivo e per sapere quali sono bisogna vedere il manuale dello strumento MIDI. Poichè ci sono diverse apparecchiature di marche differenti si è pensato di introdurre all'interno di un messaggio di sistema esclusivo un ID. Cioè un codice identificativo che indica alle apparecchiature aventi lo stesso ID che il messaggio di sistema esclusivo è per loro e tutte le altre apparecchiature collegate, ma con ID differente, ignoreranno quel messaggio.

System exclusive: Status Byte: 11110000 (Sistema esclusivo). Lo status bytes è seguito da una serie di byte che costituiscono l'informazione (compreso l'ID).

End of exclusive: Status Byte: 11110111 (Fine di messaggio di sistema esclusivo). Indica la fine di messaggio di sistema esclusivo.

Gli standard: general midi level 1

Quando si cambia uno strumento con il program change esso dà l'informazione ad un modulo sonoro che a seconda dell'implementazione può far corrispondere il suono desiderato oppure un altro suono completamente diverso da quello scelto. Ad esempio, se su di un sintetizzatore A la patch 1 corrisponde ad un organo non è detto che la patch 1 di un sintetizzatore B sia uguale, anzi potrebbe corrispondere tutt'altro strumento. Si tratta quindi di stabilire uno standard in modo che tutti i sintetizzatori, sebbene con suoni diversi, facciano corrispondere gli strumenti con patch già prestabiliti. Nel 1991 fu introdotto lo standard GM, general MIDI level 1. Uno strumento musicale per essere GM compatibile deve associare ad ogni valore di program change una path stabilita. Inoltre per convenzione, uno strumento GM deve utilizzare il canale 10 per la batteria e le percussioni.

Roland GS General System. Lo standard GS è un'evoluzione dello standard general MIDI level 1 ed è stato introdotto dalla Roland. Questo standard permette infatti di utilizzare più strumenti rispetto ai 128 consentiti dal GM, e introduce nuovi tipi di messaggi e di controller. Il GS introduce il concetto di Bank Select, esso unito al program change permette di avere 128 variazioni per ogni singolo strumento. N.B. il numero di banchi a disposizione in un sintetizzatore dipende esclusivamente dal produttore dello strumento e non dalle specifiche GS. Inoltre lo standard GS permette di modificare effetti audio come il chorus e il riverbero ecc.

Yamaha XG. Lo standard XG fu introdotto dalla Yamaha nel 1994 e rappresenta un'ulteriore evoluzione del GM e GS. Esso rappresenta lo standard più evoluto attualmente disponibile: è aumentato il numero di strumenti, il numero di Drum Kit, il numero di effetti ecc. Ovviamente lo standard XG è compatibile con lo standard GM level 1, ossia se suoniamo una sequenza che rispetta le regole GM su di un modulo XG essa verrà suonata con gli stessi strumenti che gli erano stati fissati per il GM. Anche il viceversa vale, cioè suonare una base MIDI XG su di un modulo GM, ma ovviamente si perdono tutti i vantaggi offerti dallo standard XG. Inoltre un sintetizzatore XG è in grado di riprodurre correttamente le sequenze GS, previa trasmissione di un comando GS Reset.

0	Bank Select	45	Effect control 2
1	Modulation wheel	46...47	Non definito
2	Breath control	48...51	General #1,2,3,4
3	Non definito	52...63	Non definito
4	Foot control	64	Dumper pedal
5	Portamento time	65	Portamento on/off
6	Data Entry	66	Sostenuto on/off
7	Channel Volume	67	Soft pedal on /off
8	Balance	68	Legato Footswitch
9	Non definito	69	Hold 2
10	Pan	70...79	Sound Controller 1-10
11	Expression Controller	80...83	General Purpose 5,6,7,8
12	Effect control 1	84	Portamento control
13	Effect control 2	85...90	Non definito
14...15	Non definito	91...95	Effects 1,2,3,4,5 Depth
16...19	General #1,2,3,4	96	Data entry +1
20...31	Non definito	97	Data entry -1
32	Bank Select	98	NRPN LSB
33	Modulation wheel	99	NRPN MSB
34	Breath control	1000	RPN LSB
35	Non definito	101	RPN MSB
36	Foot control	102...119	Non definito
37	Portamento time	120	All Sound Off
38	Data Entry	121	Reset All Controllers
39	Channel Volume	122	Local control on/off
40	Balance	123	All note off
41	Non definito	124	Omni mode off
42	Pan	125	Omni mode on
43	Expression	126	Poly mode on/off
44	Effect control 1	127	Poly mode on

Piano Bass Reed Synth Effects

1=Acoustic Grand Piano
2=Bright Acoustic Piano
3=Electric Grand Piano
4=Honkytonk Piano
5=Rhodes Piano
6=Chorused Piano
7=Harpichord
8=Clavinet
33=Acoustic Bass
34=Electric Bass (finger)
35=Electric Bass (pick)
36=Fretless Bass
37=Slap Bass 1
38=Slap Bass 2
39=Synth Bass 1
40=Synth Bass 2
65=Soprano Sax
66=Alto Sax
67=Tenor Sax
68=Baritone Sax
69=Oboe
70=English Horn
71=Bassoon
72=Clarinet
97=FX 1 (rain)
98=FX 2 (soundtrack)
99=FX 3 (crystal)
100=FX4 (atmosphere)
101=FX 5 (brightness)
102=FX 6 (goblins)
103=FX 7 (echoes)
104=FX 8 (sci-fi)

Chromatic Strings Pipe Ethnic

9=Celesta
10=Glockenspiel
11=Music Box
12=Vibraphone
13=Marimba
14=Xylophone
15=Tubular Bells
16=Dulcimer
41=Violin
42=Viola
43=Cello

44=Contrabass
45=Tremolo Strings
46=Pizzicato Strings
47=Orchestral Harp
48=Timpani
73=Piccolo
74=Flute
75=Recorder
76=Pan Flute
77=Bottle Blow
78=Shakuhachi
79=Whistle
80=Ocarina
105=Sitar
106=Banjo
107=Shamisen
108=Koto
109=Kalimba
110=Bagpipe
111=Fiddle
112=Shanai

Organ Ensemble Synth Lead Percussive

17=Hammond Organ
18=Percussive Organ
19=Rock Organ
20=Church Organ
21=Reed Organ
22=Accordion
23=Harmonica
24=Tango Accordion
49=String Ensemble 1
50=String Ensemble 2
51=SynthStrings 1
52=SynthStrings 2
53=Choir Aahs
54=Voice Oohs
55=Synth Voice
56=Orchestra Hit
81=Lead 1 (square)
82=Lead 2 (sawtooth)
83=Lead 3 (calliope lead)
84=Lead 4 (chiff lead)
85=Lead(charang)
86=Lead 6 (voice)

87=Lead 7 (fifths)
88=Lead 8 (bass + lead)
113=Tinkle Bell
114=Agogo
115=Steel Drums
116=Woodblock
117=Taiko Drum
118=Melodic Tom
119=Synth Drum
120=Reverse Cymbal

Guitar Brass Synth Pad Sound Effects

25=Acoustic Guitar (nylon)
26=Acoustic Guitar (steel)
27=Electric Guitar (jazz)
28=Electric Guitar (clean)
29=Electric Guitar (muted)
30=Overdriven Guitar
31=Distortion Guitar
32=Guitar Harmonics
57=Trumpet
58=Trombone
59=Tuba
60=Muted Trumpet
61=French Horn
62=Brass Section
63=Synth Brass 1
64=Synth Brass 2
89=Pad 1 (new age)
90=Pad 2 (warm)
91=Pad 3 (polysynth)
92=Pad 4 (choir)
93=Pad 5 (bowed)
94=Pad 6 (metallic)
95=Pad 7 (halo)
96=Pad 8 (sweep)
121=Guitar Fret Noise
122=Breath Noise
123=Seashore
124=Bird Tweet
125=Telephone Ring
126=Helicopter
127=Applause
128=Gunshot

4.5. Haptic technologies

Haptic From the Greek *Haphe*, means pertaining to the sense of touch. It is the earliest sense to develop in the fetus, and the development of infants' haptic senses, and how that relates to the development of the other senses such as vision, has been the target of much research. Haptic technology refers to technology which interfaces the user via the sense of touch. It is an emerging technology that promises to have wide reaching implications.

History

One of the earliest forms of haptic devices is used in large modern aircraft that use servo systems to operate control systems. Such systems tend to be "one-way" in that forces applied aerodynamically to the control surfaces are not perceived at the controls, with the missing normal forces simulated with springs and weights. In earlier, lighter aircraft without servo systems, as the aircraft approached a stall the aerodynamic buffeting was felt in the pilot's controls, a useful warning to the pilot of a dangerous flight condition. This control shake is not felt when servo control systems are used. To replace this missing cue, the angle of attack is measured, and when it approaches the critical stall point a "stick shaker" (an unbalanced rotating mass) is engaged, simulating the effects of a simpler control system. This is an example of *haptic feedback*.

Teleoperators and simulators

Teleoperators are remote controlled robotic tools, and when contact forces are reproduced to the operator, it is called "haptic teleoperation". The first electrically actuated teleoperators were built in the 1950's at the Argonne National Lab, USA, by Dr. Raymond C. Goertz, to remotely handle radioactive substances. Since then, the use of "force feedback" has become more widespread in all kinds of teleoperators such as underwater exploration devices controlled from a remote location. When such devices are simulated using a computer (as they are in operator training devices) it is useful to provide the force feedback that would be felt in actual operations. Since the objects being manipulated do not exist in a physical sense, the forces are generated using haptic (force generating) operator controls. Data representing touch sensations may be saved or played back using such haptic technologies. Haptic simulators are currently used in medical simulators and flight simulators for pilot training (2004).

Games

Some low-end haptic devices are already common. Some joysticks and game controllers provide haptic feedback. The simplest form are "rumble packs", which are simply attachments which vibrate upon command from the software. Simulated automobile steering wheels are now available which provide the road "feel" for race car simulations. The ability to change the temperature of a controlling device could also be used. However, the technology may be cost prohibitive in terms of how much power it would need to operate properly.

Haptics in virtual reality

Haptics is gaining widespread acceptance as a key part of Virtual Reality systems, adding the sense of touch to previously visual-only solutions such as 'The Wedge' and more recently in laptop-based VR solutions such as the '3D-Mobile Immersive Workstation'. Most of these solutions use stylus-based haptic rendering, where the user interfaces to the virtual world via a tool or stylus, giving a form of interaction that is computationally realistic on today's hardware

Research

Some research has been done into simulating the different kinds of tactition by means of high-speed vibrations or other stimuli. One device of this type uses a pad array of pins, where the pins vibrate to simulate a surface being touched. While this does not have a realistic feel, it does provide useful feedback, allowing discrimination between various shapes, textures, and resiliencies.

Medicine

Various haptic interfaces for medical simulation may prove especially useful for training of minimally invasive procedures (laparoscopy/interventional radiology) and remote surgery using teleoperators. In the future, expert surgeons may work from a central workstation, performing operations in various locations, with machine setup and patient preparation performed by local nursing staff. Rather than traveling to an operating room, the surgeon instead becomes a *telepresence*. A particular advantage of this type of work is that the surgeon can perform many more operations of a similar type, and with less fatigue. It is well documented that a surgeon who performs more procedures of a given kind will have statistically better outcomes for his patients.

Literature

The use of haptic devices in entertainment appeared in the 1932 futurist fiction book *Brave New World* by Aldous Huxley. The author described a future entertainment theater where the arm rests of the seats had positions for the hands to rest that gave haptic stimulation. Rather than "the movies" these theaters and shows were called "the *feelies*". The programs exhibited were of an erotic nature.

Remote sexual relations

One envisioned use of haptics is for "teledildonics". Recent developments such as the 'CyberGlove' indicate that this may be plausible and not just an interesting idea. Some sex toys are now available which can be computer controlled; normally this comes in the form of an online movie or website which sends commands to the toy at scripted moments. Members of the pornography industry are often among the early adopters of new technologies; for example, the multiple-viewpoint feature of DVDs was first widely used in pornography.

Robotics

The Shadow Dextrous Robot Hand uses the sense of touch, pressure, and position to reproduce the human grip in all its strength, delicacy, and complexity. The SDRH was first developed by Richard Greenhill and his team of engineers in Islington, London, as part of The Shadow Project, (now known as the Shadow Robot Company) an ongoing research and development program whose goal is to complete the first convincing humanoid. An early prototype can be seen in NASA's collection of humanoid robots. (<http://robonaut.jsc.nasa.gov/humanoids.html>). The Dextrous Hand has haptic sensors embedded in every joint and in every finger pad which relay information to a central computer for processing and analysis. Carnegie Mellon University in Pennsylvania and Bielefeld University in Germany in particular have found The Dextrous Hand is an invaluable tool in progressing our understanding of haptic awareness and are currently involved (2006) in research with wide ranging implications.

Arts

Touching is not limited to a feeling, but it allows interactivity in real-time with virtual objects. Thus haptics are commonly used in virtual arts, such as sound synthesis or graphic design/animation. The haptic device allows the artist to have direct contact with a virtual instrument which is able to produce real-time sound or images. We can quote the physical modelling synthesis which is an efficient modelling theory to implement cross-play interaction between sound, image, and physical objects. For instance, the simulation of a violin string produces real-time vibrations of this string under the pressure and expressivity of the bow (haptic device) held by the artist.

"I hear and I forget - I see and I remember - I do and I understand"

Chinese Proverb

Handshake Haptics

As one moves the haptic device, one can imagine that the end of the haptic device acts as a 3-D mouse. The point at which the pen connects to the final link on the Omni device is the device position. It may also be considered as the location of the cursor in virtual reality terms. One can now imagine that there is a virtual object within the Cartesian workspace of the Omni device. The computer has the ability to continually update the Cartesian location of the cursor and determine whether or not the cursor has contacted or penetrated into the virtual object. If the virtual object has been penetrated, one can model its surface as a spring and damper.

Handshake introduces the exciting Hands-On Teaching Technology™ (HOTT™) - a new way for students of all ages to actively learn. Utilizing dynamic technology developed for training medical students and surgeons, simulating surgery, and the rehabilitation of stroke patients, Handshake works with educators to bring the power of "hands-on" haptics (the sense-of-touch and forces) to classroom and home learning. Our hands-on applications provide force and touch feedback from virtual models on the computer to students. Our groundbreaking, proprietary, technology also enables students in one location to interact remotely with users (such as another student or a therapist) and feel both the learning objects and the guidance of the therapist or mentor. HOTT Learning Tools are Designed for General, Advanced and Special Needs Students: an extensive research has shown that students find "hands-on" technology exciting, engaging, and interesting regardless of the level or ability of the student.

A patented approach creates an opportunity for students to interact dynamically with their learning environment, increasing their comprehension and assisting in hard-to-learn or abstract concepts. With HOTT, students can receive guidance and assistance with tasks, customized to each student's specific level, aiding in learning and reducing frustrations. The power of HOTT is reached through our "Friend-to-Friend" program. Hands-on exercises can effectively be enabled by interacting with other students, teachers or therapists in the classroom or over a network, creating a supportive and collaborative environment. Learning is Fun When You Can Feel the Difference: HOTT Tools engage students by exploring in 2D & 3D providing them with an opportunity to touch new and exciting environments!

Literacy: The HOTT Early Literacy program is a series of five modular computer exercises for children four to seven years of age that use engaging kinesthetic activities enabling young learners to develop the skills necessary to learn to read. As a result of years of academic research, HOTT Learning Tools have been created specifically for classroom and home use bringing a new dimension to childrens' education. Students and teachers now have access to proven technology that even astronauts and surgeons use for learning new tasks! Students are able to interact dynamically with their learning environment, through our hands-on haptics approach, increasing comprehension, assisting with hard-to-learn or abstract concepts and

enhancing student independence and choice. Travel with Hoppy the Hare through full multi-sensory exercises that combine auditory, visual and kinaesthetic teaching modalities providing Phonemic awareness, Visual cuing, Kinesthetic concept development, Self-assessment.

Six Core Modules: Modules build upon each other sequentially but can also be used as stand-alone units. Carrot Patch Round-Up - Letter Identification; Bramble Scramble - Alphabet Order; Carrot Sticks - Learn-to-Print; See the letter; Feel the outline of the letter; Use assistive guidance if required; Hare Raising - Rotate & Feel the Letter; Feel the outline of the shape - learn the difference between a 2D and 3D shape; Trace the shape with help from our OMNI™ robot; Pick 'n Place - Choose the Word, Feel the Picture, Opposites; Tales from the Briar Patch - Interact with the Story; Read a story: See pictures explaining the main concepts; Interact with the pictures to learn more about the ideas. For example, the difference between hard and soft.

Science: The HOTT Science Quest Solar System program is a series of computer-based labs, for students eleven to fourteen years of age, that use engaging and immersive activities enabling students to experience abstract and hard-to-learn concepts in a kinesthetic environment. Students and teachers now have access to proven technology that even astronauts and surgeons use for learning new tasks! As a result of years of academic research, HOTT Learning Tools have been created specifically for classroom and home use bringing a new dimension to student education. Students are able to interact dynamically with their learning environment, through our hands-on haptics approach, increasing their understanding of challenging abstract concepts, applying kinesthetic models, and utilizing cost-effective teaching methods for science based programs.

Beating Heart: Feel the beating heart using a stylus; Change the pulse rate; Explore areas of the heart - rotate and examine it from all angles. **Solar System:** Many concepts in science are difficult for students to comprehend, especially those that relate to Astronomy. With the Solar System Module, the first in our Science Series covering Physics, Chemistry and Biology students will be able to: Navigate the solar system in their space ship; Feel the different gravities of each planet; Learn about planetary orbits in 3D; Experience different properties such as the rings of Saturn, the gaseous sun, meteorites, weightlessness; Change the views to zoom in on each planet, the sun, and the asteroid belt; Compare each planet to Earth on key factors such as gravity. **Physics Playground:** Interact with a 3D shape; Explore friction, mass, force; Explore perimeter and volume.

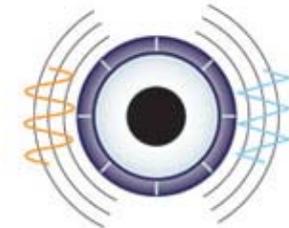
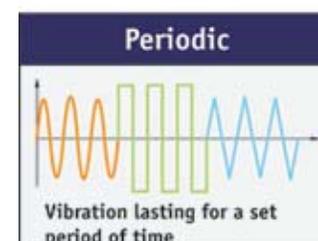
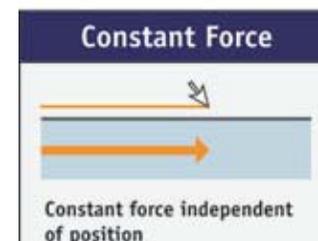
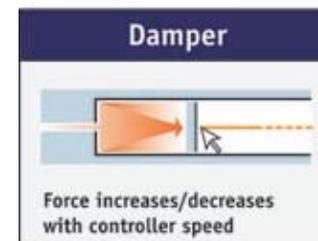
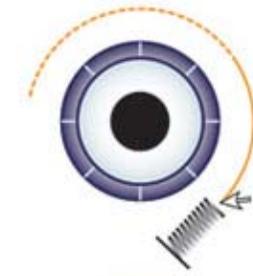
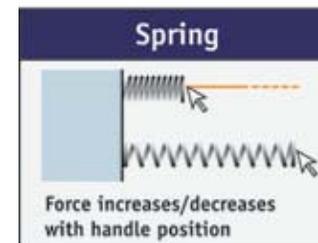
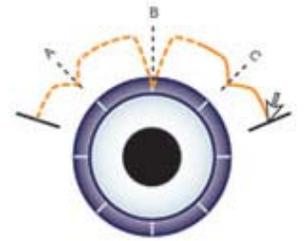
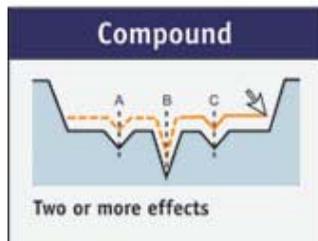
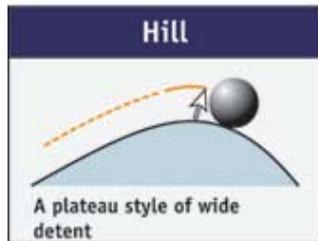
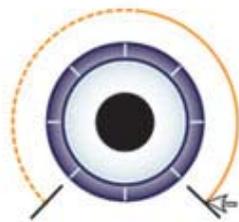
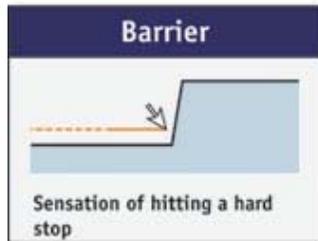
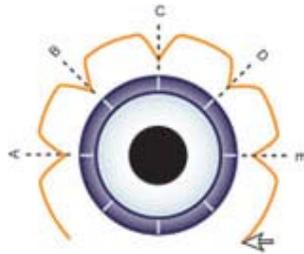
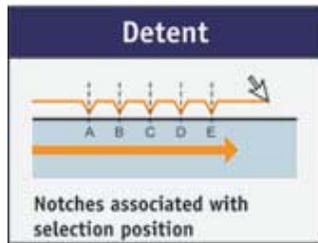
Lending a "Hand" to Special Needs Students

Handshake works with leading educators in special education to develop helpful learning applications for students with a wide variety of challenges including autism, dyslexia, ADHD and brain-based injuries or conditions. Hands-on haptic learning provides an ideal environment for special needs students to discover a new way of tackling old problems. HOTT Learning Tools captures students' attention, maintains interest and provides a lasting learning experience. Today, Handshake is pioneering the ability of therapists to work remotely with students offering as much "interactive" time with the therapist as possible and increasing the frequency of interaction.



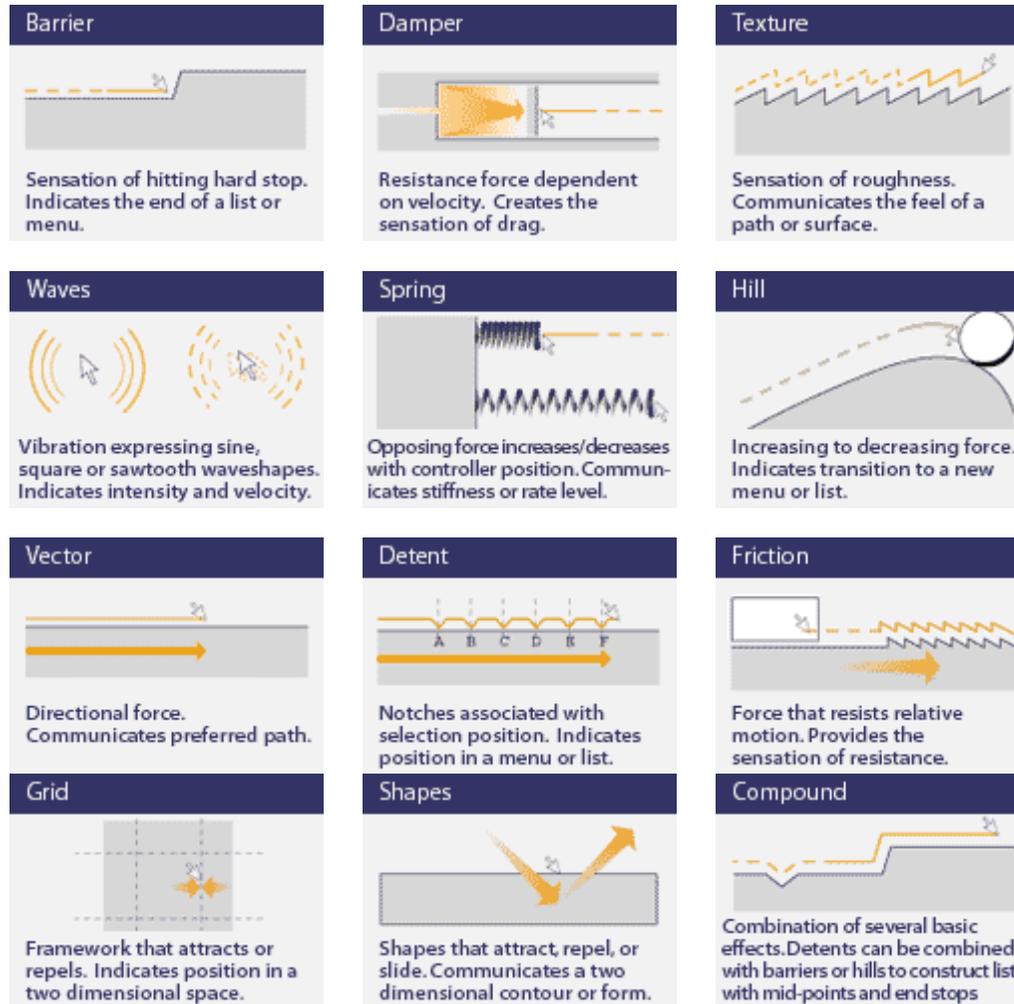
4.5.0. Sample Haptic Effects

Rotary controls can be programmed with a wide variety of tactile sensations. By combining effects, many more sensations are possible to match specific interface controls. The result is a more intuitive user experience.



Other Sample Effects

Automotive controls can be programmed with a wide variety of tactile sensations. These are just some of the sensations that can be programmed with Immersion TouchSense™ technology. By combining effects, many more sensations are possible to match specific interface controls. The result is a more intuitive driving experience. These effects are applied in BMW's I-drive knob navigation system.



5

capitolo



5.0. MicroElectroMechanicalSystems

Esistono ad oggi trasduttori per un gran numero di effetti fisici e proprietà fondamentali della materia che possono permetterci di ottenere informazioni e creare dei modelli. Possiamo approntare una piccola lista dei vari tipi di sensori disponibili:

- prossimità immateriali
- fotoelettrici
- ottici
- laser
- a fibra ottica
- sensori codificatori angolari
- spostamento lineare
- di pressione
- di temperatura
- umidità
- portata
- livello
- peso
- forza
- coppia
- accelerometri
- inclinazione
- acustici
- gas
- bussola e sensore navigazione
- dinamo tachimetrica
- sensore biometrica
- sensore di campo magnetico
- controllo rotazione
- di esplosione
- movimento ruota
- posizione ad effetto hall
- di tocco
- piezoelettrico
- reed
- tachimetro
- trasduttore densità
- trasduttore elettrostatico
- consistenza
- viscosità.

Questi sensori, integrati fra loro, fungono da interfaccia tra il mondo dell'uomo e quello della macchina. Una delle massime aspirazioni sarebbe quella di integrare questi sensori e comandare il tutto con il pensiero, senza bisogno di ulteriori componenti di interfaccia, come controller o tastiere.

La costruzione di interfacce cervello-macchina è possibile, e si stanno muovendo i primi passi in questa direzione, come dimostra l'articolo che segue.

Belle, un lemure notturno così piccolo da stare comodamente nella tasca di una giacca, sbatte le palpebre dei suoi enormi occhi mentre un tecnico gli inserisce quattro spinotti nelle prese praticate nella parte superiore del cranio. Nella stanza accanto, su uno schermo di calcolatore, pulsano le onde che rappresentano i segnali elettrici provenienti da una novantina di neuroni cerebrali di Belle. Registrati in quattro aree distinte della corteccia, questi segnali consentono di vedere ciò che accade nel suo cervello quando la bestiola allunga la mano e preme uno dei quattro bottoni che ha davanti per ottenere la ricompensa, qualche goccia di succo di mela.

Miguel Nicolelis, un neurobiologo della Duke University che compie ricerche pionieristiche sull'uso degli impianti neuronali per lo studio del cervello, indica compiaciuto i dati che compaiono sul monitor e dice: «Questa schermata è unica nel suo genere al mondo». La stessa cosa si potrebbe dire di Nicolelis, che è uno degli uomini di punta in questo campo tanto competitivo e importante. In tutto il mondo sono soltanto una mezza dozzina i gruppi di ricerca che perseguono questo stesso obiettivo: capire meglio come funziona la mente e poi sfruttare queste conoscenze per costruire sistemi da impiantare nel cervello in modo da comandare calcolatori e altri dispositivi. Nicolelis chiama questi sistemi "interfacce ibride cervello-macchina" (licm). Di recente, lavorando nel laboratorio dedicato ai Contatti uomo-macchina del Massachusetts Institute of Technology, Nicolelis ha ottenuto un risultato di primaria importanza sul fronte delle licm: è riuscito a mandare segnali da singoli neuroni del cervello di Belle a un robot, che sulla base di questi dati ha imitato in tempo reale i movimenti del braccio della scimmietta.

In futuro, prevede Nicolelis, le licm consentiranno al cervello umano di comandare dispositivi artificiali progettati per ripristinare funzioni sensoriali e motorie perdute. Chi è colpito da paralisi, per esempio, potrebbe comandare una sedia a rotelle a motore o un braccio

protetico, e magari anche i propri arti. «Pensate», dice Nicoletis, «se qualcuno potesse fare per il cervello ciò che lo stimolatore cardiaco ha fatto per il cuore». Ed è convinto che il cervello sarà capace di assimilare rapidamente i dispositivi artificiali, più o meno come un musicista si abitua a poco a poco a sentire lo strumento come parte del proprio corpo.

Gli esperimenti in corso in altri laboratori confermano che si tratta di un'idea plausibile. Alla Emory University il neurologo Phillip Kennedy ha aiutato pazienti colpiti da paralisi gravi a comunicare tramite un impianto cerebrale che permette loro di spostare un cursore sullo schermo di un computer. E questi impianti potrebbero anche far luce su alcuni dei misteri non risolti del cervello. I neurobiologi ne sanno ancora piuttosto poco sui meccanismi con cui i segnali elettrici e chimici emessi dai milioni di neuroni del cervello consentono di percepire i colori e gli odori o di attuare i movimenti di precisione dei calciatori brasiliani, le cui foto ornano le pareti dell'ufficio di Nicoletis, che è nato a San Paolo. «Non possediamo un modello compiuto del funzionamento del cervello», dice Nicoletis. «Abbiamo soltanto delle vaghe impressioni».

Gli ultimi esperimenti di Nicoletis, tuttavia, dimostrano che prelevando segnali da più neuroni situati in diverse aree del cervello è possibile ottenere abbastanza informazioni su ciò che il cervello sta facendo. Nel caso di Belle, queste informazioni sono sufficienti per capire che la scimmia ha intenzione di compiere un particolare movimento qualche decimo di secondo prima che lo esegua. E Nicoletis è riuscito nella notevole impresa del braccio del robot proprio perché il suo gruppo è stato in grado di misurare con efficacia decine di neuroni simultaneamente per molti mesi, ciò che in precedenza costituiva un impedimento tecnico fondamentale.

Ci sono ancora molti ostacoli da superare per poter costruire tra il cervello umano e i dispositivi artificiali un'interfaccia comoda e affidabile che renda gli arti protetici o i calcolatori comandati dalla mente qualcosa di più che una semplice curiosità da laboratorio. Tra i problemi più importanti c'è la messa a punto di complessi di elettrodi e di tecniche chirurgiche che consentano di registrare a lungo e senza rischi l'attività neuronale. Nicoletis ha cominciato a lavorare con il dipartimento di ingegneria biomedica della Duke University all'allestimento di un chip per telemetria che raccolga e trasmetta i dati attraverso il cranio senza il solito groviglio di prese e cavi. E quest'anno sarà nominato condirettore del nuovo Centro di Ingegneria neuronale e Calcolo neuronale della Duke University, con il compito di

studiare combinazioni inedite di informatica, elettronica integrata e neurobiologia. Per Nicoletis questa impresa rientra in un'imminente rivoluzione che potrebbe portare le licm a diffondersi quanto i dispositivi palmari e far sorgere una nuova branca industriale imperniata sul cervello.

L'attuale stato dell'arte si avvale ancora (se così si può dire, perché questo ancora nasconde grandi meraviglie) di un controllo del tipo sensore di grandezza fisica - trasduttore in grandezza elettrica o ottica, contemplando dispositivi che integrano elementi di sensing diversi a scala micrometrica, e che quindi ci avvicinano all'essere quasi invisibili all'uomo, obiettivo primario per il campo delle interfacce. Sono questi i dispositivi MEMS: Micro Electro Mechanical Systems.

In *MEMS and NEMS Systems, Devices, and Structures*, MEMS is defined as follows:

"The MEMS is the batch-fabricated integrated microscale system (motion, electromagnetic, radiating energy and optical microdevices/microstructures - driving/sensing circuitry - controlling/processing integrate circuits) that:

1. Converts physical stimuli, events, and parameters to electrical, mechanical, and optical signals and vice versa;
2. Performs actuation, sensing and other functions;
3. Comprise control (intelligence, decision-making, evolutionary learning, adaptation, self-organization, etc.), diagnostics, signal processing, and data acquisition features," (Lyshevski)

Basically, MEMS is a system that consists of microstructures, microsensors, microelectronics, and microactuators. Microstructure builds the framework of the system; microsensor detects signals; microelectronics processes the signals and gives commands to the microactuator to react to these signals.

Borrowed from IC technology, MEMS is commonly built on silicon wafer by selectively etching away parts of the wafer or add new structural layers to form the mechanical and electromechanical devices.

Temperature

ADI offers a complete range of analog and digital temperature sensors. Choose one of ADI's analog temperature sensors if a simple voltage or current out function is required. Unlike traditional conventional methods, these analog sensors eliminate the need for trimming, buffering and linearization. This saves on both cost and unnecessary circuit design. ADI also has a selection of high accurate, low cost digital temperature sensors with I²C, SPI, Mark-Space and Trip point interfaces. For more intricate system design applications, some of these digital temperature sensors have integrated DACs, ADCs, references and limit alarm registers. It may be concluded that this complete family of temperature sensors has the diversity to be used in a broad range of applications.

AD22100 - Voltage Output Temperature Sensor with Signal Conditioning

AD22103 - Voltage Output 3.3 V Temperature Sensor with Signal Conditioning

AD22105 - Low Voltage, Resistor Programmable Thermostatic Switch

Accelerometers

15 years ago, Analog Devices revolutionized automotive airbag systems with its unique MEMS[®] (integrated Micro Electro Mechanical System) accelerometers. MEMS accelerometers were the first products in an array of Motion Signal Processing™ solutions to use innovative design techniques to integrate small, robust sensors with advanced signal conditioning circuitry on a single chip. Today, ADI offers the industry's broadest accelerometer portfolio, with products addressing a range of user needs including high performance, low power consumption, and small size.

The ADXL products are the foundation for Analog Devices' accelerometer family, and utilize the MEMS surface micromachining process that has been used to ship over 200 million sensors for automotive safety, consumer and industrial applications. This technology enables volume production of reliable, high quality, cost-effective products. Available in low-*g* or high-*g* sensing ranges, MEMS

accelerometers are used to measure position, motion, tilt, shock, and vibration in a broad array of applications.

ADIS16006 - Dual Axis ±5g Accelerometer with SPI Interface

ADIS16203 - Programmable Single Axis 360° Inclinometer

ADXL330 - Small, Low Power, 3-axis ±3g iMEMS[®] Accelerometer With Analog Output

Gyroscopes

Analog Devices' MEMS[®] Gyros are part of Analog Devices' growing family of Motion Signal Processing™ solutions. The MEMS Gyros are more reliable, power efficient, user friendly, compact, cost-effective, and offer more robust performance under severe conditions than other available gyros. Fully integrated, they combine mechanical structures with signal processing circuitry onto a monolithic IC.

At the core of all Analog Devices' MEMS Gyros is the unique MEMS surface micromachining process that has been used to produce over 200 million sensors for automotive safety, consumer and industrial applications. This technology enables the production of reliable, high quality products at the lowest possible cost. By integrating both mechanical elements and electronics on a single silicon die, MEMS products employ innovative design techniques that increase performance and robustness.

ADIS16080 - ±80°/sec Yaw Rate Gyro with SPI Interface

ADIS16120 - Low Noise Angular Rate Sensor

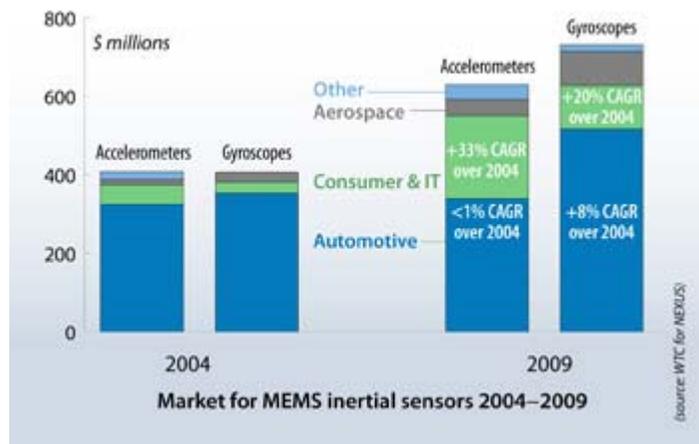
ADIS16100 - ±300°/sec Yaw Rate Gyro with SPI Interface

Mems inertial sensors

The market for MEMS inertial sensors (accelerometers and gyroscopes) is set to grow from \$835 million in 2004 to over \$1360 million in 2009 — a CAGR of 10%. Currently, the main applications

are in the **automotive industry**. These markets are well established and growth rates range from a **stagnant 1%** for airbag acceleration sensors up to 8% for gyroscopes used in ESP units and GPS navigation assistance. Much more exciting for MEMS inertial sensors is the **market opportunity for mobile applications and consumer electronics** (see [table](#)). Over the next few years, we predict annual growth rates **exceeding 30%** for accelerometers. Mobile phones in particular will provide multi-axis accelerometers with interesting opportunities in **menu navigation, gaming, image rotation, pedometers, GPS navigation** and the like. Gyroscopes are largely servicing markets for **image stabilization** and **HDD protection** in camcorders.

In contrast to the automotive sector, consumer applications feature relaxed specifications. Failure rates for automotive electronic control units (ECU) that house airbag accelerometers must be less than 50 ppm, and down to a few ppm for ASICs. Car manufacturers deploy reliable, high performance accelerometers that are relatively expensive (up to \$5 to measure lateral acceleration in ESP units, for example). Mobile phones manufacturers on the other hand tolerate failure rates of 5000 ppm. Active control of the car motion is critical to safety whereas failure in consumer applications result more in an inconvenience. Sensor manufacturers can therefore sacrifice some accuracy and reliability to lower the price.



Tri-axis accelerometers for mobile applications

At least 10 companies are working on tri-axis MEMS accelerometers, some of which were initially developed for the automotive industry. Three axes sensing is needed to fully describe three-dimensional motion, although certain applications benefit from 2-axis devices if one axis (like the horizon) is assumed fixed or for a simple 2D (tilt) game in a cell phone.

With few exceptions, accelerometers work by measuring the motion of a proof mass versus a fixed frame or reference. The main sensing approaches are capacitive (e.g. Bosch, Freescale, Kionix, oki Electric, STMicroelectronics and Analog Devices), piezoresistive (e.g. Hitachi Metals, Matsushita, Fujitsu, and Hokuriku) or the less common thermal accelerometers from sole proponent MEMSIC (with 2-axis and 3-axis in development). Each has its advantages but price is the bottom line.

Requirements for cell phones

Mobile phone integrators want the following attributes from accelerometers:

- Small packages, e.g. $3.5 \times 3.5 \text{ mm}^2$ and 1 mm thick or less
- Sensor and signal conditioning asic in the same package
- Digital output
- Power consumption < 3 mW (1 mA @ 3V). Many devices feature a microwatt-level sleep mode to reduce power consumption further
- Compatibility with lead-free soldering processes
- Availability in volume for \$1.35 or less in next 2-3 years.

These conditions have already been met. The smallest accelerometer to date is made by Hitachi Metals, a 3-axis device in a 12 mm^3 ($3.4 \times 3.7 \times 0.92$) package at least 50% smaller than its nearest competitor.

The lowest cost 3-axis sensors come from Analog Devices or Freescale (\$2 in volume). MEMSIC provides the lowest cost solution for 2-axis sensors (now under \$1 and \$0.50 expected in 2007). Meanwhile, the

power consumption prize goes to Analog Devices' ADXL330 with power consumption well under 1 mW.

Gyroscopes improve picture quality

Today, gyroscopes are used in image stabilization, where two axes of motion sensing are needed to remove vibration and shake. While other approaches are also used to damp vibration, gyros are a superior solution.

Gyroscopes have been a feature of the top range of Panasonic camcorders for some years. They are now also employed in some Pentax and Panasonic digital still cameras costing as little as \$250. Interestingly, MEMS gyroscopes are also used for HDD free-fall protection in recent camcorders from JVC and Toshiba (these camcorders retail for over \$1,000).

Most commercial gyros are vibratory and use the transfer of energy between two vibration modes caused by Coriolis force. They are challenging to manufacture (hence their higher cost compared to accelerometers), being essentially two very high performing MEMS devices in a single device, usually a self tuned resonator for drive and micro-g for the sensing. Compared to automotive applications, packages in consumer electronics are about 1/10th the size (targeting 25 mm³). The resolution requirements are lower, i.e. 5-10 °/s compared to 0.1-1 °/s in cars.

A gyro is a larger device than an accelerometer, and can require as much as 10 times more power due to the higher drive requirements (5V).

There are many different sensing approaches and materials employed to make MEMS gyroscopes. Silicon micromachined capacitive gyros are produced by Analog Devices and Kionix. Piezo-ceramic devices are available from NEC Tokin and Murata, thin-film resonators deposited with PZT from Matsushita, quartz from Seiko Epson and Microcomponents (part of Swatch). These are all single axis solutions. Currently, Murata manufactures a piezoceramic single-axis solution that costs \$7-8. Sony and Invensense are examples of companies with two-axis solutions estimated to cost under \$10.

Today, the smallest packaged device is 21 mm³ – a single axis quartz piezoelectric gyroscope from Seiko Epson. At 6 mW, it is also the lowest power solution, in part due to a lower drive voltage compared to 5V drive capacitive gyroscopes.

The cell phone remains the dangling carrot for gyroscopes, but they are still too expensive for this application. Cell phone companies tell us the acceptable price is somewhere around \$3-5 for 2-axis sensing for an application like GPS navigation.

Currently it is not clear which technology or solution will win. While Sony is manufacturing a quartz 2-axis gyro, Murata is developing silicon capacitive micromachined sensing elements (in parallel to its ceramic gyros) to meet the consumer price targets. We believe the route for low cost 2-axis gyros will likely be a monolithic silicon capacitive solution. This is an important consideration to reduce package size and for integration of the ASIC.

Inertial measurement units arrive to consumer markets

There is a trend in consumer markets to integrate accelerometers and gyroscopes into inertial measurement unit (IMU) modules. This follows the military muster and recent automotive trends. In automotive, modules are partitioned to provide measurement in four degrees-of-freedom (DOF) to be distributed between navigation and vehicle dynamics (e.g. vehicles from Volvo and Fiat).

IMUs are targeting navigation applications in PDAs and mobile phones. AKM already supplies an electronic compass, albeit with a 3-axis magnetometer and mixed-signal processor in phones from NEC, Sanyo, SonyEricsson and LG. One application that has gained traction in Japan is an electronic compass that allows the calculation of the position of the constellations. AKM is working on compasses that include a 3-axis accelerometer, due in 2007. The package size is 4.8 × 4.8 × 1.3 mm³. A leading Asian supplier is also developing a navigation module, first using a 3-axis accelerometer and magnetometers, and subsequently a 2-axis gyro.

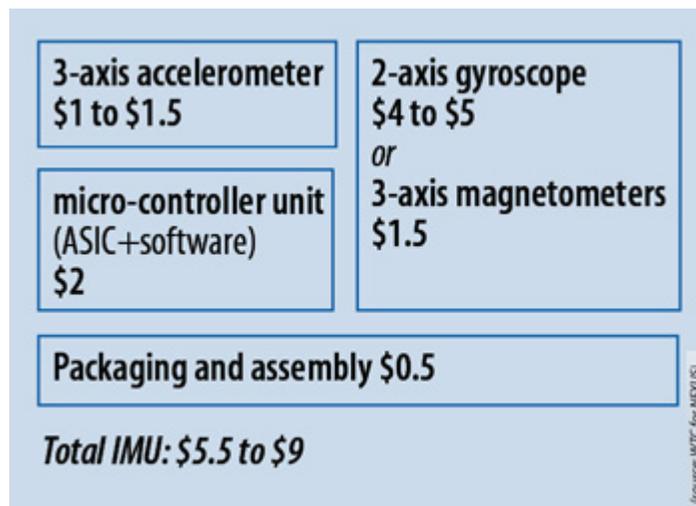
The first IMUs with MEMS sensors are estimated to cost \$20, e.g. for high-end PDAs (\$800). In 2-3 years time we believe \$6-8 is the goal,

whereby a gyro will need to cost under \$5 for 2-axes. Eventually, prices for mobile phone IMUs will not exceed \$6. An important requirement is the signal processing algorithm – which is non-trivial for five or six degrees-of-freedom. In contrast to the automotive industry, consumer integrators prefer to plug and play.

Three ways to deal with tight margins in consumer markets

Finally, let's look at how to make money in the consumer sensor market, where margins are tight.

- Prices can be pressured downwards and better margins generated when sensors are developed to be transparent to CMOS processes. An example is MEMSIC who uses fabs to produce its low cost dual axis accelerometers in high volume
- The margin for an IMU module is higher than for sensors alone.
- There are also opportunities for companies providing algorithms for IMUs. These are required to harness the power of a 5- to 6-axis functionality by turning the complex signal signatures into commands and for signal recalibration in real time.



Summary

Accelerometers have already met the requirements in terms of size, power and cost to open the markets for mobile consumer electronic. However, we believe it is will be a considerable challenge to reach the price target of \$1 and below set by cell phone integrators for 3-axis devices.

Gyroscopes remain challenging to manufacture. There is great interest in a 2-axis gyro costing under \$5. Such a price would expand the applications of gyroscopes from image stabilization in digital still cameras and camcorders to cell phone cameras, navigation imus and even further into protection of hard disc drive.

Audio MEMS

BioMEMS

Fabrication Equipment

Industry Trends

Inertial MEMS

MEMS Displays

MEMS Manufacturing

MEMS Memory

MEMS Software

MOEMS

RF MEMS

Applications of MemS:

- Aerospace: Landing gear control, flight simulators, formation flying, fly-by-wire modeling and design
- Automotive: Fuel emission control, window controllers, AC control, engine management, ABS, formation driving, steer by wire
- Military: Simulated missile trajectories, weapons control, flight simulators

- Medical and Research: Optics, remote surgery, laparoscopic surgery, clinical lab equipment
- Robotics: Teleoperation, telesurgery, tool grinding, crane control, manipulator nonlinear position and force control

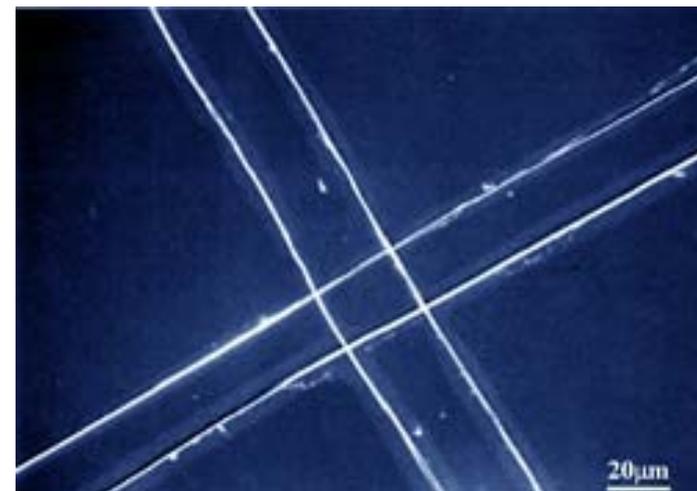
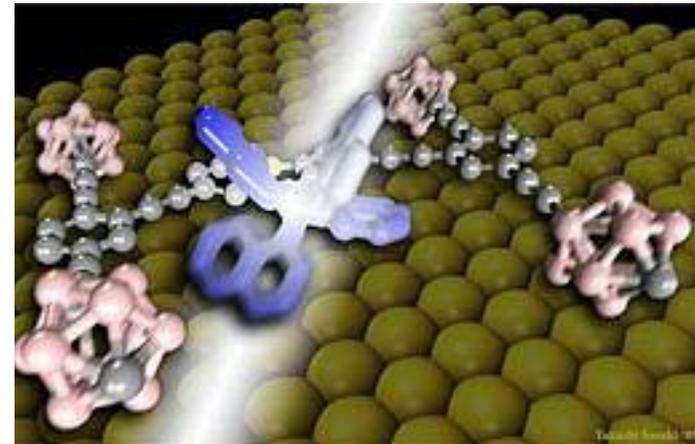
- Manufacturing: Audio filtering techniques, sound cancellation, noise control, vibration control, hard disk drive control, lens grinding, packaging, web-tension control, wafer manufacturing, product testing.

WASHINGTON, DC, April 17, 2006 (ENS)

Cars made of a single molecule now are motorized, and mechanical energy from body movement is being harnessed to generate electricity and eventually, power implantable medical devices. These are just two of the nanotechnology breakthroughs announced in the past seven days.

Nanotechnology is the science of creating or modifying materials at the atomic and molecular level to develop new or enhanced materials and products. The breakthroughs are based on work with nanometers - each one billionth of a meter in size - about one ten-thousandth the diameter of a human hair, or a thousand times smaller than a red blood cell. At Rice University in Houston, Texas, chemists who last year invented the world's first single-molecule car have now produced the first motorized version of their tiny nanocar. The research is published in the April 13 issue of the journal "Organic Letters." "We want to construct things from the bottom up, one molecule at a time, in much the same way that biological cells use enzymes to assemble proteins and other supermolecules," said lead researcher James Tour, the Chao Professor of Chemistry, professor of mechanical engineering and materials science and professor of computer science. The nanocar's light-powered motor is attached mid-chassis. When struck by light, it rotates in one direction, pushing the car along like a paddlewheel. The nanocar consists of a rigid chassis and four alkyne axles that spin freely and swivel independently of one another. The motorized model of the nanocar is powered by light. Its rotating motor, a molecular framework that was developed by Ben Feringa at the University of Groningen in the Netherlands, was modified by Tour's group so that it would attach in-line with the nanocar's chassis. When light strikes the motor, it rotates in one direction, pushing the car along like a paddlewheel.

The nanocars, which measure just 3-by-4 nanometers, are about the same width as a strand of DNA, but much shorter than DNA. About 20,000 of these nanocars could be parked, side-by-side, across the diameter of human hair. They are the first nanoscale vehicles with an internal motor. The nanocar research was funded by the National Science Foundation, the Welch Foundation, Honda, the Natural Sciences and Engineering Research Council of Canada, and Le Fonds Québécois de la Recherche sur la Nature et les Technologies.



A different nanotech advance has emerged this week from the Georgia Institute of Technology (GIT). Researchers there have developed a new technique for powering nano-scale devices without the need for bulky energy sources such as batteries: tiny nanowires that generate electricity when they vibrate. By converting mechanical energy from body movement, muscle stretching or water flow into electricity, these "nanogenerators" could make possible a new class of self-powered implantable medical devices, sensors and portable electronics.

"There is a lot of mechanical energy available in our environment," said Zhong Lin Wang, a Regents Professor in the GIT School of Materials Science and Engineering. "Our nanogenerators can convert this mechanical energy to electrical energy. This could potentially open up a lot of possibilities for the future of nanotechnology."

Described in the April 14th issue of the journal "Science," the nanogenerators produce electric current by bending and then releasing zinc oxide nanowires - which are both piezoelectric and semiconducting. The nanogenerators could also be used wherever mechanical energy - such as the hydraulic motion of seawater, wind or the motion of a foot inside a shoe - is available. The nanowires can be grown not only on crystal substrates, but also on polymer-based films. Use of flexible polymer substrates could one day allow portable devices to be powered by the movement of their users. A scanning electron microscope image shows an array of zinc oxide nanowires. Middle image shows a schematic of how an AFM tip was used to bend nanowires to produce current. Bottom image depicts output voltages produced by the array as it is scanned by the probe. "You could envision having these nanogenerators in your shoes to produce electricity as you walk. This could be beneficial to soldiers in the field, who now depend on batteries to power their electrical equipment. As long as the soldiers were moving, they could generate electricity," said Wang, who also holds affiliated faculty positions at Peking University and the National Center for Nanoscience and Technology of China. "Our bodies are good at converting chemical energy from glucose into the mechanical energy of our muscles," Wang explained. "These nanogenerators can take that mechanical energy and convert it to electrical energy for powering devices inside the body. This could open up tremendous possibilities for self-powered implantable medical devices." The research was sponsored by the National Science Foundation, the NASA Vehicle Systems Program and the Defense Advanced Research Projects Agency.

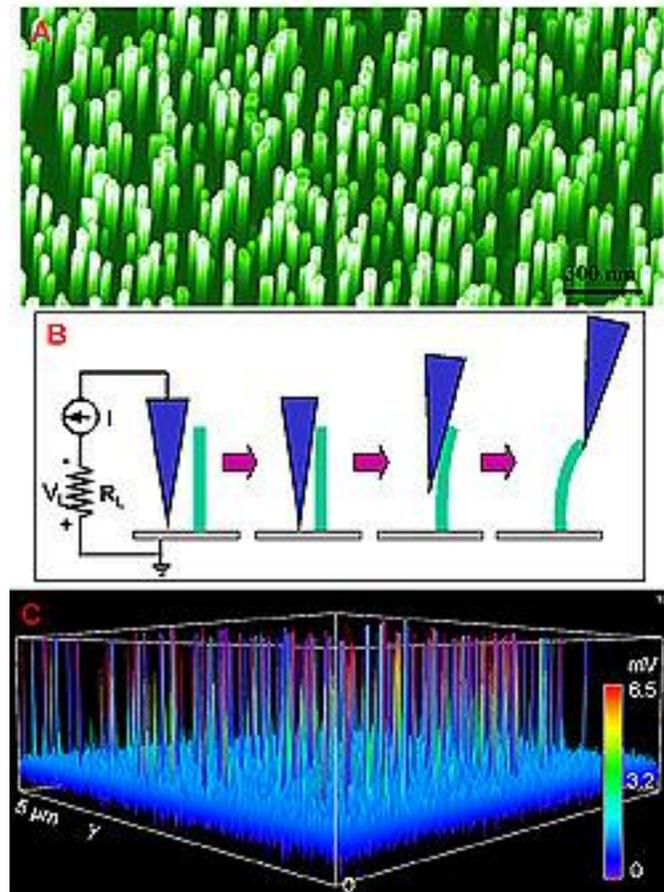
Also on the nano-medical front, in Australia, Professor Mark Kendall at the University of Queensland (UQ) has received funding to research how nanotechnology may replace syringes in administering medications. His work could eventually replace needles with tiny nano-patches on the skin. "There is an explosion of designer drugs requiring precise delivery to specific locations in the skin and we are producing new delivery methods that are practical and needle-free," Kendall said Saturday. "We are targeting immunologically sensitive cells to produce improved immune responses in the treatment of major diseases such as HIV, malaria and allergies," said Kendall. "This has enormous potential, including for the delivery of cheap and more effective vaccinations in the developing world." Professor Kendall is a UQ graduate who recently returned from the University of Oxford, where he achieved commercial success with a bioballistic gene gun. He was the associate director of the PowderJect Centre for Gene and Drug Delivery at Oxford. He is jointly appointed to UQ's Australian Institute for Bioengineering and Nanotechnology, Centre for Immunology and Cancer Research, and Faculty of Health Sciences. His work on needle-free drug delivery last week won a three year Queensland Government Smart State Senior Fellowship.

Back in the United States, researchers at University of California-Berkeley, have found a way to use the electric-field process to make nanofibers in a direct, continuous and controllable manner. Their study is in the April issue of the journal "Nano Letters." The new technique, known as near-field electrospinning, offers the possibility of producing out of nanofibers new, specialized materials with organized patterns that can be used for such applications as wound dressings, filtrations and bio-scaffolds. For 72 years, scientists have been able to use electric fields to spin polymers into tiny fibers. But the fibers tangled randomly almost as soon as they are created. In the mid-1990s, the emerging field of nanotechnology rekindled interest in electrospinning. Since then, scientists have spun more than 100 synthetic and natural polymers into fibers with diameters ranging from tens of nanometers to a few microns. A micron is one-thousandth of a millimeter. A nanometer is one-thousandth of a micron, or about the width of

10 atoms. When Daoheng Sun, a professor of mechanical and electrical engineering from China's Xiamen University came to Liwei Lin's laboratory at UC Berkeley for two years with the Berkeley Scholars Program in 2004, he and Lin, a professor of mechanical engineering, came up with the idea of trying to tame the electrospinning process to make orderly arrays of fibers.

"I'd been doing work with nanotechnology, but nothing on electrospinning before then," Lin said. "We were really outsiders in the field, so we didn't have any preconceived notions. We just tried things that others may have never thought about. And in the end, they worked just fine." These orderly rows of nanofibers were created using the new near-field electrospinning process. Until now, electrospinning produced random tangles of fibers. Lin said he foresees applications that require precise deposition of the nanofibers, such as making nanosensors for biological measurements - a glucose monitor, for instance. Another will be to make non-woven fabrics with organized patterns that can have many applications, such as scaffolds for living cells. Near-field electrospinning may also be useful in nanolithography for making next-generation microchips, Lin predicted. But, he said, this will require more effort to develop. The Berkeley Scholars Program is a privately funded program founded by the Tang Family Foundation. Today, nanomaterials are used in paints and coatings to protect against corrosion, scratches and radiation; protective and glare-reducing coatings for eyeglasses and cars; metal-cutting tools; sunscreens and cosmetics; longer-lasting tennis balls; light-weight, stronger tennis racquets; stain-free clothing and mattresses; dental-bonding agent; burn and wound dressings; ink; and automobile catalytic converters. The federal government is spending about \$1 billion per year from 2004 - 2008 to promote nanotechnology. Useful and intriguing as nanomaterials are, the U.S. Environmental Protection Agency says there may be dangers to human health and the environment. To date, EPA has funded 65 grants for more than \$22 million related to the environmental applications or implications of manufactured nanomaterials.

"This emerging field has the potential to transform environmental protection. Researchers are now testing iron nanoparticles that could clean up pollutants in large areas of groundwater cheaper and more effectively than any existing techniques," said George Gray, assistant administrator for EPA's Office of Research and Development, announcing another round of research grants March 16. "At the same time, we must understand whether nanomaterials could negatively impact health or the environment," said Gray. "This research will help determine the viability of nanotechnology as a tool for protecting our environment." Even though there are hundreds of nano products already on the market in the United States, there are currently no regulations to ensure that new nanomaterials are safe for human health and the environment. Human health research underway is testing the absorption and toxicity of nanoparticles on skin, the effects of nanoparticles on drinking water, and the effects of nanoparticles on lung tissues. Environmental researchers are investigating the impacts on marine and freshwater sediments and on aquatic bacteria, algae and plankton; and the conditions under which nanoparticles absorb or release environmental contaminants.



STMicroelectronics

ST's MEMS (Micro Electro-Mechanical Systems) products - already available or in development - target a wide array of applications in computer peripherals, telecommunications, automotive, and consumer markets. MEMS technology has become widely popular in sensors for measuring motion, acceleration, and inclination. MEMS sensors offer multiple-axis response with high resolution and low power consumption in a single package.

ST's linear accelerometers have long been used in vehicle safety systems to detect high g-levels and prompt the deployment of passenger-protecting airbags at the right moment with the right force. MEMS can also be used in 'dead-reckoning' systems to back up a GPS signal inside buildings and in urban areas, by monitoring motion and distance traveled and correcting digital-compass readings in the absence of the GPS signal.

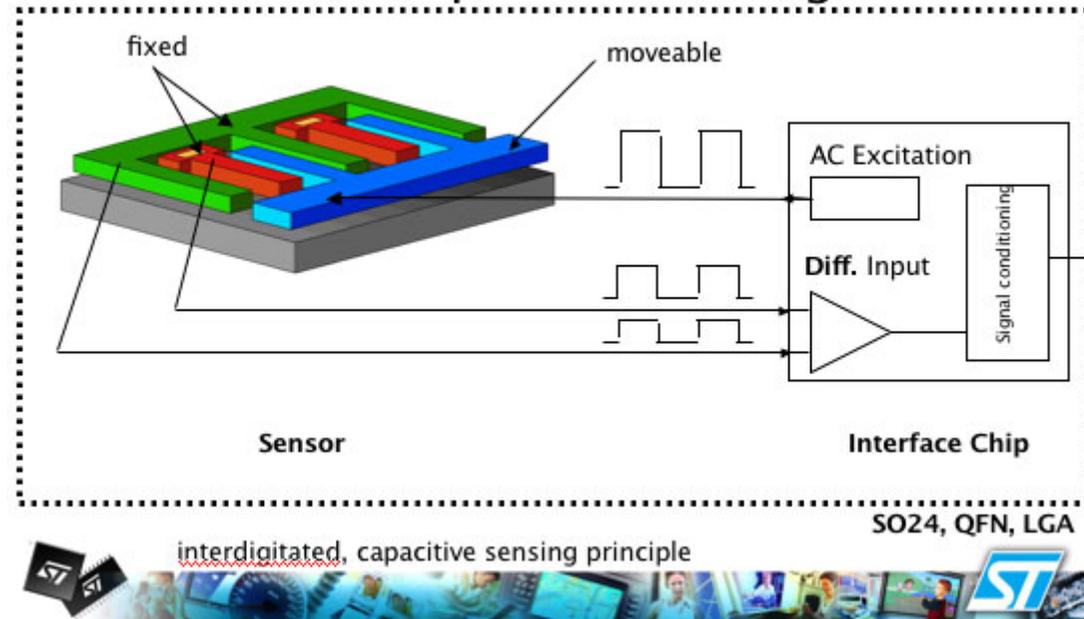
ST accelerometers are also integrated as vibration detectors in today's electronic home appliances, such as washing machines or dryers, to alert users to unbalanced loads and to protect against excessive wear of parts, before a failure occurs. ST has developed the world's first 3-axis accelerometer with analog or serial digital output, which won a prestigious Innovation of the Year Award in 2003 from EDN, the premier electronics design magazine. ST experts believe that universal all-in-one sensors, combining linear and angular accelerometers and gyroscopes with microprocessor and memory in a single package, could soon become leading products. These "Inertial Measurement Units" enable the cost-effective creation and success of various motion-activated devices, from mobile phones and PDAs to new products, based on an intuitive user interface that efficiently calibrates motion.

The high shock-resistance of MEMS devices, together with their low cost and high stability, makes MEMS an ideal solution for a variety of applications in the radio frequency (RF) circuits of wireless products, such as mobilephones and Bluetooth-connected peripherals. In MEMS-based RF filters, silicon's physical properties and the small size of the moving parts make it possible to achieve stable resonance at very high frequencies. Thanks to the extreme precision and large scale manufacturing process, MEMS components can be fabricated with very close tolerances at very competitive costs.

Silicon micromachining has been successfully applied in developing various thermofluidic MEMS devices. ST has been a long-time supplier of thermally operated inkjet print heads to Hewlett Packard, the world's leader in the printing and imaging industry. The print head contains a chip with hundreds of microscopic channels connected to corresponding miniature ink-filled chambers. The heating elements and the control circuitry, formed in the same chip, vaporize and propel the tiny ink bubbles onto the paper.

Another ingenious application of thermofluidic MEMS is demonstrated in a light-switching matrix device, developed by STMicroelectronics in cooperation with Agilent Technologies. This innovative Photonic Switching Platform uses fast temperature cycles to create (and uncreate) tiny bubbles in a transparent liquid to reflect the light beams inside an optical channel.

Two Chip - one Package



Similarly, silicon's electrical and thermal properties are utilized in ST's groundbreaking Lab-on-Chip, which is able to replace expensive and time-consuming laboratory analysis with cost-effective, disposable tools. In the 3rd quarter of 2004, STMicroelectronics launched its In-Check™ product platform, the first market-available biochip for amplifying tiny DNA samples by precisely controlled heating. Since silicon allows electronic controls on the same chip as the analyzed substance and the Lab-on-Chip heats and cools a very small sample, reaction times are faster, and the analysis more accurate and self-adjustable. MEMS miniature geometries also save costs by consuming less power and using smaller quantities of costly reagents, compared with conventional diagnostic systems.

STMicroelectronics has identified and addressed several issues that are critical for mass deployment of MEMS in a broad range of applications. In order to accelerate time-to-market and achieve economies of scale, standardizing the core technology is critically important. At the forefront of the necessary standardization efforts, the Company has developed THELMA (Thick Epitaxial Layer for Microactuators and Accelerometers), a 0.8-micron, surface micro-machining process, combining variably thick and thin polysilicon layers for structures and interconnections, respectively. Aside from THELMA, ST's goal is to establish two additional micro-machining standards, offering enough flexibility to satisfy diverse market needs.

MEMS are leaving the labs and revolutionize the industry by exploiting both the unique electrical and outstanding mechanical properties of silicon. STMicroelectronics is well-positioned to play a major role in this process, with its impressive array of MEMS products, innovative applications, and sound business cases, together with successful standardization efforts, aimed at achieving the economies of scale and winning customer confidence, critical in stimulating market demand for applications of this sophisticated (and potentially disruptive) technology.

La STMicroelectronics presenta un accelerometro su tre assi in tecnologia MEMS, con uscita digitale seriale.

Un sensore MEMS su tre assi e un chip di interfaccia riuniti in un unico package, per il mercato dei terminali portatili. La STMicroelectronics sarà presente al 3GSM World Congress che si terrà a Cannes dal 23 al 26 febbraio 2004 Stand K21, Sala4.

Ginevra, 16 febbraio 2004 - La STMicroelectronics ha annunciato oggi di aver iniziato le consegne dei primi campioni di un accelerometro su tre assi in tecnologia MEMS che, per la prima volta, riunisce in un solo package la funzione di misura dell'accelerazione lungo i tre assi (X,Y,Z) e un'uscita digitale. Il dispositivo sarà presentato allo stand della ST presente al 3GSM World Congress che si terrà a Cannes dal 23 al 26 febbraio.

Il nuovo dispositivo, LIS3L02D, è progettato soprattutto per terminali portatili e palmari, dove può essere usato con le nuove interfacce d'utente che rilevano i movimenti della mano, e possono essere fatti funzionare con una sola mano, senza l'utilizzo di apposite penne, tastierine tattili o altri tipi di interfaccia di ingresso. Altre possibili applicazioni sono nel settore dei giocattoli, apparecchiature industriali, robotica, sistemi automatizzati e apparecchiature che richiedono la rilevazione del movimento, dell'accelerazione o dell'inclinazione.

Il dispositivo LIS3L02D comprende un sensore MEMS su chip singolo, più un chip calibrato di interfaccia che è in grado di rilevare le variazioni di capacità del sensore e di tradurle in uscite digitali seriali di tipo SPI o I2C. Rispetto a un sistema composto da diversi sensori mono-asse, il nuovo componente risulta essere più semplice, meno costoso e più preciso, con costi di produzione inferiori.

Una versione più recente del LIS3L02D è stata selezionata come finalista nel Premio dell'Anno 2003 nella categoria dei componenti nella rivista EDN Magazine. Questo programma premia i migliori prodotti elettronici in circolazione tra una vasta gamma di OEM e nuovi prodotti.

Gli accelerometri basati sulla tecnologia di microlavorazione del silicio MEMS sfruttano le variazioni di capacità causate dal movimento relativo di strutture fisse e mobili, realizzate utilizzando le tecniche di lavorazione dei wafer di silicio. La STMicroelectronics offre una vasta gamma di sensori in tecnologia MEMS, fra cui sensori lineari per uno, due e tre assi. Questi dispositivi sono associati ad un chip di interfaccia che genera le uscite analogiche, PWM o digitali seriali, e sono opportunamente calibrati in fabbrica per garantire una tolleranza molto rigorosa, che non richiede ulteriori aggiustamenti nella linea di produzione del prodotto finale.



Il dispositivo LIS3L02D viene alimentato con una tensione compresa fra 2,7 e 3,6V. Ha una densita' di rumore equivalente migliore di 500 milionesimi di 'g'. Durante il trasporto e la manutenzione può supportare senza danni accelerazioni fino a 3000 g per 0.5 secondi, ed è quindi sufficientemente resistente agli urti da essere utilizzato per applicazioni di telefonia cellulare.

New 2-Axis Accelerometers from STMicroelectronics Pack Increased Performance into Smaller, Thinner Package

Geneva, September 29, 2005 - STMicroelectronics (NYSE: STM), one of the world's leading manufacturers of MEMS (Micro-Electro-Mechanical System) devices, has extended its range of dual-axis acceleration sensors with the introduction of two new devices housed in the tiny, lead-free LGA (Land Grid Array) package measuring just 5 x 5 x 1.5mm.

The LIS2L02AL and LIS2L06AL are 2-axis ultra-compact linear accelerometers that complement ST's popular existing 'low g' accelerometers, while offering tighter offset tolerances, improved temperature drift performance, and a package whose smaller footprint and height is ideally suited to applications such as mobile phones, PDA's, laptops, and PCMCIA cards characterized by slim product form factors. The LIS2L02AL provides a full-scale output range of +/-2.0g (typical), while the LIS2L06AL provides a dynamically selectable choice of +/-2.0g and +/-6.0g.

The new devices incorporate in a single package both a robust, dual-axis MEMS sensor and a CMOS interface chip that provides two simultaneous, real-time analogue outputs: one for side-to-side (lateral) and one for forward-backward (longitudinal) directions. The sensor elements, which are fabricated using a dedicated process developed and optimized by ST for the production of inertial and rotational sensors, can measure both DC accelerations (static tilt) and AC accelerations with high resolution, can withstand shocks of up to 10,000g, and incorporate a built-in self-test capability that allows the user to verify functionality of the device. The interface chip incorporates all the multiplexer, charge amplifier, demultiplexer, sample-and-hold, and reference circuits required to provide the two analog output signals, with a typical bandwidth of 2.0kHz. To maximize robustness and immunity against external disturbances, the complete signal processing chain uses a fully differential structure, with a final stage converting the differential signal into a single-ended one to be compatible with the external world.

Key benefits of the new devices include: very low noise and consequent high resolution of 0.3mg (typical); a low power consumption (typically 850µA supply current at 3.3V), which makes the devices ideal for battery-operated applications; the small package, which reduces board space and therefore the system cost; high temperature stability and low offset, which eliminate the need for calibration in most applications; the ability to adjust the bandwidth via a single external capacitor per channel to optimize system performance, and operation over a -40o to +85oC temperature range.

Accelerometers based on silicon-micromachined MEMS technology are highly cost-effective because they provide high sensitivity, robustness, and precision with the economies of scale of the high volume wafer process used to make integrated circuits. ST's 2-axis acceleration sensors incorporate a suspended silicon structure that is able to move with two degrees of freedom above a silicon substrate, with movement along each axis producing a change in capacitance that is communicated to the interface chip. The interface chip converts these capacitance changes into calibrated analog voltages proportional to the movement. As a result, board mounting is much simpler than for two separate linear sensors. In addition, the interface chip is factory-trimmed to ensure repeatable performance without production-line adjustments in the end product. The built-in Self-Test facility allows the user to verify both the mechanical and electrical parts of the device. Applying a logic signal to the Self-Test input pin causes an electric field to be established inside the device that has the same effect as a real physical acceleration, causing the sensor to move and generating a resulting change in capacitance that is detected and conditioned by the interface chip and than can be compared with the expected value.

The interface chip is factory calibrated for Sensitivity and Zero-g Offset. The trimming values are stored inside the device in a non-volatile structure. Whenever the device is turned on, the trimming parameters are downloaded into registers for use during operation. This allows the user to employ the device without further calibration.



Ideal for use in mobile terminals, gaming, and virtual-reality devices, free-fall detection for data protection in laptops and other portable equipment, inertial navigation and anti-theft systems, and a wide range of appliance, robotics, and other industrial applications, the LIS2L02AL and LIS2L06AL are available now at prices of US\$2.80 and US\$3.10, respectively, in quantities of 10,000 pieces.

Smaller, Slimmer, and Smarter - New Silicon Sensors from STMicroelectronics

Geneva, October 12, 2005 - STMicroelectronics (NYSE: STM), a leading supplier of MEMS (Micro-Electro-Mechanical System) devices, today announced that it has expanded its portfolio of three-axis accelerometers with three new models, embedding world-class performance in ever smaller packages. From ultra-compact analog devices to 'smart sensors' with two digital output formats (SPI/I2C), ST's low-power, high-resolution MEMS accelerometers address different system needs and hardware characteristics in a wide range of low-g applications, from hard-disk drive protection to motion user interfaces.

ST's [LIS3LV02DQ](#) is the market's most advanced 3-axis low-g accelerometer with a standard SPI/I2C digital interface. Its novel features, including adjustable bandwidth and direction detection, mark the transition towards the concept of "smart sensors."

Adjustable bandwidth provides optimized performance for different frequencies, raising the resolution by eliminating interference from irrelevant frequency areas. In ST's [LIS3LV02DQ](#), the bandwidth can be flexibly changed with a software command, allowing designers to effectively combine measurements of slow and fast motions, such as tilts and vibrations, in one application.

The embedded feature of direction detection immediately signals threshold crossing along any of the three axes. This information helps quickly understand in which direction the sensor is moving before any data calculations are made. Areas of application include intuitive user interfaces in various computer and consumer applications, such as 3D mice or gaming consoles.

ST's new analog-output [LIS3L02AL](#)/[LIS3L06AL](#) sensors are the market's first devices in slim-and-small plastic LGA (Land Grid Array) packages. Their 5x5x1.5 mm³ dimensions meet the space constraints of the booming small-form-factor applications, including mini disc drives, mobile phones, and other portable electronic devices.

ST's leadership in MEMS miniaturization goes hand in hand with superior performance parameters. The new accelerometers can measure both tilt and acceleration along all three (x, y, z) axes at a very low noise level and minimized power consumption, which is especially critical in battery-operated portable systems.

The LIS3L0xAL devices also offer high temperature stability and tight offset tolerances, and their robust design provides very high immunity to vibration and a shock resistance up to 10,000g.

ST's new three-axis analog sensor is available in two acceleration ranges, the standard +/- 2g [LIS3L02AL](#) for tilt applications, and a dynamically selectable choice of +/-2g and +/-6g in the [LIS3L06AL](#), suited for measuring both static and dynamic accelerations, such as vibrations or rollover motions.

"Shrinking dimensions, coupled with increasing performance, and embedded flexibility, open up a fascinating range of computer and consumer applications for our MEMS-based sensors," said Hubert Geitner, Marketing and Business Development Manager at the MEMS Business Unit of ST. "ST's silicon expertise and leadership in micro-machining technologies translate into the highest levels of sensitivity and robustness at very competitive prices. Furthermore, full ownership of the device manufacturing flow enables us to guarantee our customers long-term availability and roadmap continuity."

Available now in production volumes, unit pricing is \$4.25 for the [LIS3L02AL](#), \$4.65 for the [LIS3L06AL](#), and \$5.35 for the [LIS3LV02DQ](#), all for quantities of 10,000 pieces.



5.1. Accelerometro

I sensori sono dispositivi che si pongono all'interfaccia tra la strumentazione elettronica ed il mondo fisico. Aiutano l'elettronica a "vedere", "ascoltare", "odorare", "assaggiare" e "toccare". Convertono tipicamente quantità chimiche o fisiche non elettriche in segnali elettrici o ottici. La forza principale che ha portato allo sviluppo dei sensori è l'evoluzione della "signal processing", divenuta poco costosa, accurata ed affidabile, permettendo alle apparecchiature elettroniche di aumentare il loro grado di "intelligenza". Oggi possiamo trovare sensori più o meno ovunque, dai forni a microonde agli aspirapolvere, dai registratori alle consolle videoludiche.

Dal punto di vista della misurazione, tutte le quantità fisiche e chimiche di interesse possono essere divise in sei domini di segnali:

1. termici: temperatura, calore e flusso di calore;
2. meccanici: forza, pressione, accelerazione, velocità e posizione;
3. chimico: quantità di materia, concentrazione, composizione e velocità di reazione;
4. magnetico: intensità di campo magnetico, densità di flusso e magnetizzazione;
5. elettrico: tensione, corrente e carica;
6. segnali radianti: ovvero quantità caratterizzanti le onde elettromagnetiche come intensità, lunghezza d'onda, polarizzazione e fase.

Alcuni sensori possono inoltre generare un segnale elettrico di uscita anche senza una fonte energetica ausiliaria di alimentazione: sono detti auto-generanti o passivi, come la termocoppia. Sensori che invece necessitano di alimentazione sono detti modulanti o attivi: l'energia ausiliaria fornita è modulata dalla quantità fisica misurata; i sensori attivi hanno migliori prestazioni nella misurazione di segnali deboli, esempi includono magnetotransistori e fototransistori.

I parametri che descrivono le prestazioni dei sensori sono:

- sensibilità assoluta: rapporto tra le variazioni del segnale d'uscita e segnale da misurare;
- sensibilità relativa: quella assoluta normalizzata al valore dell'uscita con segnale nullo;
- cross sensitivity: variazione del segnale d'uscita causata da più di un segnale da misurare;
- sensibilità in dipendenza della direzione: rispetto l'angolo tra la quantità da misurare e il sensore;
- risoluzione: minima quantità misurabile;
- accuratezza: rapporto percentuale tra l'errore massimo d'uscita e la scala di output;
- errore di linearità: massima deviazione della curva di calibrazione dell'uscita dalla linea che descrive al meglio il segnale d'uscita;
- isteresi: mancanza di capacità di mostrare la stessa uscita per un dato valore della quantità da misurare in relazione alla direzione della variazione del misurando;
- offset: segnale d'uscita quando il misurando è nullo;
- rumore: segnale d'uscita casuale non riferibile al misurando;
- frequenza di cut-off: frequenza alla quale il segnale d'uscita del sensore scende sotto il 70.7% del suo massimo;
- range dinamico: distanza tra i valori massimo e minimo del misurando che il sensore può percepire;
- temperatura d'esercizio: intervallo di temperature nel quale il segnale d'uscita del sensore resta nel prefissato intervallo d'errore.

What is an accelerometer? An accelerometer is an electromechanical device that will measure acceleration forces. These forces may be static, like the constant force of gravity pulling at your feet, or they could be dynamic - caused by moving or vibrating the accelerometer.

What are accelerometers useful for? By measuring the amount of static acceleration due to gravity, you can find out the angle the device is tilted at with respect to the earth. By sensing the amount of dynamic acceleration, you can analyze the way the device is moving. At first, measuring tilt and acceleration doesn't seem all that exciting. However, engineers have come up with many ways to make really useful products using them.

An accelerometer can help your project understand its surroundings better. Is it driving uphill? Is it going to fall over when it takes another step? Is it flying horizontally or is it dive bombing your professor? A good programmer can write code to answer all of these questions using the data provided by an accelerometer. An accelerometer can help analyze problems in a car engine using vibration testing, or you could even use one to make a musical instrument. In the computing world, IBM and Apple have recently started using accelerometers in their laptops to protect hard drives from damage. If you accidentally drop the laptop, the accelerometer detects the sudden freefall, and switches the hard drive off so the heads don't crash on the platters. In a similar fashion, high g accelerometers are the industry standard way of detecting car crashes and deploying airbags at just the right time.

How do accelerometers work? There are many different ways to make an accelerometer! Some accelerometers use the piezoelectric effect - they contain microscopic crystal structures that get stressed by accelerative forces, which causes a voltage to be generated. Another way to do it is by sensing changes in capacitance. If you have two microstructures next to each other, they have a certain capacitance between them. If an accelerative force moves one of the structures, then the capacitance will change. Add some circuitry to convert from capacitance to voltage, and you will get an accelerometer. There are even more methods, including use of the piezoresistive effect, hot air bubbles, and light.

What things should I consider when buying an accelerometer? *Analog vs digital* - First and foremost, you must choose between an accelerometer with analog outputs or digital outputs. This will be determined by the hardware that you are interfacing the accelerometer with. Analog style accelerometers output a continuous voltage that is proportional to acceleration. E.g. 2.5V for 0g, 2.6V for 0.5g, 2.7V for 1g. Digital accelerometers usually use pulse width modulation (PWM) for their output. This means there will be a square wave of a certain frequency, and the amount of time the voltage is high will be proportional to the amount of acceleration.

If you are using a BASIC Stamp, or any other microcontroller with purely digital inputs, you will most likely need to go for a digital output accelerometer. The disadvantage here is that it requires you to use the timing resources of the microcontroller to measure the duty cycle, as well as performing a computationally intensive division operation. If you are using a PIC/AVR/OOPIC/Javelin with analog inputs, or a completely analog based circuit, analog is almost always the best way to go. Depending on the compiler, measuring analog acceleration can be as simple as `acceleration=read_adc();` and can be done in a few microseconds.

Number of axis - For most projects, two is enough. However, if you want to attempt 3d positioning, you will need a 3 axis accelerometer, or two 2 axis ones mounted at right angles.

Maximum swing - If you only care about measuring tilt using earth's gravity, a $\pm 1.5g$ accelerometer will be more than enough. If you are going to use the accelerometer to measure the motion of a car, plane or robot, $\pm 2g$ should give you enough headroom to work with. For a project that experiences very sudden starts or stops, you will need one that can handle $\pm 5g$ or more.

Sensitivity - Generally speaking, the more sensitivity the better. This means that for a given change in acceleration, there will be a larger change in signal. Since larger signal changes are easier to measure, you will get more accurate readings.

Bandwidth - This means the amount of times per second you can take a reliable acceleration reading. For slow moving tilt sensing applications, a bandwidth of 50Hz will probably suffice. If you intend to do vibration measurement, or control a fast moving machine, you will want a bandwidth of several hundred Hz.

Impedance/buffering issues - This is by far the single most common source of problems in projects involving analog accelerometers, because so few people thoroughly read the required documentation. Both PIC and AVR datasheets specify that for A-D conversion to work properly, the connected device must have an output impedance under 10k Ω . Unfortunately, Analog Devices' analog accelerometers have an output impedance of 32k Ω . The solution to this is to use a low input offset rail to rail op amp as a buffer to lower the output impedance. As far as we know, the DE-ACCM is the only accelerometer solution that takes care of this problem for you.

Accelerometro

Un accelerometro è uno strumento di misura in grado di rilevare e/o misurare l'accelerazione. Negli ultimi anni l'importanza di questi sensori è notevolmente accresciuta, questo perché, accanto alle tradizionali applicazioni in ambito scientifico e aerospaziale, si è sviluppato il loro uso in molti campi civili (automotive, testing, analisi meccanica, eccetera). Con il moltiplicarsi delle applicazioni, sono accresciute anche le tipologie di questi strumenti, e oggi se ne possono contare decine tipi, ognuno con caratteristiche funzionali e costruttive differenti.

Principio di funzionamento

Nella maggior parte degli accelerometri, il principio di funzionamento è il medesimo: si basa sulla rilevazione dell'inerzia di una massa quando viene sottoposta ad una accelerazione. La massa viene sospesa ad un elemento elastico, mentre un qualche tipo di sensore ne rileva lo spostamento rispetto alla struttura fissa del dispositivo. In presenza di un'accelerazione, la massa (che è dotata di una propria inerzia) si sposta dalla propria posizione di riposo in modo proporzionale all'accelerazione rilevata. Il sensore trasforma questo spostamento in un segnale elettrico acquisibile dai moderni sistemi di misura. Un prima classificazione si può così fare suddividendo questi strumenti a seconda del principio di funzionamento del sensore di posizione.

Accelerometro estensimetrico

L'accelerometro a ponte estensimetrico sfrutta come principio di rilevazione lo stesso delle celle di carico, cioè la variazione di resistenza di un estensimetro dovuta alla variazione della sua lunghezza. In questi dispositivi una massa viene sospesa su dei sottili lamierini, su quest'ultimi sono fissati degli estensimetri collegati a ponte di Wheatstone. In presenza di un'accelerazione la massa si sposta, flettendo i lamierini e conseguentemente gli estensimetri subiscono un allungamento. Con un voltmetro è possibile leggere una tensione di sbilanciamento del ponte di Wheatstone proporzionale all'accelerazione.

Accelerometro piezoresistivo

L'accelerometro a ponte piezoresistivo è una variante dell'accelerometro a ponte estensimetrico, dove al posto degli estensimetri sono utilizzati sensori piezoresistivi. Questi sensori si comportano in modo analogo agli estensimetri, ma permettono allungamenti superiori, dunque sono più adatti a realizzare accelerometri per forti accelerazioni. Spesso, in questi strumenti la massa viene sospesa su una membrana plastica, sulla quale sono stati attaccati gli elementi piezoresistivi.

Accelerometro LVDT

L'accelerometro LVDT sfrutta, come principio per la rilevazione dello spostamento della massa, un sensore LVDT integrato nella struttura dell'accelerometro stesso. In questi dispositivi, la massa stessa costituisce il nucleo ferromagnetico del sensore LVDT, e scorre (sospesa su molle o altri elementi elastici) all'interno di un canale, attorno al quale sono avvolte le bobine destinate alla rilevazione della posizione della massa. Un apposito circuito rileva la posizione del nucleo rispetto alle bobine e genera un segnale elettrico proporzionale allo spostamento rispetto alla posizione di riposo.

Accelerometro capacitivo

L'accelerometro capacitivo sfrutta, come principio per la rilevazione dello spostamento della massa, la variazione della capacità elettrica di un condensatore al variare della distanza tra le sue armature. In questi accelerometri, la massa (realizzata con materiale conduttivo) costituisce un'armatura, mentre l'altra è realizzata sulla struttura fissa del dispositivo, nell'immediata prossimità della massa. La massa viene sospesa su un elemento elastico relativamente rigido (tipicamente una membrana). Un apposito circuito rileva la capacità del condensatore così realizzato e genera un segnale elettrico proporzionale alla posizione della massa.

Accelerometro piezoelettrico

L'accelerometro piezoelettrico sfrutta, come principio per la rilevazione dello spostamento della massa, il segnale elettrico generato da un cristallo piezoelettrico quando è sottoposto ad una compressione. In questi accelerometri la massa viene sospesa sul cristallo piezoelettrico, che, in questo caso, costituisce sia il sensore, che l'elemento elastico. In presenza di un'accelerazione la massa (che presenta una certa inerzia) comprime il cristallo, il quale genera un segnale elettrico proporzionale alla compressione. Visto che l'elemento elastico è un cristallo, le caratteristiche di questi dispositivi sono peculiari:

- presentano una sensibilità relativamente bassa;
- possono rilevare accelerazioni elevatissime senza danneggiarsi (anche 1000 g);
- non possono rilevare accelerazioni continue e statiche.

L'ultima caratteristica è da rimarcare: come detto, il cristallo genera un segnale elettrico proporzionale alla compressione, ma se la compressione sul cristallo permane, il segnale generato tende a dissiparsi dopo un breve periodo. In conseguenza di ciò questi accelerometri non sono in grado di rilevare un'accelerazione continua e statica: dopo qualche secondo dall'applicazione di una tale accelerazione, il segnale si dissipa e in uscita non

sarà presente nessun segnale. Questi accelerometri trovano impiego in applicazioni dove si deve rilevare accelerazioni dinamiche come quelle che si generano nelle vibrazioni e negli shock meccanici.

Banda passante

Gli accelerometri possono essere divisi in due grandi categorie:

- per misure di accelerazione statica;
- per misure di accelerazione dinamica.

Alla prima categoria appartengono gli strumenti che presentano una banda passante con una caratteristica passa basso, mentre alla seconda appartengono gli strumenti che presentano una caratteristica passa banda. Gli accelerometri per misure di accelerazione statica sono in grado di rilevare dalle accelerazioni continue e statiche (grandezza d'ingresso con frequenza a 0 Hz) fino ad accelerazioni che variano con frequenze basse (normalmente fino a 500 Hz). Questa caratteristica è tipica degli accelerometri realizzati con il principio estensimetrico, LVDT o capacitivi. Esempi d'applicazione per questi strumenti sono misure d'accelerazione gravitazionale, d'accelerazione centrifuga, di un veicolo in movimento, nella guida inerziale.

Gli accelerometri per misure di accelerazione dinamica sono dei dispositivi che non sono in grado di rilevare accelerazioni statiche (ad esempio l'accelerazione gravitazionale), ma sono in grado di rilevare le accelerazioni che variano nel tempo, ad esempio quelle generate da oggetti che vibrano o quelle che si generano negli urti. La banda passante di questi strumenti può andare da qualche Hz a 50 kHz. Tipici accelerometri di questo tipo sono quelli realizzati con tecnologia piezoelettrica.

Condizionatori di segnale integrati

Il progresso della miniaturizzazione elettronica, ha permesso di integrare all'interno del contenitore di diversi accelerometri i relativi condizionatori di segnale. Questi circuiti possono servire ad alimentare correttamente i sensori, oppure ad amplificare, filtrare e linearizzare il segnale d'uscita. L'integrazione di questi circuiti ha semplificato l'uso degli accelerometri, che possono essere efficacemente usati in una più vasta gamma di applicazioni in modo semplice ed economico.

Accelerometro laser

L'accelerometro laser è un tipo particolare d'accelerometro, usato quando è necessario effettuare misure estremamente precise, non ottenibili con altri tipi di strumenti. Il principio di funzionamento è concettualmente differente rispetto a quelli sopra esposti, è si basa sul principio fisico che l'accelerazione è la derivata della velocità nel tempo. In questo dispositivo un interferometro laser misura istante per istante lo spostamento dell'oggetto in moto, un computer ad esso collegato effettua la derivata seconda rispetto al tempo, ottenendo così direttamente il valore d'accelerazione. I problemi di questi dispositivi è che sono costosi, piuttosto ingombranti, richiedono che l'interferometro sia montato a terra (o su un luogo da considerarsi fisso) e il laser deve essere costantemente puntato verso l'oggetto in moto.

Gravitometro

Il gravitometro è un tipo particolare d'accelerometro realizzato appositamente per misurare l'accelerazione di gravità. Secondo il principio d'equivalenza della relatività generale, gli effetti della gravità e dell'accelerazione sono gli stessi, perciò un accelerometro non può fare distinzione tra le due cause. Come gravitometri si possono usare delle versioni migliorate di accelerometri per misure statiche, in cui sono state particolarmente curate le caratteristiche di sensibilità, di precisione e di stabilità. Infatti, in questa applicazione, necessita rilevare variazioni d'accelerazione estremamente ridotte. Dove, a fini scientifici, è necessario effettuare misure estremamente precise, si ricorre ad uno strumento che lavora con lo stesso principio dell'accelerometro laser: in questo caso, si rileva l'accelerazione della caduta di un grave in una camera sottovuoto, usando un interferometro laser per misurare lo spostamento, e un orologio atomico per misurare il tempo di caduta. La rilevazione dell'accelerazione gravitazionale, oltre ad avere interesse in campo scientifico (specie in fisica e in geologia), è una pratica dell'industria estrattiva (specie per la ricerca di giacimenti petroliferi).

In biologia

Molti animali usano le statocisti, che sono speciali strutture cellulari con pesi penduli e organi di senso innervati, per misurare l'accelerazione e la gravità.

Giroscopio

Il giroscopio è un dispositivo fisico rotante che, per effetto della legge di conservazione del momento angolare, tende a mantenere il suo asse di rotazione orientato in una direzione fissa. Essenzialmente è costituito da una ruota in rotazione intorno al suo asse. Quando la ruota è in rotazione il suo asse tende a mantenersi parallelo a se stesso ed ad opporsi ad ogni tentativo di cambiare la sua orientazione. Questo meccanismo fu inventato nel 1852 dal fisico Jean Bernard Léon Foucault nell'ambito dei suoi studi sulla rotazione terrestre. Se un giroscopio è installato su una sospensione cardanica, ovvero un sistema che permette al sistema di orientarsi liberamente nelle tre direzioni dello spazio, il suo asse si manterrà orientato nella stessa direzione anche se il supporto cambia orientazione.

Impieghi

Il principio del giroscopio è sfruttato per costruire la girobussola, un dispositivo in grado di sostituire la bussola magnetica. La bicicletta è stabilizzata dalla rotazione delle ruote, che agendo come giroscopi tendono a mantenere il veicolo verticale. Il giroscopio è impiegato per stabilizzare le navi contro il rollio, gli aerei e i veicoli in generale. Nei satelliti artificiali, nelle sonde spaziali e nella navi spaziali è alla base del sistema di guida inerziale, che mantiene orientato il veicolo rispetto alle stelle fisse. In particolare nel Telescopio Spaziale Hubble è usato per mantenere puntato con precisione il telescopio verso il punto di osservazione. Il principio del giroscopio è sfruttato nello yo-yo ed è anche presente nel moto dei proiettili, infatti le canne di molte armi da fuoco presentano una rigatura interna leggermente elicoidale che imprime al proiettile una rotazione in grado di conferirgli un'azione penetrante e resistente alla forza di gravità. Negli Stati Uniti, nei primi anni 2000 è stato inventato un piccolo veicolo a due ruote mantenuto verticale grazie a un sistema di giroscopi e sistemi di retroazione sui motori, chiamato segway HT.

L'effetto giroscopico è presente come effetto collaterale in tutti i dispositivi in rapida rotazione quali i volani e gli hard disk per computer e deve essere tenuto in considerazione nella progettazione. Un giroscopio mostra una serie di fenomeni, tra cui la precessione e la nutazione.

Descrizione matematica

L'equazione fondamentale che descrive un qualunque giroscopio è:

$$\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} = \frac{d(I\boldsymbol{\omega})}{dt} = I\boldsymbol{\alpha}$$

dove:

- il vettore $\boldsymbol{\tau}$ è il momento torcente
- il vettore \mathbf{L} è il momento angolare
- il valore scalare I è il momento di inerzia
- il vettore $\boldsymbol{\omega}$ è la velocità angolare
- il vettore $\boldsymbol{\alpha}$ è l'accelerazione angolare

Dall'equazione deriva che se viene applicato un momento torcente $\boldsymbol{\tau}$ perpendicolarmente all'asse di rotazione, quindi perpendicolare ad \mathbf{L} , si sviluppa una forza perpendicolare sia a $\boldsymbol{\tau}$ che ad \mathbf{L} . Il moto che ne deriva è detto precessione. La velocità angolare del moto di precessione Ω_P , è data da:

$$\boldsymbol{\tau} = \Omega_P \times \mathbf{L}$$

La precessione può essere dimostrata sperimentalmente facendo ruotare un giroscopio con l'asse orizzontale rispetto al terreno, sostenendolo per una estremità dell'asse e liberando l'altra. Invece di cadere come ci si aspetterebbe, la ruota rimane sospesa in aria, sostenuta per un unico estremo dell'asse e l'estremità libera dell'asse descrive lentamente un cerchio sul piano orizzontale, come previsto dalla prima equazione. Il momento torcente è in questo caso prodotto dalla forza di gravità agente sul centro di massa del sistema e dalla forza di reazione che tiene sollevato il giroscopio. Come previsto dalla seconda equazione, sotto l'effetto di un momento torcente costante come quello causato dalla gravità, la velocità di precessione è inversamente proporzionale al momento angolare. Questo significa che mano a mano che il giroscopio rallenta in seguito alle perdite per attrito, la velocità di precessione aumenta, fino a che il sistema non è più in grado di sostenere il proprio peso e cade dal sostegno.

In una girobussola l'asse del giroscopio è vincolato a muoversi su un piano passante per il centro terrestre. In questo modo la rotazione terrestre genera un momento torcente sull'asse stesso che tende a ruotarlo fino a renderlo parallelo all'asse di rotazione del pianeta. Il risultato è che l'asse stesso indica sempre (a regime) la direzione nord-sud. Un problema posto dal giroscopio nell'ambito della relatività, è rispetto a quale sistema inerziale il sistema si possa definire in rotazione. Una risposta, conosciuta come principio di Mach, sostiene che il riferimento è costituito dall'insieme di tutte le masse contenute nel cosmo.

Altri tipi di giroscopi

Alternativi al giroscopio tradizionale sono i giroscopi a fibre ottiche o laser (noti come giroscopi ottici). In essi due fasci di luce sono indirizzati lungo due percorsi curvi o poligonali sul perimetro di una figura perpendicolare all'asse di cui si vuole evidenziare la rotazione. Si basano sul principio relativistico che la velocità della luce è costante in ogni riferimento inerziale. Se il sistema descritto subisce una rotazione intorno all'asse, i due raggi di luce impiegheranno tempi differenti per compiere i due percorsi, ed un interferometro potrà rilevare questa differenza. Altri sistemi impiegano sensori cristalli piezoelettrici estremamente sensibili. Tre di questi sensori disposti parallelamente ai tre assi cartesiani sono in grado di rilevare minime variazioni di orientamento. Rispetto al giroscopio meccanico tradizionale questi sistemi sono molto più sensibili e, non avendo parti in movimento, più rapidi nella risposta.

L'accelerometro Vernier

L'accelerometro Vernier, che può essere usato con CBL o con altre interfacce (LabPro, Coach, ULI...) è un sensore analogico che fornisce una tensione proporzionale alla forza esercitata su una piccola massa, sospesa nel sensore stesso. Per effetto della sua inerzia la massa, sospesa ad un supporto flessibile solidale con l'involucro del sensore, applica al supporto una forza proporzionale all'accelerazione ad essa impressa, e questo permette di utilizzare il sensore di forza come sensore di accelerazione.

Ovviamente il sensore è direzionale: quando esso è tenuto fermo con la freccia marcata sull'involucro puntata verso il basso, la massa sente nel campo gravitazionale una forza verso il basso che fornisce all'uscita una tensione il cui valore è compreso tra 0 e 5V.

Misura di accelerazioni orizzontali. Se il sistema di acquisizione dati utilizza parametri di calibrazione che fanno corrispondere alla tensione in uscita il valore dell'accelerazione di gravità (9.8 m/s^2) quando la freccia è verso il basso e il valore zero quando la freccia è orizzontale, il sensore risulta tarato come accelerometro per accelerazioni orizzontali (e va usato con la freccia in orizzontale).

Per misurare correttamente accelerazioni in moti lungo la verticale (il sensore va tenuto con la freccia in verticale) la taratura deve essere diversa: va sottratto il valore (9.8 m/s^2) che il sensore fornisce quando esso è fermo (accelerazione nulla) con la freccia rivolta verso il basso. E' bene osservare che l'accelerazione misurata ha il segno corrispondente ad un sistema di riferimento con asse diretto nella direzione opposta a quella indicata dalla freccia!

Se infatti lasciamo cadere liberamente il sensore (accelerazione pari a 9.8 m/s^2 verso il basso) l'interfaccia fornisce il valore -9.8 m/s^2 . E quando il sensore decelera bruscamente al toccar terra il segnale corrispondente è un picco positivo, ovvero una variazione positiva di velocità nel breve intervallo di tempo in cui avviene la collisione. In altri termini, dato che dopo l'urto la velocità è nulla, la velocità prima dell'urto era negativa, e quindi il sistema di riferimento è rivolto verso l'alto.

Se il sensore viene posizionato con la freccia puntata verso l'alto la taratura va modificata: il valore (-9.8 m/s^2) che il sensore fornisce quando esso è fermo (accelerazione nulla) con la freccia rivolta verso l'alto medesimo valore va ancora sottratto, ovvero va sommato il valore $+9.8 \text{ m/s}^2$.

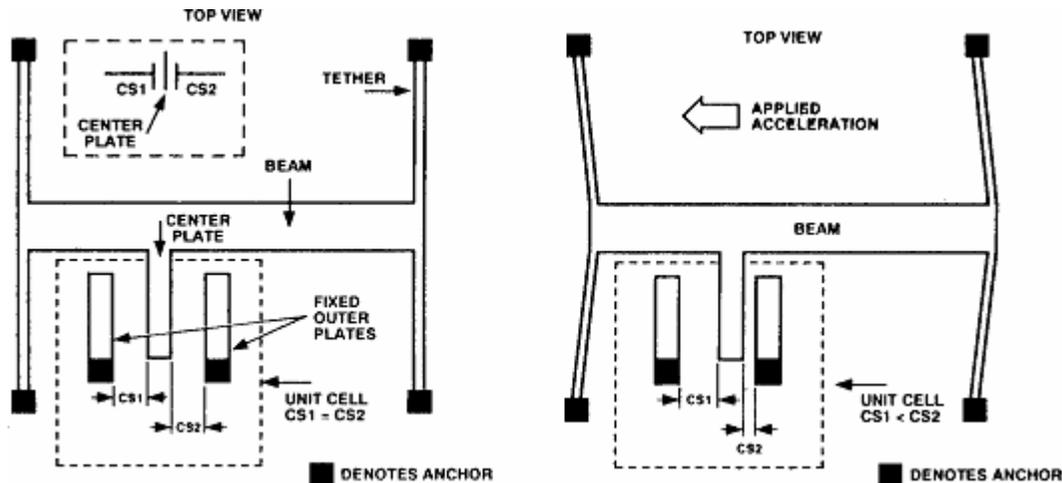
In questo caso il sistema di riferimento ha asse diretto verso il basso (sempre in direzione opposta a quella della freccia) e in caduta libera il sensore fornisce il valore $+g$ e la collisione con il terreno è registrata come picco negativo (accelerazione verso l'alto).

Un modo pratico di eseguire una corretta calibrazione per misurare accelerazioni lungo un asse verticale (una volta deciso come si vuole orientare l'asse del sistema di riferimento) è quello di fornire il valore zero quando il sensore è fermo e con la freccia diretta verticalmente, e fornire il valore $\pm g$ (a seconda della direzione scelta per l'asse) quando il sensore è posto con la freccia orizzontale (segno meno se l'asse punta in alto e segno più se l'asse punta in basso). Infatti se il sensore è fermo con la freccia orizzontale esso sente una forza nulla, la stessa che sente se si trova in caduta libera, con la freccia in qualsiasi direzione.

Note tecniche. Per CBL esistono tre tipi di accelerometri: due monoassiali, che si distinguono solo per il fondo scala (rispettivamente 5 g e 50g), e uno triassiale composto da tre distinti sensori che misurano l'accelerazione lungo tre assi ortogonali. Tutti sono ottenuti impiegando circuiti integrati prodotti dalla Analog Devices (i monoassiali impiegano gli integrati ADXL05 e ADXL50).

I sensori ADXL05 e ADXL50 sono sensori di forza. Più esattamente si tratta di una capacità variabile costituita da un elettrodo mobile (massa inerziale) posto tra due armature fisse: quindi la misura della accelerazione (in caso di movimento del sensore) ovvero della forza applicata (in caso di sensore fermo) che provoca lo spostamento dell'elettrodo centrale, si riduce ad una misura di variazione di capacità. Un oscillatore pilota, in controfase alla frequenza di 1MHz, le due capacità in serie costituite dai tre elettodi che formano il sensore.

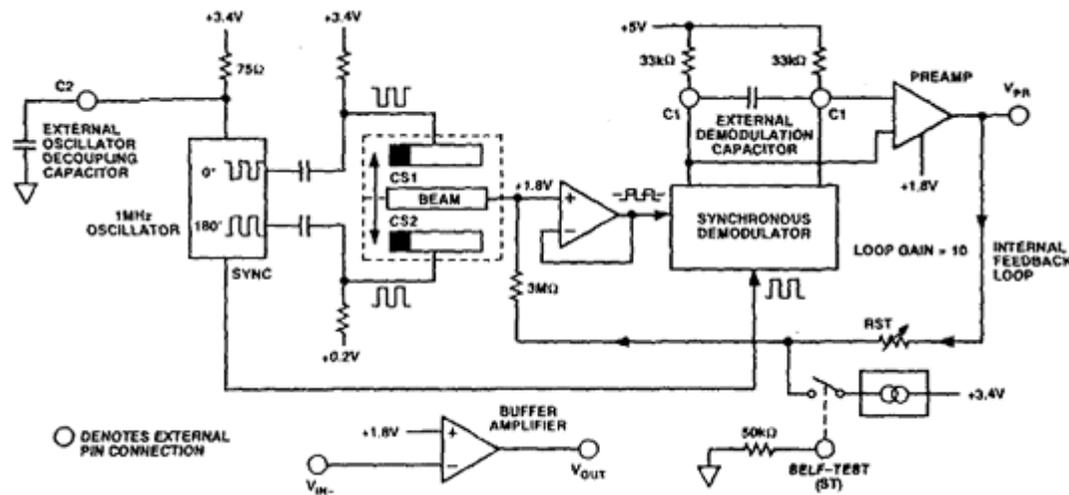
Schema dell'elemento sensibile dell'accelerometro:



Le due capacità sono ricavate mediante microlitografia da un chip di silicio ritagliando l'elettrodo centrale (l'elemento sensibile che si muove per effetto dell'accelerazione) come doppio pettine i cui denti si affacciano ai due pettini che costituiscono le altre due armature fisse. La direzione in cui l'accelerometro è sensibile è quindi l'asse di simmetria del pettine centrale, che è tenuto in posizione da 4 sottili leve alle estremità.

Il segnale prelevato all'elettrodo centrale viene rivelato da un amplificatore a sensibilità di fase (synchronous demodulator) che rimuove i segnali spuri e guida un amplificatore che insieme fornisce il segnale di uscita e forza elettrostaticamente l'elettrodo centrale nella posizione di bilanciamento delle capacità.

Schema dell'elettronica interna al sensore



Questo sistema di rivelazione del segnale e di retroazione garantisce una buona immunità al rumore e una buona linearità, e consente di discriminare tra le componenti positive o negative della accelerazione rispetto all'asse del sensore.

Accelerometers and Inclinometers

The principle of acceleration sensing is simple and reliable: mass inertia according to Newton's second law is the phenomenon that gives people a sense of acceleration. The basic elements of the accelerometer are the body, the spring and the proof mass. When the speed of the sensor body changes, the mass is forced to follow the change via the spring coupling. A force is needed to change the motion of the proof mass. Due to the force the spring is bent and the distance between the body and the proof mass changes in proportion to the body acceleration.

The operating principles of sensors differ in how the movement between the body and the mass is detected. In a capacitive sensor the body and the mass are insulated from each other and their capacitance, or charge storage capacity, is measured. As the distance decreases, the capacitance increases and electric current travels towards the sensor; when the distance increases, the opposite occurs. The sensor converts the acceleration of the body into an electric current, charge or voltage. The superb performance of VTI's sensors based on capacitive measurement technology makes them highly suitable for detecting small changes in motion. The acceleration sensing element is made of single crystal silicon and glass. This gives the sensor exceptional reliability, unprecedented accuracy and outstanding stability in terms of time and temperature. Typically a 1 g element can withstand more than 50,000 g of acceleration (1 g = the acceleration caused by the Earth's gravity). The capacitive sensing element measures acceleration in both positive and negative directions and is sensitive to static acceleration as well as vibration.

The core of the VTI accelerometer is a symmetrical, bulk-micromachined acceleration sensing element, which has two sensing capacitors. Symmetry decreases temperature dependence and cross-axis sensitivity, and improves linearity. Hermeticity has been achieved by using anodic bonding to attach the wafers to each other. This facilitates the packaging of the element, improves reliability and enables the use of gas damping in the sensor element

New 3-axis Accelerometer Concept

VTI's three axis accelerometer concept offers a single element solution to measure the low-g acceleration, the magnitude and the direction accurately. The three axis sensor element takes advantage of proven VTI single axis accelerometer element technology by utilising bulk micromachined inertial masses and the capacitive measurement principle.

The sensor design is based on multiple proof masses suspended asymmetrically with torsion springs attached close to the surface of the silicon wafer. An inertial mass with a relatively high thickness attached in this way is sensitive in vertical and horizontal directions of the wafer. The resultant sensing direction depends on the location of the proof mass center of gravity with respect to the suspension points. The horizontal surfaces of the mass and metal electrodes on the capping wafers on both sides of the mass form a differential capacitor system. When the mass is excited, it will rotate. The movement changes the electrode distances and thus the capacitances.

The three orthogonal components of the acceleration vector can be calculated as linear combinations of the position signals of the proof mass-electrode systems. The concept enables very similar linear responses in all three measurement axes.

Accelerometri in silicio.

Se ne possono produrre migliaia in un unico wafer di silicio. L'output è dell'ordine dei mV/g e gli elementi resistivi sono sensibili alla temperatura, quindi è necessaria circuiteria addizionale che provveda alla calibrazione, alla compensazione della temperatura e che funga da interfaccia verso altri circuiti. Negli approcci capacitivi l'accelerazione, perpendicolare alla superficie del substrato, fa muovere le armature del condensatore, producendo un segnale d'uscita proporzionale all'intensità dell'accelerazione. L'uscita è in fF/g, e viene convertita in una frequenza per interfacciarsi direttamente ad un microcontrollore.

Uno dei vantaggi degli accelerometri in silicio microlavorato è che possono integrare caratteristiche di auto-test. Un campo elettrostatico applicato ad un condensatore di test, realizzato nell'accelerometro frammentando l'armatura più bassa, può dare informazioni sul buon funzionamento del dispositivo e sulla calibrazione.

Accelerometro in bulk micromachining.

Un accelerometro realizzato con tecnica "bulk micromachining" consiste in un sensore con due coperture di silicio. Il sensore ha una massa di prova quadrata con lato 3.6mm, realizzata in silicio, sospesa da quattro vincoli. Le dimensioni totali dell'accelerometro sono di 7.7 x 7.2 x 1.2 mm ed è progettato per accelerazioni da 1 a 500 g. Si utilizzano piezoresistori piazzati sui vincoli per misurare la flessione indotta dal movimento della massa

di prova sotto l'accelerazione. I piezoresistori sono realizzati mediante "ion implantation". La massa prova invece con etching anisotropico dalla parte posteriore del wafer e con etching a secco dalla parte anteriore. L'etching a secco determina anche le dimensioni finali dei vincoli.

L'asse di sensibilità è perpendicolare alla superficie del chip. Poiché il centro della massa di prova non è sullo stesso piano dei vincoli flessibili, qualsiasi accelerazione in direzione diversa da quella desiderata tende a indurre torsioni sull'asse perpendicolare. Per compensare questa sensibilità sugli altri assi, due resistori sono inseriti sopra ogni vincolo flessibile: uno sulla parte terminale della massa e l'altro sulla fine della montatura. Sono poi connessi come in un ponte di Wheatstone con i due resistori di sensing nelle gambe del ponte. Con questa configurazione, la sensibilità fuori asse è ridotta a meno dell'1% della risposta.

Il dispositivo è ricoperto da entrambi i lati: la parte cava è realizzata etchando la profondità del wafer con controllo di $\pm 1 \mu\text{m}$. I tappi superiore ed inferiore sono legati al wafer centrale poiché permettono alla massa di prova di muoversi in uno spazio isolato, proteggono il sensore durante il taglio delle parti di scarto e fungono da vincolo meccanico per la massa prova in caso di forti accelerazioni. La profondità della cavità inoltre caratterizza l'escursione del sensore. La cavità restringe il movimento dell'aria attorno alla massa, che aumenta l'escursione e fa decrescere le oscillazioni.

Accelerometro capacitivo realizzato con surface micromachining.

La struttura base è un capacitore differenziale doppio che utilizza tre strati di polisilicio. La struttura è "double pinned" poiché è ancorata al substrato ai due opposti della massa sospesa. Dispositivi simili si possono costruire con un design a cantilever, anche se ci sono significative differenze nella realizzazione: per i "double pinned" è desiderabile che la risultante delle forze di stress residuali nel film di polisilicio sia di tipo tensile, in modo da produrre una struttura piana e non distorta. Nel design a cantilever il gradiente di stress attraverso lo spessore del polisilicio è il fattore critico che deve essere minimizzato per evitare la nascita di momenti di torsione. La massa sospesa è sempre ancorata a quattro supporti, a loro volta solidali con il substrato. Il capacitore differenziale è costituito da tre livelli di polisilicio: quello inferiore è tra il primo ed il secondo livello di polisilicio, quello superiore tra il terzo ed il secondo. Il primo ed il terzo livello sono fissi: sotto l'accelerazione si muove quindi il secondo livello di polisilicio, che è la massa sospesa. Questo movimento avviene in direzione perpendicolare alla superficie del chip, e si ripercuote in un cambiamento della capacità.

Una caratteristica dell'accelerometro realizzato in questa maniera è la sua capacità di auto-test, che si ottiene facilmente partizionando il primo livello di polisilicio. L'armatura di auto-test è isolata elettricamente dall'armatura del capacitore: quando si applica un voltaggio all'armatura di auto-test si ottiene una deflessione del secondo livello di polisilicio, ovvero dell'armatura ad esso parallela che rappresenta la seconda armatura del condensatore, e quindi si ottiene un cambio di capacità che causa un cambiamento nel voltaggio d'uscita. Si può quindi calibrare questa tensione d'auto-test per farla corrispondere ad un dato livello di accelerazione.

Accelerometro piezoelettrico.

Sono sensori analogici, e non interruttori come le unità meccaniche, che si basano principalmente sull'attivazione di un contatto elettrico. L'elemento sensibile consiste in due barre piezoelettriche unite tra loro, polarizzate una inversamente all'altra. La struttura a cantilever è posizionata nella direzione del movimento del veicolo. Quando si decelera, la barra superiore si allunga e la seconda è compressa: poiché sono polarizzate in maniera opposta, il segnale raddoppia. All'interno ci sono dei blocchi che fungono da convertitore, filtro passa basso e amplificatore, necessari per ottenere un'uscita utilizzabile.

Dispositivi piezoelettrici possono soffrire di offset e derive dovute alla temperatura, poiché le capacità di auto-test sono difficili da implementare. L'impedenza d'uscita inoltre è molto alta, e ciò rende sensibile il dispositivo al carico ed alla radiazione elettromagnetica.

Il principio base è quello del materiale piezoelettrico che induce carica, oppure sviluppa al suo interno una tensione, quando è deformato da una forza. Il silicio cambia la sua resistività sotto l'azione di forze. I resistori di sensing possono essere indipendentemente regioni drogate p o n. Utilizzando silicio orientato (100), l'effetto piezoelettrico può essere descritto come:

$$E = r_0 \int \pi_x S$$

con E = campo elettrico, r_0 = resistività del silicio senza applicazione di forze, I = densità della corrente di eccitazione, π_x = coefficiente piezoresistivo dipendente dall'orientamento del cristallo, S = tensore di stress dovuto alla forza. L'anisotropia del coefficiente piezoresistivo è studiata per costruire un'approssimazione nella forma di un ponte di quattro resistori orientati. Poiché i resistori non sono posizionati nello stesso punto, π_x ed S saranno differenti per ogni resistore, generando non-linearità nella sensibilità alla forza e alla temperatura.

L'approccio capacitivo permette invece l'utilizzo di elementi di piccole dimensioni, un più ampio raggio operativo per le temperature, buona linearità. È possibile misurare cambiamenti nella capacità di 0.05 fF, anche se in un ambiente industriale il limite diventa 0.5 fF. Un accelerometro con risoluzione dell'1% avrà quindi un scala di cambiamento della capacità di 50 fF. Le tecniche per misurare la variazione di capacità richiedono un ingresso in corrente alternata, che produce un campo elettrostatico, e di conseguenza una forza attrattiva tra le armature del condensatore. Altre forze possono essere generate da flusso della corrente di misura, che produce un campo magnetico, oppure dal rumore. Il rumore di sistema genera tensione di rumore a media nulla, ma la forza elettrostatica, che è proporzionale al quadrato della tensione di rumore, non sarà nulla. Questo causa nell'accelerometro una deflessione della massa e un offset allo zero. Se l'ampiezza della tensione della forma d'onda è troppo grande, le forze elettrostatiche potrebbero anche spingere la massa contro il substrato, rendendo il dispositivo inservibile.

Sensori differenziali, che utilizzando un'armatura semovibile tra due armature fisse, fanno sì che le forze elettrostatiche si cancellino quando l'armatura di mezzo è nella posizione centrale, a riposo. L'approccio differenziale ad anello aperto aumenta la linearità del sistema e minimizza gli effetti della temperatura. In un sistema ad anello chiuso, ovvero con controeazione, la massa è mantenuta nella posizione centrale dalle forze elettrostatiche: questo si ottiene applicando un appropriato voltaggio all'armatura superiore ed inferiore per bilanciare le forze d'accelerazione della massa. Si riduce così lo sforzo al quale sono sottoposti i vincoli della massa sospesa, ottenendo un'ampia banda dinamica ed intervallo di temperature di funzionamento, con riduzione di tutti i parametri, esclusa la massa, a termini di secondo ordine.

Principalmente tre sono i sistemi per realizzare la massa sospesa:

1. cantilever (come un trampolino di piscina), ha la più alta sensibilità per una data lunghezza del braccio sospeso e, a data sensibilità, permette di realizzare un dispositivo più piccolo degli altri
2. barra di torsione (fissato con una base molto piccola rispetto alla larghezza del trampolino);
3. massa sospesa da quattro cantilevers che si dipartono dalla colonna centrale.

L'impacchettamento degli accelerometri deve soddisfare ad un numero di requisiti unico nella branca dell'assemblaggio di dispositivi micromeccanici. La principale caratteristica è provvedere alla protezione dall'intrusione di particelle e dal danneggiamento fisico, contemporaneamente all'assicurare un ambiente definito dove controllare l'escursione meccanica. Il dispositivo viene protetto già al livello del wafer, legando il substrato ad una custodia microlavorata via "anodic bonding", oppure viene sigillato ermeticamente con la tecnica "glass sealing". Si inserisce poi il tutto in un contenitore convenzionale, che deve proteggere l'elemento sensibile da materiali corrosivi che tendono a far diminuire la vita del dispositivo, e dagli stress applicati che possono causare deviazioni nell'uscita dell'accelerometro.

Sensore infrarosso basato su termopile in silicio realizzate con bulk micromachining.

Una matrice composta da un gruppo di cantilever (lunghi 3 mm, spessi 10 μm e larghi 440 μm) è vincolata ad un supporto che funge da pozzo di calore. La struttura base è realizzata con un combinazione di tecnologie standard nell'ambito dei circuiti integrati e di bulk micromachining. Un strato epitassiale di tipo n spesso 10 μm viene utilizzato per definire lo spessore dei cantilevers; in seguito una diffusione di materiale di tipo p è utilizzata per creare le strisce della termopila nella zona dei cantilever. Le termopile consistono in una combinazione di p-silicio ed alluminio. Una maschera posteriore di nitrato di silicio viene utilizzata per etching anisotropico. Etching elettrochimico crea la cavità dalla parte posteriore del wafer con l'epilayer di tipo n che funge da limite per l'etching. Per definire i cantilevers si utilizza un etching al plasma di silicio dalla parte

anteriore. Infine, le microtravi risultanti sono unite con uno strato di materiale assorbente per formare la struttura. Ogni cantilever è coperto per metà con il materiale assorbente mentre l'altra metà ospita la termopila di p-Si/Al.

Il principio operativo è quello dell'effetto Seebeck. Il drogaggio della striscia di tipo p è $35\Omega/\text{quadrato}$ per ottenere un alto coefficiente di Seebeck (0.7 mV/K). Esposto a radiazione, l'energia assorbita crea un gradiente di temperatura nel cantilever, che può essere misurato dalla matrice di termopile poiché il calore è distribuito per conduzione lungo la lunghezza della microtrave.

Switch a soglia di pressione.

L'interruttore a pressione deve presentare un'isteresi tra la posizione aperta e chiusa per prevenire cambiamenti di stato indesiderati dovuti a piccole variazioni della pressione intorno al valore prefissato per la transizione. L'interruttore è formato da un elettrodo superiore di silicio, di forma sferica e deformato plasticamente, separato da un livello di isolamento dall'elettrodo di silicio inferiore. L'elettrodo superiore si piega, per via della pressione, verso il basso formando il contatto elettrico con l'elettrodo inferiore. L'isteresi può essere realizzata con variazioni nella grandezza e nello spessore dell'elettrodo. Il processo di fabbricazione parte con la creazione di una cavità nel wafer di n-silicio. Dopo l'ossidazione e la realizzazione dei pattern, si ricopre il wafer con isolante e si deposita un sottile strato sottile di p-silicio. Questo ossido a bassa temperatura definisce l'elettrodo superiore. Durante una fase ad alta temperatura (1050°C per 1 ora in azoto) l'elettrodo superiore è plasticamente deformato per via dell'espansione dei gas intrappolati nella cavità, che è sigillata. Dopo il raffreddamento l'elettrodo superiore mantiene la sua forma sferica. Si deposita poi alluminio per contattare entrambi gli elettrodi.

L'involucro protegge i circuiti integrati dall'ambiente esterno, dallo sporco, da gas, liquidi e composti chimici aggressivi. In alcune applicazioni questi contenitori sono ulteriormente protetti da composti epossidici, strati acrilici o poliuretanic e schermi ai raggi ultravioletti. I sensori però devono in genere essere a contatto con l'ambiente, quindi il completo isolamento non è certamente una buona soluzione. Per particolari applicazioni poi i sensori richiedono contenitori compatibili con il mezzo nel quale andranno inseriti: l'approccio a questo tipo di compatibilità comincia già a livello del wafer, dove l'area attiva del semiconduttore è protetta da uno strato di passivazione (nitrato o biossido di silicio drogati con boro o fosforo) depositato alla fine del processo di fabbricazione del wafer. Bisogna anche tener conto delle proprietà meccaniche di questi materiali, come ad esempio nei sensori di pressione, nei quali il diaframma rimane non protetto per evitare offset dovuti alle micro-tensioni prodotte dai diversi coefficienti di espansione termica tra lo strato di passivazione e il silicio. I contenitori più robusti per i sensori di pressione sono quelli in acciaio inossidabile riempiti con olio siliconico che permette di trasferire la pressione al sensore di silicio: minimizzare la quantità di olio vuol dire anche minimizzare gli effetti della temperatura. Questi contenitori sono costosi e quindi rappresentano una buona parte del costo complessivo del dispositivo.

S. Sherman "A low cost monolithic accelerometer", IEDM '92, Technical Digest, IEEE, 1992.

Massa in movimento forma un elettrodo del condensatore (elettrodo a pettine), il secondo elettrodo è fisso. L'accelerazione fa cambiare la capacità. Si utilizza la tecnologia BiCMOS Integrated Circuits con processi addizionali per formare la struttura in polisilicio del sensore: la tecnica "sacrificial layer" è utilizzata per rilasciare la massa in movimento (surface micromachining).

Bulk micromachining: fabrication of three-dimensional silicon structures with exactly defined and predictable shapes.

-dry etching: etching in a gas plasma atmosphere (plasma, reactive ion)

-wet etching: etching in a liquid acid or a liquid base

-anisotropic wet etching: etch rate depends on the crystal orientation (EDP EthyleneDiamine Pyrocatechol, KOH potassium hydroxide in water 6-10 mol/l). Higher crystallographic index planes get attacked by the etchants: this allow underetching of convex corner and thus the fabrication of free standing dielectric silicon dioxide structures.

Etch Rate (ER): ER(110) > ER(100) >> ER(111) -> 60:45:1

-isotropic wet etching: uniform etch rate in all crystal directions (HF); used for silicon dioxide and metals.

Surface micromachining:

-Sacrificial layer techniques: to completely remove an underlying layer to obtain free movable structures.

Etching silicon dioxide with HF or BOE provides polysilicon structures (for example accelerometer ADXL50 from Analog Devices)



Etching aluminium with H3PO3 / HNO3 / CH3COOH provides dielectric structures (for example thermal pressure sensor from PEL).

Etch rate is expressed in $\mu\text{m}/\text{min}$ at 25°C (changes 35% for every 5°C around 25°C). The etchants must be selective.

Encapsulation techniques:

-Wafer bonding (silicon fusion, anodic bonding)

Underetching of convex corners: convex corners are underetched due the non-zero etch rate of high index planes such as (211).

Etch stop mechanism

- etching by time: medium control level for fresh etchants and low etch rates;
- p++ etch stop: KOH and EDP have significantly reduced etch rates for boron doped silicon with concentrations above $2 \cdot 10^{19} \text{cm}^{-3}$;
- electrochemical etch stop: the etch of p-type silicon stops at a reversed biased pn-junction; complicated, but perfect control;
- etch stop at oxide and nitride layers
- etch stop at (111) crystal planes

Post processing deposition of special materials increases sensor sensitivity.

- capacitive humidity sensor: spin-on of water-absorbing polyimide film increases sensitivity
- magnetic field transistor: electroplating of ferromagnetic material (permalloy) increases sensitivity;
- gasfet: deposition of Pd on transistor gate by lift-off technology provides hydrogen sensitive fet.

Packaging of piezoresistive pressure sensor:

Anodic bonding between pyrex glass and silicon, followed by soft gluing on ceramic substrate. But the stress due to thermal mismatch between sensor (silicon) and packaging (glass, glue, ceramic) leads to offsets and temperature dependence.
Silicon fusion bonding between silicon and silicon: there is no thermal mismatch, but is a high temperature process.

Accelerometro

Massa sismica sospesa ad una molla. Definiamo:

X_{in} = velocità assoluta del quadro rigido
 X_r = velocità relativa della massa sismica
 X_{abs} = velocità assoluta della massa M
 V_r = tensione d'uscita del circuito elettrico equivalente

Nel dominio delle variabili di Laplace:

$$V_r/V_{in} = (1/(sL^{-1} + G)) / (1/(sL^{-1} + G) + 1/(sC)) = s^2 * L * C / (s^2 * L * C + s * L * G + 1)$$

Ponendo $w_0^2 = 1/m * C_m$ [w_0 = pulsazione naturale] ; $E = 1/(2 * m * R_m * w_0)$ [E = fattore di smorzamento] si passa alle quantità meccaniche:

$$x_r(s)/X'_{in} = s^2 / (s^2 + 2 * E * w_0 * s + w_0^2)$$

$$x_r(s)/X''_{in} = 1 / (s^2 + 2 * E * w_0 * s + w_0^2)$$

Si è effettuato in quest'ultimo passaggio un integrale, che comporta un fattore $1/s$ a numeratore, e una derivata, che porta un fattore s al denominatore. Le ultime due equazioni descrivono l'accelerometro a massa sismica, permettendo il calcolo dell'accelerazione assoluta, della velocità assoluta e dello spostamento.

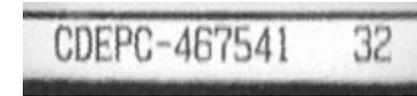
5.2. Visione artificiale

Strumento fondamentale per applicazioni di sorveglianza e sicurezza, controllo della qualità dei prodotti anche in ambiti strettamente legati alla salute dell'uomo, quali la medicina, la microchirurgia e il settore farmaceutico, i sistemi di visione artificiale trovano il loro impiego più consistente nell'industria e, in particolare, nell'automazione dei processi produttivi per verificarne il corretto funzionamento e per rilevare eventuali errori di fabbricazione dei prodotti.

La storia dell'ispezione ottica è breve ma ricca di passaggi evolutivi. Nata utilizzando barrette di sensori di luce (fotodiodi) illuminati da una lampada, che al passaggio dei pezzi da ispezionare ne determinavano la dimensione in base al numero di fotodiodi oscurati, si è rapidamente sviluppata fino ad offrire oggi fotosensori miniaturizzati selettivi ai diversi colori e incorporati in fibre ottiche per poter lavorare nelle più disparate condizioni ed essere inseriti anche in posizioni di difficile accesso nel macchinario e nelle linee di produzione. Sensori a infrarossi e a laser abbinati a microprocessori riescono a realizzare controlli dimensionali di maggior precisione e per distanze più lunghe. Inoltre l'emissione infrarossa calibrata su opportune lunghezze d'onda può riflettersi su superfici, o attraversarne altre, come tessuti, carta o plastica, permettendo così di verificare che il materiale non contenga impurità, difetti superficiali, bolle d'aria o altri difetti di finitura. Le prime telecamere utilizzate nell'industria per l'ispezione ottica erano del tipo a matrice quadrata adatta a rilevare la forma di un pezzo e quindi a verificare per confronto con un disegno l'esattezza delle quote e la completezza della lavorazione. Questo sistema viene ancora utilizzato nelle produzioni con controllo a campionatura fuori linea. Oggi il mercato mette a disposizione diversi tipi di telecamere: lineari, matriciali, multilineari e a colori, il cui impiego viene scelto in base alla necessità dell'applicazione.

Un Sistema di Visione è costituito dall'intelligenza di componenti ottiche, elettroniche e meccaniche che permettono di acquisire, archiviare, analizzare ed elaborare immagini sia della luce visibile che al di fuori di essa: infrarosso, ultravioletto, raggi X, ultrasuoni, ecc. allo scopo di riconoscere, classificare, selezionare determinate caratteristiche dell'immagine. Le principali componenti di un Sistema di Visione sono:

- l'oggetto da esaminare
- il sistema di illuminazione
- l'ottica e la telecamera
- il sistema di acquisizione e di elaborazione dell'immagine
- le interfacce uomo macchina
- le interfacce con l'ambiente esterno

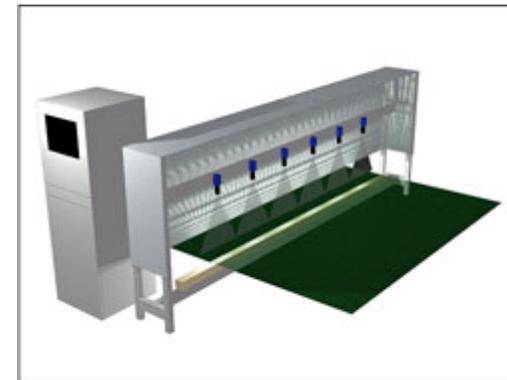


Nelle applicazioni industriali le parti da ispezionare vengono posizionate, spesso in modo automatico, di fronte a una o più telecamere ed illuminati in modo appropriato, in modo cioè da evidenziare il più possibile le caratteristiche da esaminare. Il sistema ottico proietta l'immagine sul sensore della telecamera e, il relativo segnale elettrico in uscita viene digitalizzato, convertito in forma numerica e memorizzato da una apposita scheda chiamata in gergo "frame grabber". L'immagine acquisita forma una matrice numerica che, utilizzando appositi algoritmi software, può essere elaborata da una calcolatrice, allo scopo di evidenziare quelle caratteristiche topologiche dell'immagine quali, ad esempio, contorni, spigoli, forme, strutture, che analizzate ci permettono di prendere la corretta decisione in merito alla destinazione dell'oggetto, e fornirà le informazioni opportune al resto del sistema produttivo. Per applicazioni non particolarmente complesse, sono da qualche anno disponibili sistemi compatti e

modulari e sistemi completamente integrati (Smart camera - telecamere intelligenti) che inglobano in un unico contenitore tutto il sistema di visione. Entrambe le tipologie di sistema (compatti e telecamere intelligenti) dispongono di moduli software di visione in grado di svolgere determinate funzioni, come ad esempio verifiche di posizionamento, lettura di caratteri o di codici a barre, assenza di parti ecc. Concettualmente possono essere considerati veri e propri sensori intelligenti, a metà strada fra i Sistemi di Visione veri e propri ed i sensori tradizionali. Non a caso questi prodotti sono sviluppati e commercializzati dai principali produttori di sensori

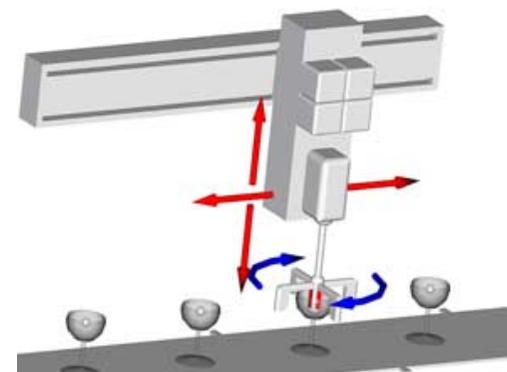
OTTICHE E SISTEMI DI ILLUMINAZIONE. L'ottica ed i sistemi di illuminazione sono componenti essenziali e allo stesso tempo critici del Sistema di Visione completo. A seconda della struttura fisica e delle dimensioni del sistema su cui il Sistema di Visione deve operare (ad esempio nastri trasportatori) e a seconda dei difetti da rilevare, ottiche ed illuminazione devono essere progettati in modo da riprendere gli oggetti e da evidenziarne le caratteristiche da controllare.

- **INTERFACCE CON IL PROCESSO PRODUTTIVO.** Nelle applicazioni industriali, il Sistema di Visione è spesso inserito in linea all'interno del processo produttivo, sia esso di tipo continuo che discreto. L'integrazione con il processo è un aspetto importante del sistema totale e riguarda ad esempio: il controllo delle fasi di movimentazione automatica, per posizionare le parti da esaminare nel raggio delle telecamere
- la sincronizzazione temporale con il processo
- i comandi alle fasi successive del processo sulla base dei controlli effettuati (ad esempio al sistema di smistamento posto a valle del Sistema di Visione)
- la gestione degli allarmi, nel caso in cui il Sistema di Visione rilevi uno scostamento del processo dai parametri normali di funzionamento
- la trasmissione dei dati relativi al controllo o l'elaborazione e la trasmissione periodica dei dati statistici



INTERFACCIA UOMO MACCHINA. Gli attuali Sistemi di Visione sono nella maggior parte dei casi completamente trasparenti all'operatore (in altre parole il personale che controlla la linea non si accorge della loro presenza, nè deve eseguire operazioni su di essi). Questo avviene sia quando la linea lavora normalmente, sia in fase di setup di produzione, il Sistema di Visione può acquisire il tipo di produzione ed i controlli da effettuare direttamente dal sistema di controllo del processo senza alcun intervento umano. In altre fasi invece, è necessaria l'interazione con l'operatore, in particolare quando:

- è necessario configurare il Sistema di Visione per nuovi tipi di produzione (fase di apprendimento)
- si verificano anomalie al processo che richiedono l'intervento umano (gestione degli allarmi impianto)
- il Sistema di Visione non opera correttamente (gestione degli allarmi e autodiagnosi)



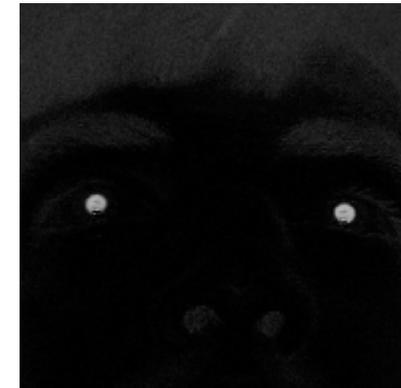
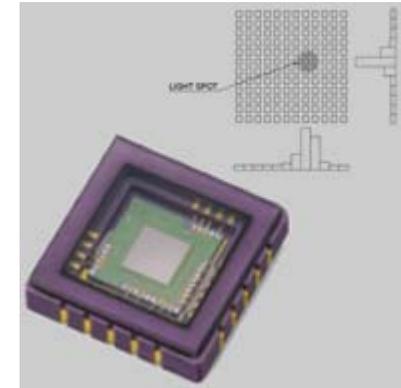
LE FUNZIONALITA'

Le funzionalità principali che i Sistemi di Visione sono in grado di assolvere sono:

- verifica scritte (Optical Character Verification)
- riconoscimento di caratteri (Optical Character Recognition)
- lettura di codici a barre
- controllo superficiale
- misure metriche
- riconoscimento e classificazione forme
- verifica presenza/assenza di parti
- verifica di colori e tonalità

High-Speed, Low-Cost Eyetracking

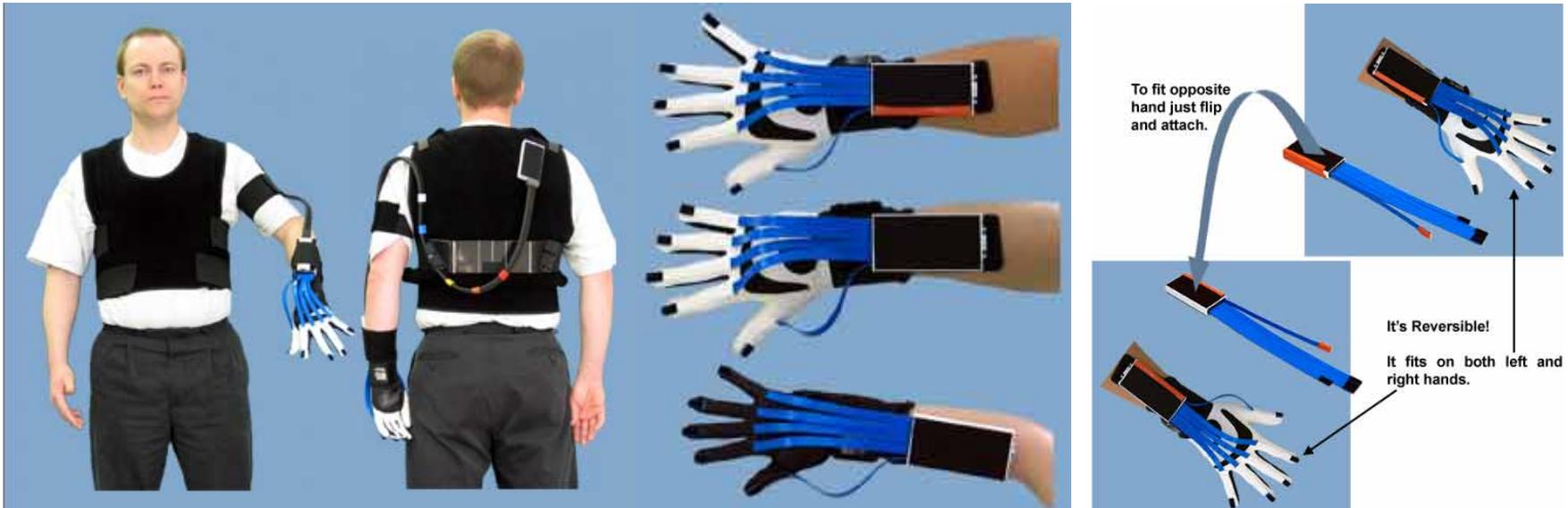
Utilizing a new generation of CMOS imaging sensors that feature on-board signal processing functionality, we are experimenting with creating a 1000fps non-invasive eye-tracker for under \$100.



5.3. Motion Capture

ShapeHand™ & ShapeHandPlus™

ShapeHand is a portable, lightweight hand motion capture system of flexible ribbons that capture hand and finger motion. ShapeHand's flexible sensors are not physically built into a glove and attach to virtually any glove and hand size. ShapeHand's glove system has two components, a data capturing component of sensors and a glove. The glove component of ShapeHand is essentially that of a leather sports glove. Gloves can be interchanged to accommodate different wearers and sizes and disposed of when desired. ShapeHand's sensor component is easily attached and detached to the glove component.



Finger & Thumb Capture

ShapeHand and ShapeHandPlus capture a full range of finger and thumb movement including abduction & adduction, .flexion & extension..



Wrist & Hand Capture



Position and Orientation Measurement

ShapeHandPlus captures complete hand and arm movements and poses including rotations and deviations.



6

capitolo



6.0. Gesture Recognition

A primary goal of virtual environments is to support natural, efficient, powerful, and flexible interaction. If the interaction technology is overly obtrusive, awkward, or constraining, the user's experience with the synthetic environment is severely degraded. If the interaction itself draws attention to the technology, rather than the task at hand, or imposes a high cognitive load on the user, it becomes a burden and an obstacle to a successful virtual environment experience.

The traditional two-dimensional, keyboard- and mouse-oriented graphical user interface (GUI) is not well-suited for virtual environments. Instead, synthetic environments provide the opportunity to utilize several different sensing modalities and technologies and integrate them into the user experience. Devices which sense body position and orientation, direction of gaze, speech and sound, facial expression, galvanic skin response, and other aspects of human behaviour or state can be used to mediate communication between the human and the environment. The cross product of communication modalities and sensing devices begets a wide range of unimodal and multimodal interface techniques. The potential for these techniques to support natural and powerful interfaces for communication in virtual environments appears promising [1].

To more fully support natural communication, we want to not only track human movement, but interpret that movement in order to recognize semantically meaningful gestures. While tracking of a user's head position or hand configuration may be quite useful for directly controlling objects or inputting parameters, people naturally express communicative acts through higher-level constructs such as speech and gesture. We want to interpret the output of position (and other) sensing to allow users to communicate more naturally and effortlessly.

Gesture is used for control and navigation in CAVEs [2][3] and in other virtual environments such as smart rooms [4], virtual work environments [5], performance spaces [6]. In addition, gesture may be perceived by the environment in order to be transmitted elsewhere, e.g., as a compression technique, to be reconstructed at the receiver. Gesture recognition may also influence - intentionally or unintentionally - a system's model of the user's state. For example, a look of frustration may cause the system to slow down its presentation of information, or the urgency of a gesture may cause the system to speed up. Gesture may also be used as a communication backchannel¹ to indicate agreement, participation, attention, conversation turn-taking, etc. Given that the human body can express a huge variety of gestures, what is appropriate to sense? Clearly the position and orientation of each body part - the parameters of an articulated body model - would be useful, as well as features that are derived from those measurements, such as velocity and acceleration. Facial expressions are very expressive. More subtle cues such as hand tension, overall muscle tension, locations of self-contact, and even pupil dilation may be of use.

Technologies to track the head, hands, and body include instrumented gloves, body suits, and marker-based optical tracking. Most of the gesture recognition work applied to virtual environments has used these tracking technologies as input. Eye-tracking devices have their limitations in tracking gaze direction. In this chapter, we cover the interpretation of tracking data from such devices in order to recognize gestures. We focus additional attention on passive sensing from cameras, using computer vision techniques to recognize gestures.

[1] Visual or verbal behaviours such as nodding or saying "uh-huh" to indicate "I'm with you, continue", or raising a finger to indicate the desire to interrupt.

6.1. The Nature of Gesture

Gestures are expressive, meaningful body motions - i.e., physical movements of the fingers, hands, arms, head, face, or body with the intent to convey information or interact with the environment. Cadoz [ref] described three functional roles of human gesture:

- Semiotic - to communicate meaningful information
- Ergotic - to manipulate the environment
- Epistemic - to discover the environment through tactile experience.

Gesture recognition is the process by which gestures made by the user are made known to the system [8]. One could argue that in GUI-based systems, standard mouse and keyboard actions used for selecting items and issuing commands are gestures; here we are interested in less trivial cases. While static position (also referred to as posture, configuration, or pose) is not technically considered gesture, we include it for the purposes of this chapter. In virtual environments, users need to communicate in a variety of ways, to the system itself and also to other users or remote environments. Communication tasks include specifying commands and/or parameters for:

- Navigating through a space
- Specifying items of interest
- Manipulating objects in the environment
- Changing object values
- Controlling virtual objects
- Issuing task-specific commands

In addition to user-initiated communication, the virtual environment system may benefit from observing a user's behavior for purposes such as:

- analysis of usability
- analysis of user tasks
- monitoring of changes in a user's state
- better understanding a user's intent or emphasis
- communicating user behaviour to other users or environments.

Messages can be expressed through gesture in many ways. For example, an emotion such as sadness can be communicated through facial expression, a lowered head position, relaxed muscles, and lethargic movement. Similarly, a gesture to indicate "Stop!" can be simply a raised hand with the palm facing forward, or an exaggerated waving of both hands above the head. In general, there exists a many-to-one mapping from concept to gesture (i.e., gestures are ambiguous); there is also a many-to-one mapping from gesture to concept (i.e., gestures are not completely specified). And, like speech and handwriting, gestures vary among individuals, they vary from instance to instance for a given individual, and they are subject to the effects of co-articulation.

An interesting real-world example of the use of gestures in visual communications is a U.S. Army field manual [9] that serves as a reference and guide to commonly used visual signals, including hand and arm gestures for a variety of situations. Despite the richness and complexity of gestural communication, researchers have made progress in beginning to understand and describe the nature of gesture. Kendon [10] described a "gesture continuum" defining five different kinds of gestures:

- Gesticulation - spontaneous movements of the hands and arms that accompany speech
- Language-like gestures - gesticulation that is integrated into a spoken utterance, replacing a particular spoken word or phrase
- Pantomimes - gestures that depict objects or actions, with or without accompanying speech
- Emblems - familiar gestures such as "V for victory", "thumbs up", and assorted rude gestures (these are often culturally specific)
- Sign languages - Linguistic systems, such as American Sign Language, which are well defined.

As the list progresses (moving from left to right in the figure), the association with speech declines, language properties increase, spontaneity decreases, and social regulation increases.

Within the first category - spontaneous, speech-associated gesture - McNeill [11][12] defined four gesture types:

- Iconic - representational gestures depicting some feature of the object, action or event being described
- Metaphoric - gestures that represent a common metaphor, rather than the object or event directly
- Beat - small, formless gestures, often associated with word emphasis
- Deictic - pointing gestures that refer to people, objects, or events in space or time.

These types of gesture modify the content of accompanying speech and may often help to disambiguate speech - similar to the role of spoken intonation. Cassell et al. [13] describe a system which models the relationship between speech and gesture and generates interactive dialogs between three-dimensional animated characters who gesture as they speak.

These spontaneous gestures (gesticulation in Kendon's Continuum) make up some 90% of human gestures. People even gesture when they are on the telephone, and blind people regularly gesture when speaking to one another. Across cultures, speech-associated gesture is natural and common. For human-computer interface to be truly natural, we need to develop technology to understand both speech and gesture together. Despite the importance of this type of gesture in normal human-to-human interaction, most research to date in HCI, and most virtual environment technology, focuses on the right side of Figure 2, where gestures tend to be less ambiguous, less spontaneous and natural, more learned, and more culture-specific. Emblematic gestures and gestural languages, although perhaps less spontaneous and natural, carry more clear semantic meaning and may be more appropriate for the kinds of command-and-control interaction that virtual environments tend to support. The main exception to this is work in recognizing and integrating deictic (mainly pointing) gestures, beginning with the well-known Put That There system by Bolt [14]. The remainder of this chapter will focus on symbolic gestures (which includes emblematic gestures and predefined gesture languages) and deictic gestures.

6.2. Representations of Gesture

The concept of gesture is loosely defined, and depends on the context of the interaction. Recognition of natural, continuous gestures requires temporally segmenting gestures. Automatically segmenting gestures is difficult, and is often finessed or ignored in current systems by requiring a starting position in time and/or space. Similar to this is the problem of distinguishing intentional gestures from other "random" movements. There is no standard way to do gesture recognition - a variety of representations and classification schemes are used. However, most gesture recognition systems share some common structure.

Gestures can be static, where the user assumes a certain pose or configuration, or dynamic, defined by movement. McNeill [11] defines three phases of a dynamic gesture: pre-stroke, stroke, and post-stroke. Some gestures have both static and dynamic elements, where the pose is important in one or more of the gesture phases; this is particularly relevant in sign languages. When gestures are produced continuously, each gesture is affected by the gesture that preceded it, and possibly by the gesture that follows it. These co-articulation may be taken into account as a system is trained, if the representation supports it. There are several aspects of a gesture which may be relevant and therefore may need to be represented explicitly. Hummels and Stappers [15] describe four aspects of a gesture which may be important to its meaning:

- Spatial information - where it occurs, locations a gesture refers to
- Pathic information - the path which a gesture takes
- Symbolic information - the sign that a gesture makes
- Affective information - the emotional quality of a gesture

In order to infer these aspects of gesture, we first have to sense human position, configuration, and movement. This can be done directly with sensing devices such as magnetic field trackers, instrumented gloves, and data-suits, which are attached to the user, or indirectly using cameras and computer vision techniques. Each sensing technology differs along several dimensions, including accuracy, resolution, latency, range of motion, user comfort, and cost. The integration of multiple sensors in gesture recognition is a complex task, since each sensing technology varies along these dimensions. Although the output from these sensors can be used to directly control parameters such as navigation speed and direction or movement of a virtual object, we are interested here in the interpretation of sensor data to recognize gestural information. The output of initial sensor processing is a time-varying sequence of parameters describing position, velocities, and angles of the relevant body part. These should (but often do not) include a representation of uncertainty that indicates the limitations of the sensor and the processing algorithms. Recognizing gestures from these parameters is a pattern recognition task which typically involves transforming the input into the appropriate representation (feature space) and then classifying it from a database of predefined gesture representations, as shown in Figure 3. The parameters

produced by the sensors may be transformed into a global coordinate space, processed to produce sensor-independent features, or used directly in the classification step.

Because gestures are highly variable, from one person to another and from one example to another with a single person, it is essential to capture the essence of the gesture - its invariant properties - and use this to represent the gesture. Besides the choice of representation itself, a significant issue in building gesture recognition systems is how to create and update the database of known gestures. Hand-coding the gestures to be recognized only works for trivial systems; in general, a system needs to be trained through some kind of learning. As with speech recognition systems, there is often a trade-off between accuracy and generality - the more accuracy desired, the more user-specific training is required. In addition, systems may be fully trained when in use, or they may adapt over time to the current user.

Static gesture, or pose, recognition can be accomplished by a straightforward implementation, using template matching, geometric feature classification, neural networks, or other standard pattern recognition techniques to classify pose. Dynamic gesture recognition, however, requires consideration of temporal events. This is typically accomplished through the use of techniques such as time-compressing templates, dynamic time warping, hidden Markov models (HMMs), and Bayesian networks.

A nice overview of work up to 1995 in hand gesture modeling, analysis, and synthesis is presented by Huang and Pavlovic [16].

6.3. Pen-based Gesture Recognition

Recognizing gestures from two-dimensional input devices such as a pen or mouse has been considered for some time. The early Sketchpad system in 1963 [17] used light-pen gestures, for example. Some commercial systems have used pen gestures since the 1970s. There are examples of gesture recognition for document editing [18][19], for air traffic control [20], and for design tasks such as editing splines [21]. More recently, systems such as the OGI QuickSet system [22] have demonstrated the utility of pen-based gesture recognition in concert with speech recognition to control a virtual environment. QuickSet recognizes 68 pen gestures, including map symbols, editing gestures, route indicators, area indicators, and taps. Oviatt [23] has demonstrated significant benefits of using both speech and pen gestures together in certain tasks. Zeleznick et al. [24] and Landay and Myers [25] developed interfaces that recognize gestures from pen-based sketching.

There have been commercially available Personal Digital Assistants (PDAs) for several years, starting with the Apple Newton, and more recently the 3Com PalmPilot and various Windows CE devices. These PDAs perform handwriting recognition and allow users to invoke operations by various, albeit quite limited, pen gestures. Long, Landay, and Rowe [26] survey problems and benefits of these gestural interfaces and provide insight for interface designers.

Although pen-based gesture recognition is promising for many HCI environments, it presumes the availability of, and proximity to, a flat surface or screen. In virtual environments, this is often too constraining - techniques that allow the user to move around and interact in more natural ways are more compelling. The next two sections cover the two primary technologies for gesture recognition in virtual environments: instrumented gloves and vision-based interfaces.

6.4. Tracker-based Gesture Recognition

There are a number of commercially available tracking systems (covered in Chapters 7 and 8) which can be used as input to gesture recognition, primarily for tracking eye gaze, hand configuration, and overall body position. Each sensor type has its strengths and weaknesses in the context of virtual environment interaction. While eye gaze can be useful in a gestural interface, we focus here on gestures based on input from tracking the hands and the body.

Data gloves. People naturally use their hands for a wide variety of manipulation and communication tasks. Besides being quite convenient, hands are extremely dexterous and expressive, with approximately 29 degrees of freedom (including the wrist). In his comprehensive thesis on whole hand input, Sturman [27] showed that the hand can be used as a sophisticated input and control device in a wide variety of application domains,

providing real-time control of complex tasks with many degrees of freedom. He analyzed task characteristics and requirements, hand action capabilities, and device capabilities, and discussed important issues in developing whole-hand input techniques. Sturman suggested a taxonomy of whole-hand input that categorizes input techniques along two dimensions:

- Classes of hand actions: continuous or discrete
- Interpretation of hand actions: direct, mapped, or symbolic

The resulting six categories describe the styles of whole-hand input. A given interaction task, can be evaluated as to which style best suits the task. Mulder [28] presented an overview of hand gestures in human-computer interaction, discussing the classification of hand movement, standard hand gestures, and hand gesture interface design. For several years, commercial devices have been available which measure, to various degrees of precision, accuracy, and completeness, the position and configuration of the hand. These include instrumented gloves and exoskeleton devices mounted on the hand and fingers (we use the term “data glove” here to include both types). Some advantages of data gloves include:

- Direct measurement of hand and finger parameters (joint angles, 3D spatial information, wrist rotation)
- Provides data at a high sampling frequency
- Easy to use
- No line-of-sight occlusion problems
- Relatively low cost versions available
- Data is translation-independent (within the range of motion)

Disadvantages of data gloves include:

- Calibration can be difficult
- Tethered gloves reduce range of motion and comfort
- Inexpensive systems can be very noisy
- Accurate systems are expensive
- The user is forced to wear a somewhat cumbersome device

Many projects have used hand input from data gloves for “point, reach, and grab” operations or more sophisticated gestural interfaces. Latoschik and Wachsmuth [29] present a multi-agent architecture for detecting pointing gestures in a multimedia application. Väänänen and Böhm [30] developed a neural network system that recognized static gestures and allows the user to interactively teach new gestures to the system. Böhm et al. [31] extend that work to dynamic gestures using a Kohonen Feature Map (KFM) for data reduction.

Baudel and Beaudouin-Lafon [32] developed a system to provide gestural input to a computer while giving a presentation - this work included a gesture notation and a set of guidelines for designing gestural command sets. Fels and Hinton [33] used an adaptive neural network interface to translate hand gestures to speech. Kadous [34] used glove input to recognize Australian sign language; Takahashi and Kishino [35] for the Japanese Kana manual alphabet. The system of Lee and Xu [36] could learn and recognize new gestures online. Despite the fact that many, if not most, gestures involve two hands, most of the research efforts in glove-based gesture recognition use only one glove for input. The features that are used for recognition, and the degree to which dynamic gestures are considered vary quite a bit.

The HIT Lab at the University of Washington developed GloveGRASP [37], a C/C++ class library that allows software developers to add gesture recognition capabilities to SGI systems, including user-dependent training and one- or two-handed gesture recognition. A commercial version of this system is available from General Reality [38].

Body suits. It is well known that by viewing only a small number of strategically placed dots on the human body, people can easily perceive complex movement patterns such as the activities, gestures, identities, and other aspects of bodies in motion [39][40]. One way to approach the recognition of human movements and postures is to optically measure the 3D position of several such markers attached to the body and then recover the time-varying articulated structure of the body. The articulated structure may also be measured more directly by sensing joint angles and positions using electromechanical body sensors. Although some of the optical systems only require dots or small balls to be placed on top of a user’s clothing, we will refer to all of these body motion capture systems generically as “body suits”.

Body suits have advantages and disadvantages which are similar to those of data gloves: they can provide reliable data at a high sampling rate (at least for electromagnetic devices), but they are expensive and very cumbersome. Calibration is typically non-trivial. The optical systems typically use several cameras and process their data offline - their major advantage is the lack of wires and a tether.

Body suits have been used, often along with data gloves, in several gesture recognition systems. Wexelblat [41] implemented a continuous gesture analysis system using a data suit, data gloves, and an eye tracker. Data from the sensors is segmented in time (between movement and inaction), key features are extracted, motion is analyzed, and a set of special-purpose gesture recognizers look for significant changes. Marrin and Picard [42] have developed an instrumented jacket for an orchestral conductor which includes physiological monitoring to study the correlation between affect, gesture, and musical expression.

Although current optical and electromechanical tracking technologies are cumbersome and therefore contrary to the desire for more natural interfaces, it is likely that advances in sensor technology will enable a new generation of devices (including stationary field sensing devices, gloves, watches, and rings) that are just as useful as current trackers but much less obtrusive. Similarly, instrumented body suits, which are currently exceedingly cumbersome, may be displaced by sensing technologies embedded in belts, shoes, eyeglasses, and even shirts and pants [43][44][45]. While sensing technology has a long way to go to reach these ideals, passive sensing using computer vision techniques is beginning to make headway as a user-friendly interface technology.

Vision-based Gesture Recognition. The most significant disadvantage of the tracker-based systems is that they are cumbersome. This detracts from the immersive nature of a virtual environment by requiring the user to don an unnatural device that cannot easily be ignored, and which often requires significant effort to put on and calibrate. Even optical systems with markers applied to the body suffer from these shortcomings, albeit not as severely. What many have wished for is a technology that provides real-time data useful for analyzing and recognizing human motion that is passive and non-obtrusive. Computer vision techniques have the potential to meet these requirements.

Vision-based interfaces use one or more cameras to capture images, at a frame rate of 30 Hz or more, and interpret those images to produce visual features that can be used to interpret human activity and recognize gestures. Typically the camera locations are fixed in the environment, although they may also be mounted on moving platforms or on other people. For the past decade, there has been a significant amount of research in the computer vision community on detecting and recognizing faces, analyzing facial expression, extracting lip and facial motion to aid speech recognition, interpreting human activity, and recognizing particular gestures.

Unlike sensors worn on the body, vision approaches to body tracking have to contend with occlusions. From the point of view of a given camera, there are always parts of the user's body that are occluded and therefore not visible - e.g., the back side of the user is not visible when the camera is in front. More significantly, self-occlusion often prevents a full view of the fingers, hands, arms, and body from a single view. Multiple cameras can be used, but this adds correspondence and integration problems.

The occlusion problem makes full body tracking difficult, if not impossible, without a strong model of body kinematics and perhaps dynamics. However, recovering all the parameters of body motion may not be a prerequisite for gesture recognition. The fact that people can recognize gestures leads to three possible conclusions: (1) we infer the parameters that we cannot directly observe, (2) we don't need these parameters to accomplish the task, or (3) we infer some and ignore others.

It is a mistake to consider vision and tracking devices (such as data gloves and body suits) as alternative paths to the same end. Although there is overlap in what they can provide, these technologies in general produce qualitatively and quantitatively different output which enable different analysis and interpretation. For example, tracking devices can in principle detect fast and subtle movements of the fingers while a user is waving his hands, while human vision in that case will at best get a general sense of the type of finger motion. Similarly, vision can use properties like texture and color in its analysis of gesture, while tracking devices do not. From a research perspective, these observations imply that it may not be an optimal strategy to merely substitute vision at a later date into a system that was developed to use a data glove or a body suit - or vice versa. Unlike special devices which measure human position and motion, vision uses a multipurpose sensor; the same device used to recognize gestures can be used to recognize other objects in the environment and also to transmit video for teleconferencing, surveillance, and other purposes. There is a growing interest in CMOS-based cameras [46], which promise miniaturized, low cost, low power cameras integrated with

processing circuitry on a single chip. With its integrated processing, such a sensor could conceivably output motion or gesture parameters to the virtual environment. Currently, most computer vision systems for recognition look something like Figure 3. Analog cameras feed their signal into a digitizer board, or framegrabber, which may do a DMA transfer directly to host memory. Digital cameras bypass the analog-to-digital conversion and go straight to memory. There may be a preprocessing step, where images are normalized, enhanced, or transformed in some manner, and then a feature extraction step. The features - which may be any of a variety of two- or three-dimensional features, statistical properties, or estimated body parameters - are analyzed and classified as a particular gesture if appropriate. Vision-based systems for gesture recognition vary along a number of dimensions, most notably:

- Number of cameras - How many cameras are used? If more than one, are they combined early (stereo) or late (multi-view)?
- Speed and latency - Is the system real-time (i.e., fast enough, with low enough latency, to support interaction)?
- Structured environment - Are there restrictions on the background, the lighting, the speed of movement, etc.?
- User requirements - Must the user wear anything special (e.g., markers, gloves, long sleeves)? Is anything disallowed (e.g., glasses, beard, rings)?
- Primary features - What low-level features are computed (edges, regions, silhouettes, moments, histograms, etc.)?
- Two- or three-dimensional representation - Does the system construct a three-dimensional model of the body part(s), or is classification done on some other (view-based) representation?
- Representation of time - How is the temporal aspect of gesture represented and used in recognition (e.g., via a state machine, dynamic time warping, HMMs, time-compressed template)?

Head and face gestures. When people interact with one another, they use an assortment of cues from the head and face to convey information. These gestures may be intentional or unintentional, they may be the primary communication mode or backchannels, and they can span the range from extremely subtle to highly exaggerated. Some examples of head and face gestures include:

- Nodding or shaking the head
- Direction of eye gaze
- Raising the eyebrows
- Opening the mouth to speak
- Winking
- Flaring the nostrils
- Looks of surprise, happiness, disgust, anger, sadness, etc.

People display a wide range of facial expressions. Ekman and Friesen [47] developed a system called FACS for measuring facial movement and coding expression; this description forms the core representation for many facial expression analysis systems. A real-time system to recognize actions of the head and facial features was developed by Zelinsky and Heinzmann [48], who used feature template tracking in a Kalman filter framework to recognize thirteen head/face gestures. Moses et al. [49] used fast contour tracking to determine facial expression from a mouth contour. Essa and Pentland [50] used optical flow information with a physical muscle model of the face to produce accurate estimates of facial motion. This system was also used to generate spatio-temporal motion-energy templates of the whole face for each different expression - these templates were then used for expression recognition. Oliver et al. [51] describe a realtime system for tracking the face and mouth which recognized facial expressions and head movements. Otsuka and Ohya [52] model coarticulation in facial expressions and use an HMM for recognition.

Hand and arm gestures. Hand and arm gestures receive the most attention among those who study gesture - in fact, many (if not most) references to gesture recognition only consider hand and arm gestures. The vast majority of automatic recognition systems are for deictic gestures (pointing), emblematic gestures (isolated signs) and sign languages (with a limited vocabulary and syntax). Some are components of bimodal systems, integrated with speech recognition. Some produce precise hand and arm configuration while others only coarse motion.

Stark and Kohler [53] developed the ZYKLOP system for recognizing hand poses and gestures in real-time. After segmenting the hand from the background and extracting features such as shape moments and fingertip positions, the hand posture is classified. Temporal gesture recognition is then performed on the sequence of hand poses and their motion trajectory. A small number of hand poses comprises the gesture catalog, while a sequence of these makes a gesture. Similarly, Maggioni and Kämmerer [54] described the GestureComputer, which recognized both hand gestures and head movements. Other systems which recognize hand postures amidst complex visual backgrounds are reported by Weng and Cui [55] and Triesch and von der Malsburg [56].

There has been a lot of interest in creating devices to automatically interpret various sign languages to aid the deaf community. One of the first to use computer vision without requiring the user to wear anything special was built by Starner [57], who used HMMs to recognize a limited vocabulary of ASL sentences. A more recent effort which uses HMMs to recognize Sign Language of the Netherlands is described by Assan and Grobel [58]. The recognition of hand and arm gestures has been applied to entertainment applications. Freeman et al. [59] developed a real-time system to recognize hand poses using image moments and orientations histograms, and applied it to interactive video games. Cutler and Turk [60] described a system for children to play virtual instruments and interact with lifelike characters by classifying measurements based on optical flow.

Body gestures. This section includes tracking full body motion, recognizing body gestures, and recognizing human activity. Activity may be defined over a much longer period of time than what is normally considered a gesture; for example, two people meeting in an open area, stopping to talk, and then continuing on their way may be considered a recognizable activity.

Bobick [61] proposed a taxonomy of motion understanding in terms of:

- Movement - the atomic elements of motion
- Activity - a sequence of movements or static configurations
- Action - high-level description of what is happening in context

Most research to date has focused on the first two levels. The Pfinder system [62] developed at the MIT Media Lab has been used by a number of groups to do body tracking and gesture recognition. It forms a two-dimensional representation of the body, using statistical models of color and shape. The body model provides an effective interface for applications such as video games, interpretive dance, navigation, and interaction with virtual characters. Lucente et al. [63] combined Pfinder with speech recognition in an interactive environment called Visualization Space, allowing a user to manipulate virtual objects and navigate through virtual worlds. Paradiso and Sparacino [64] used Pfinder to create an interactive performance space where a dancer can generate music and graphics through their body movements - for example, hand and body gestures can trigger rhythmic and melodic changes in the music.

Systems that analyze human motion in virtual environments may be quite useful in medical rehabilitation and athletic training. For example, a system like the one developed by Boyd and Little [65] to recognize human gaits could potentially be used to evaluate rehabilitation progress. Yamamoto et al. [66] describe a system which used computer vision to analyze body motion in order to evaluate the performance of skiers.

Davis and Bobick [67] used a view-based approach by representing and recognizing human action based on "temporal templates," where a single image template captures the recent history of motion. This technique was used in the KidsRoom system [68], an interactive, immersive, narrative environment for children. A nice online description of this project can be found at <http://vismod.www.media.mit.edu/vismod/demos/kidsroom/>.

Video surveillance and monitoring of human activity has received significant attention in recent years. For example, the W4 system developed at the University of Maryland [69] tracks people and detects patterns of activity.

6.5. Suggestions for System Design

There has been little work in evaluating the utility and usability of gesture recognition systems. However, those developing gestural systems have learned a number of lessons along the way. Here we present a few guidelines, in the form of "dos and don'ts" for gestural interface designers. Do inform the user. As discussed in Section 10.2, people use different kinds of gestures for many purposes, from spontaneous gesticulation associated

with speech to structured sign languages. Similarly, gesture may play a number of different roles in a virtual environment. To make compelling use of gesture, the types of gestures allowed and what they effect must be clear to the user.

Do give the user feedback. Feedback is essential to let the user know when a gesture has been recognized. This could be inferred from the action taken by the system, when that action is obvious, or by more subtle visual or audible confirmation methods.

Do take advantage of the uniqueness of gesture. Gesture is not just a substitute for a mouse or keyboard.

Do understand the benefits and limits of the particular technology. For example, precise finger positions are better suited to data gloves than vision-based techniques. Tethers from gloves or body suits may constrain the user's movement.

Do usability testing on the system. Don't just rely on the designer's intuition.

Do avoid temporal segmentation if feasible. At least with the current state of the art, segmentation of gestures can be quite difficult.

Don't tire the user. Gesture is seldom the primary mode of communication. When a user is forced to make frequent, awkward, or precise gestures, the user can become fatigued quickly. For example, holding one's arm in the air to make repeated hand gestures becomes tiring very quickly.

Don't make the gestures to be recognized too similar. For ease of classification and to help the user.

Don't use gesture as a gimmick. If something is better done with a mouse, keyboard, speech, or some other device or mode, use it - extraneous use of gesture should be avoided.

Don't increase the user's cognitive load. Having to remember the whats, wheres, and hows of a gestural interface can make it oppressive to the user. The system's gestures should be as intuitive and simple as possible. The learning curve for a gestural interface is more difficult than for a mouse and menu interface, since it requires recall rather than just recognition among a list.

Don't require precise motion. Especially when motioning in space with no tactile feedback, it is difficult to make highly accurate or repeatable gestures.

Don't create new, unnatural gestural languages. If it is necessary to devise a new gesture language, make it as intuitive as possible.

6.6. Summary and Future Directions

Although several research efforts have been referenced in this chapter, these are just a sampling; many more have been omitted for the sake of brevity. Good sources for much of the work in gesture recognition can be found in the proceedings of the Gesture Workshops and the International Conferences on Automatic Face and Gesture Recognition.

There is still much to be done before gestural interfaces, which track and recognize human activities, can become pervasive and cost-effective for the masses. However, much progress has been made in the past decade and with the continuing march towards computers and sensors that are faster, smaller, and more ubiquitous, there is cause for optimism. As PDAs and pen-based computing continue to proliferate, pen-based two-dimensional gestures should become more common, and some of the technology will transfer to three-dimensional hand, head, and body gestural interfaces. Similarly, technology developed in surveillance and security areas will also find uses in gesture recognition for virtual environments.

Each technology discussed in this chapter has its benefits and limitations. Devices that are worn or held - pens, data gloves, body suits - are currently more advanced, as evidenced by the fact that there are many commercial products available. However, passive sensing (using cameras or other sensors) promised to be more powerful, more general, and less obtrusive than other technologies. It is likely that both camps will continue to improve and co-exist, often used together in systems, and that new sensing technologies will arise to give even more choice to virtual environment developers.

6.7. Bibliographic References

- [1] Billinghamurst, M. and Savage, J. (1996). Adding Intelligence to the Interface. In Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium (pp. 168-176). Piscataway, NJ: IEEE Press.
- [2] Cruz-Neira, D.J., Sandin, T.A., DeFanti, "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE," Computer Graphics (Proceedings of SIGGRAPH '93), ACM SIGGRAPH, August 1993, pp. 135-142.
- [3] V. Pavlovic, R. Sharma, and T. Huang, "Gestural interface to a visual computing environment for molecular biologists," Proc. Second International Conference on Automatic Face and Gesture Recognition, Killington, VT, Oct. 1996.
- [4] S. Shafer, J. Krumm, B. Brumitt, B. Meyers, M. Czerwinski, and D. Robbins, "The New EasyLiving Project at Microsoft Research," Proc. Joint DARPA/NIST Smart Spaces Workshop, Gaithersburg, Maryland, July 30-31, 1998.
- [5] W. Krüger, C. A. Bohn, B. Fröhlich, H. Schüth, W. Strauss, and G. Wesche, "The responsive workbench: A virtual work environment," IEEE Computer, 28(7), 1995.
- [6] Fels, S., Nishimoto, K., and Mase, K., Musikalscope: A Graphical Musical Instrument, IEEE Multimedia Magazine, Vol. 5, No. 3, Jul/Sep, pp. 26--35, 1998.
- [7] Cadoz, "Les réalités virtuelles," Dominos, Flammarion, 1994.
- [8] G. Baratoff and D. Searles, "Gesture Recognition," <http://www.hitl.washington.edu/sciww/EVE/I.D.2.b.GestureRecognition.html>.
- [9] "Visual Signals," U.S. Army Field Manual FM-2160, Sept. 1987, available from <http://155.217.58.58/atdls.html>.
- [10] Kendon, "Some relationships between body motion and speech," in A. W. Siegman and B. Pope (eds.), Studies in Dyadic Communication, New York, Pergamon Press, 1972.
- [11] D. McNeill, Hand and Mind: What Gestures Reveal about Thought. Chicago: University of Chicago Press, 1992.
- [12] McNeill, D. (1985). So you think gestures are nonverbal?. Psychological Review. Vol 92(3) 350-371, Jul.
- [13] J. Cassell, M. Steedman, N. Badler, C. Pelachaud, M. Stone, B. Douville, S. Prevost, and B. Achorn, "Modeling the interaction between speech and gesture," Proceedings of the Sixteenth Conference of the Cognitive Science Society, August, 1994.
- [14] Bolt, R. A. "Put-That-There": Voice and gesture at the graphics interface, Computer Graphics 14.3, 1980, 262- 270.
- [15] C. Hummels and P. Stappers, "Meaningful gestures for human computer interaction: beyond hand gestures," Proc. Third International Conference on Automatic Face and Gesture Recognition, Nara, Japan, Apr. 1998.
- [16] T. Huang and V. Pavlovic, "Hand gesture modeling, analysis, and synthesis," Proc. International Workshop on Automatic Face- and Gesture-Recognition, Zurich, June 1995.
- [17] Johnson, T. "Sketchpad III: Three Dimensional Graphical Communication with a Digital Computer," in AFIPS Spring Joint Computer Conference. 1963. 23. pp. 347-353.
- [18] G. Kurtenbach, and W. Buxton, (1991). GEdit: a testbed for editing by contiguous gesture. SIGCHI Bulletin, 23(2), 22 - 26.
- [19] D. Rubine, (1991) The Automatic Recognition of Gestures. Ph.D. Dissertation, Carnegie-Mellon University.
- [20] C. P. Mertz and P. Lecoanet, "GRIGRI: gesture recognition on interactive graphical radar image," in P. Harling and A. Edwards (eds.), Progress in Gestural Interaction: Proceedings of Gesture Workshop '96, Springer- Verlag, 1997.
- [21] T. Baudel, "A Mark-Based Interaction Paradigm for Free-Hand Drawing," Proc. ACM Symposium on User Interface Software and Technology (UIST), 1994
- [22] Cohen, P. R., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., and Clow, J. (1997). QuickSet: Multimodal interaction for distributed applications, Proceedings of the Fifth Annual International Multimodal Conference (Multimedia '97), (Seattle, WA, November 1997), ACM Press, pp 31-40.
- [23] Oviatt, S. L. Multimodal interfaces for dynamic interactive maps, Proceedings of CHI'96 Human Factors in Computing Systems. ACM Press, NY, 1996, 95-102.
- [24] Zeleznik, R.C., Herndon, K.P., and Hughes J.F. Sketch: An interface for sketching 3D scenes. Computer Graphics (Proceedings of SIGGRAPH '96), August, 1996
- [25] J.A. Landay and B.A. Myers. Interactive sketching for the early stages of user interface design. Proceedings of CHI'95, pages 43-50, 1995.
- [26] A. Long, J. Landay, and L. Rowe PDA and Gesture Uses in Practice: Insights for Designers of Pen-Based User Interfaces, Report #CSD-97-976, CS Division, EECS Department, UC Berkeley, Berkeley, CA. January, 1998.
- [27] J. Sturman, "Whole-hand Input," Ph.D. Thesis, MIT Media Laboratory, Cambridge, MA, February 1992.
- [28] A. Mulder, "Hand gestures for HCI," Technical Report 96-1, School of Kinesiology, Simon Fraser University, February 1996.
- [29] M. Latoschik and I. Wachsmuth, "Exploiting distant pointing gestures for object selection in a virtual environment," in I. Wachsmuth and M. Fröhlich (eds.), Gesture and Sign Language in Human-Computer Interaction, Proc. International Gesture Workshop, Bielefeld, Germany, Sept. 1997.
- [30] K. Väänänen and K. Böhm, "Gesture driven interaction as a human factor in virtual environments - an approach with neural networks," Proc. Virtual Reality Systems, British Computer Society, Academic Press, 1992.
- [31] K. Böhm, W. Broll, and M. Solokewicz, "Dynamic gesture recognition using neural networks; a fundament for advanced interaction construction," in S. Fisher, J. Merrit, and M. Bolan (eds.), Stereoscopic Displays and Virtual Reality Systems, SPIE Conference on Electronic Imaging Science and Technology, Vol. 2177, San Jose, CA, 1994.
- [32] T. Baudel and M. Beaudouin-Lafon, "CHARADE: remote control of objects using free-hand gestures," Communications of the ACM, vol. 36, no. 7, p. 28-35, July, 1993.
- [33] S. Fels and G. Hinton, "Glove-TalkII: An Adaptive Gesture-to-Formant Interface," CHI'95, Denver, CO, May 1995.
- [34] W. Kadous, "Computer recognition of Auslan signs with PowerGloves," Proc. Workshop on the Integration of Gesture in Language and Speech, Wilmington, DE, October 1996.
- [35] T. Takahashi and F. Kishino. Gesture coding based in experiments with a hand gesture interface device. SIGCHI Bulletin, 23(2):67-73, April 1991.
- [36] C. Lee, Y. Xu. "Online, Interactive Learning of Gestures for Human/Robot Interfaces." 1996 IEEE International Conference on Robotics and Automation, Minneapolis, MN. vol. 4.
- [37] <http://www.hitl.washington.edu/people/grof/GestureGRASP.html>
- [38] <http://www.genreality.com>

- [39] G. Johansson, "Spatio-temporal differentiation and integration in visual motion perception," *Psychological Research*, 38, pp. 379-393, 1976.
- [40] Rashid, R., "Toward a System for the Interpretation of Moving Light Display," *PAMI(2)*, No. 6, November 1980, pp. 574-581.
- [41] A. Wexelblat, "A Feature-Based Approach to Continuous-Gesture Analysis," M.S. Thesis, MIT Media Laboratory, Cambridge, MA, June 1994.
- [42] T. Marrin and R. Picard. "The Conductor's Jacket: a Testbed for Research on Gestural and Affective Expression." Presented at the XII Colloquium for Musical Informatics, in Gorizia, Italy, September 1998.
- [43] J. Paradiso, E. Hu, K. Hsiao, "The CyberShoe: A Wireless Multisensor Interface for a Dancer's Feet," *Proc. International Dance and Technology 99*, Tempe AZ, Feb. 26-28, 1999.
- [44] J. Smith, T. White, C. Dodge, D. Allport, J. Paradiso, N. Gershenfeld. Electric Field Sensing for Graphical Interfaces *IEEE Computer Graphics and Applications*, Vol. 18, No. 3, May-June 1998, pp. 54-60.
- [45] R. Fletcher, J. Levitan, J. Rosenerg, N. Gershenfeld; Application of Smart Materials to Wireless ID Tags and Remote Sensors, *Materials Research Society Fall Meeting*, 1996.
- [46] S.K. Mendis, S.E. Kemeny, R.C. Gee, B. Pain, Q. Kim, and E.R. Fossum, CMOS active pixel image sensors for highly integrated imaging systems, accepted for publication in *IEEE J. Solid-State Circuits*.
- [47] Ekman, P. and Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, Calif.: Consulting Psychologists Press.
- [48] A. Zelinsky and J. Heinzmann, "Real-time visual recognition of facial gestures for human-computer interaction," *Proc. Second International Conference on Automatic Face and Gesture Recognition*, Killington, VT, Oct. 1996.
- [49] Y. Moses, D. Reynard, and A. Blake, "Determining facial expressions in real time," *Proc. Fifth International Conference on Computer Vision*, Cambridge, MA, 1995.
- [50] Essa, I., and A. Pentland. "Coding, Analysis, Interpretation and Recognition of Facial Expressions.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 19 (7), *IEEE Computer Society Press*, July, 1997.
- [51] N. Oliver, A. Pentland, and F. Bérard, "LAFTER: Lips and face real time tracker," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Puerto Rico, June 1997.
- [52] T. Otsuka and J. Ohya, "Recognizing abruptly changing facial expressions from time-sequential face images," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, CA, June 1998.
- [53] M. Stark and M. Kohler, "Video based gesture recognition for CHI," in W. D.-Fellner (ed.), *Modeling - Virtual Worlds - Distributed Graphics*, November 1995.
- [54] C. Maggioni and B. Kämmerer, "GestureComputer - history, design and applications," in Cipolla and Pentland (eds.), *Computer Vision for Human-Machine Interaction*, Cambridge University Press, 1998.
- [55] J. Weng and Y. Cui, "Recognition of hand signs from complex backgrounds," in Cipolla and Pentland (eds.), *Computer Vision for CHI*, Cambridge University Press, 1998.
- [56] J. Triesch and C. von der Malsburg, "Robust classification of hand postures against complex backgrounds," *Proc. Second International Conference on Automatic Face and Gesture Recognition*, Killington, VT, Oct. 1996.
- [57] T. Starner and A. Pentland, "Visual recognition of American Sign Language using hidden Markov models," *Proc. International Workshop on Automatic Face- and Gesture-Recognition*, Zurich, June 1995.
- [58] M. Assan and K. Grobel, "Video-based sign language recognition using hidden Markov models," in I. Wachsmuth and M. Fröhlich (eds.), *Gesture and Sign Language in Human-Computer Interaction*, *Proc. International Gesture Workshop*, Bielefeld, Germany, Sept. 1997.
- [59] W. Freeman, K. Tanaka, J. Ohta, and K. Kyuma, "Computer vision for computer games," *Proc. Second International Conference on Automatic Face and Gesture Recognition*, Killington, VT, Oct. 1996.
- [60] R. Cutler and M. Turk, "View-based interpretation of real-time optical flow for gesture recognition," *Proc. Third International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, Apr. 1998.
- [61] A. Bobick, "Movement, activity, and action: the role of knowledge in the perception of motion," *Royal Society Workshop on Knowledge-based Vision in Man and Machine*, London, England, February, 1997.
- [62] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," *Proc. Second International Conference on Automatic Face and Gesture Recognition*, Killington, VT, Oct. 1996.
- [63] M. Lucente, G. Zwart, A. George, "Visualization Space: a testbed for deviceless multimodal user interface," *Intelligent Environments Symposium*, AAAI Spring Symposium Series, Stanford, CA, March 1998.
- [64] J. Paradiso and F. Sparacino, "Optical Tracking for Music and Dance Performance," *Fourth Conference on Optical 3-D Measurement Techniques*, Zurich Switzerland, September 29-October 2, 1997.
- [65] J. Boyd and J. Little, *Shape of Motion and the Perception of Human Gaits*, *IEEE Empirical Evaluation Methods in Computer Vision*, CVPR 98, Santa Barbara, CA, June 1998.
- [66] M. Yamamoto, T. Kondo, T. Yamagiwa, and K. Yamanaka, "Skill recognition," *Proc. Third Intern. Conf. on Automatic Face and Gesture Recognition*, Nara, Japan, Apr. 1998.
- [67] J. Davis and A. Bobick, "The representation and recognition of human movement using temporal trajectories," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Puerto Rico, June 1997.
- [68] A. Bobick, S. Intille, J. Davis, F. Baird, C. Pinhanez, L. Campbell, Y. Ivanov, A. Schütte, and A. Wilson, "The KidsRoom: a perceptually-based interactive and immersive story environment," *MIT Media Lab Perceptual Computing Section Technical Report #398*, November 1996. (Revised June 1998).
- [69] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Who? When? Where? What? A real time system for detecting and tracking people," *Proc. Third International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, Apr. 1998.
- [71] J. Coutaz and J. L. Crowley, "Interpreting human gesture with computer vision," *Workshop on Gesture at the User Interface*, CHI'95, Denver, CO, 1995.
- [72] Emering L., Boulic R., Balcisoy S., Thalmann D., *Multi-Level Modelling and Recognition of Human Actions Involving Full Body Motion*, *First ACM Conf. on Autonomous Agents'97*, Marina Del Rey, 1997.

7

capitolo



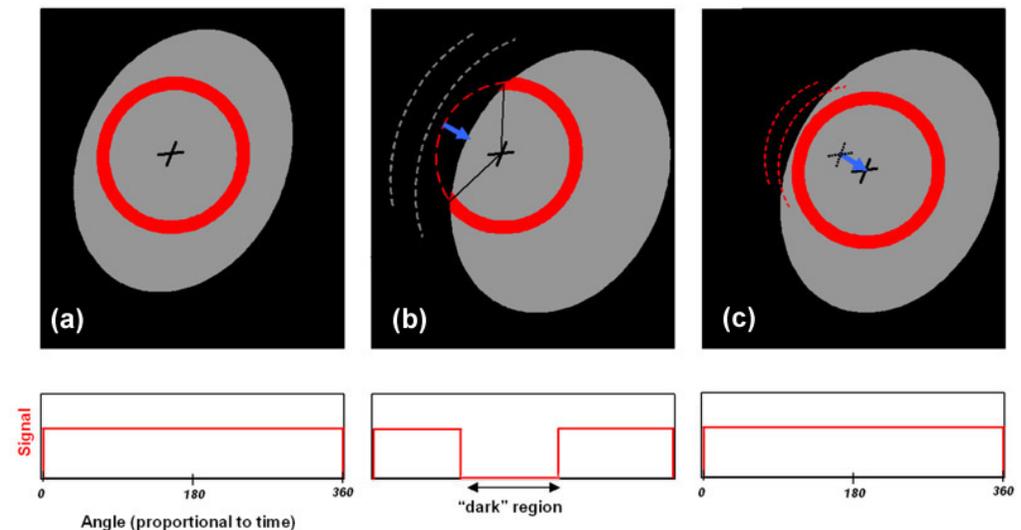
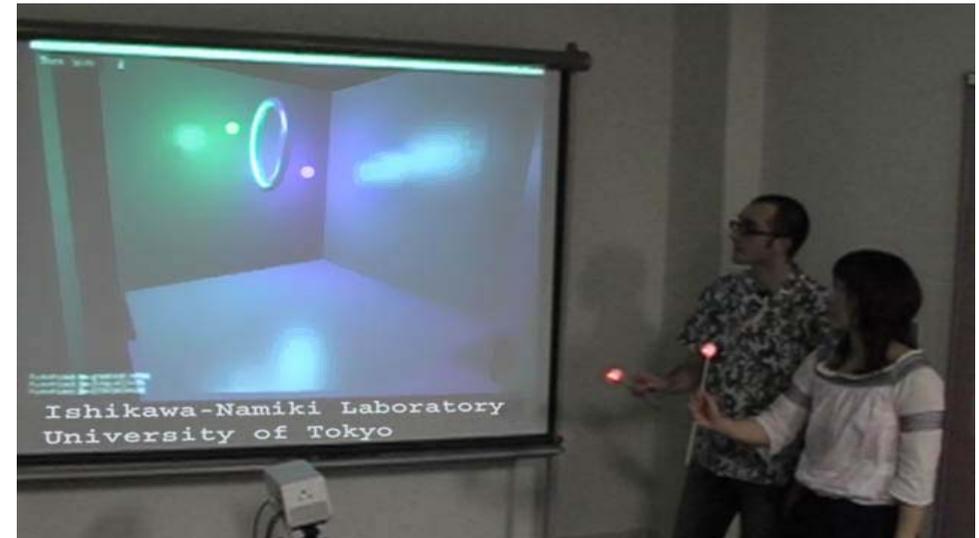
7.0. Smart Laser Scanner for Human-Computer Interface

Research description

The problem. Input of information is becoming a challenging task as portable electronic devices become smaller and smaller. Alternatives to the keyboard/mouse are necessary on handheld computers; for instance, on personal digital assistants (PDAs), input of text data is often done through a touch sensitive screen and a stylus using a prescribed input method, such as Graffiti™. The input space is thus merged with the viewing space, allowing a substantial saving of physical space. The next logical step would be to remove the need for any (dedicated or merged) input space, as well as the need for any additional input device (stylus, data-gloves, etc). This would allow inputting data by just executing bare-handed gestures in front of a portable device - that could then be embodied in a keyboardless wrist-watch. Ultimately, coupling such interface with a projective display (e.g. an eyeglass-mounted display or even a laser-based display) capable of creating an arbitrary large virtual space would completely remove the need of any real interaction area (for viewing or inputting data), allowing further downsizing of mobile computing devices (see figure).

Our approach. The problem of tracking hands and fingers on natural scenes has received much attention using passive acquisition vision systems and computationally intense image processing. We are currently studying a simple active tracking system using a laser diode (visible or invisible light), steering mirrors, and a single non-imaging photodetector, which is capable of acquiring three dimensional coordinates in real time without the need of any image processing at all. Essentially, it is a smart rangefinder scanner that instead of continuously scanning over the full field of view, restricts its scanning area, on the basis of a real-time analysis of the backscattered signal, to a very narrow window precisely the size of the target. The proposed active-tracking mechanism has some appealing features over other conventional (passive or active) tracking systems:

- Few constraints on the user/environment. The interface frees the user for adopting a special input posture; it does not require the user to carry any special device (such as active transceivers or reflective optical markers); it does not require any special projection area (other than the hand/finger itself); it is an active-illumination system, so it does not imposes stringent illumination conditions.



- Real-time, full 3D acquisition. Although depth estimation of distant targets may be difficult (full 3D measurement may be limited to about 160 cm in our present configuration), tracking and three-dimensional data acquisition is achieved without the need for a stereoscopic camera - in fact, no imaging device is needed at all. For tracking distant targets, more sophisticated telemetric techniques could be readily adopted in the system (i.e. time of flight measurement).

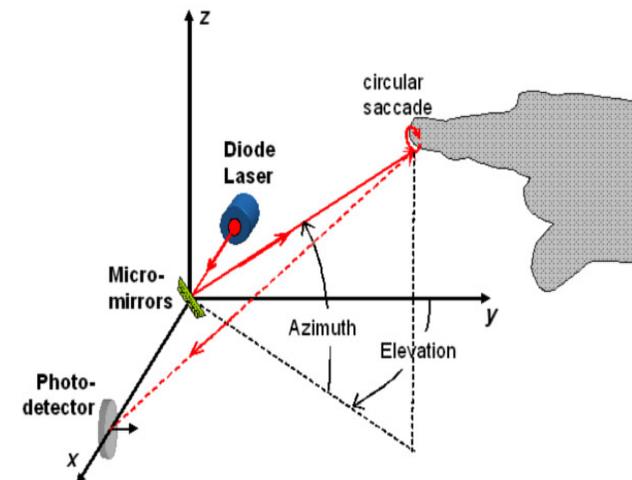
- System on-chip integration. Hardware simplicity is such that, using state-of-the-art Micro-Opto-Electro-Mechanical-System (MOEMS) technology, it should be possible to integrate the whole system on a single chip, making a versatile human-machine gestural input interface for use in mobile computing devices.

Tracking Principle

Tracking is based on the temporal analysis of the backscattered light measured during a millisecond-long circular laser saccade generated around the current position of the tracked object, thanks to a beam-steering mechanism. Adaptive saccade shapes are also under study (these would change so as to match important features of the tracked object). The figure shows a simplified, step-by-step tracking sequence.

Since the scan has a finite duration and the object cannot be expected to wait for the scan to complete, there is a systematic drift on the computed recentering vector. Thanks to elementary prediction strategies (e.g. Kalman filtering), this drift can be readily eliminated. Additional processing is required to adjust the size of the saccade as the object moves back and forth from the tracking system, or, for instance, to maintain the tracking of a fingertip without seeing the saccade slipping towards the base of the finger - but this additional processing remains elementary.

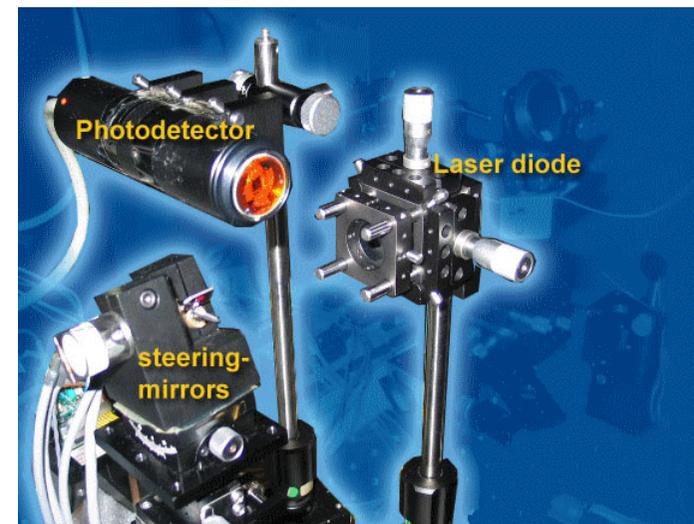
(a) The saccade is supposed to remain fully inside the object while tracking; (b) As the object moves, a small portion of the saccade may fall outside the object, and the backscattered signal will momentarily drop. Due to the synchronous operation of the beam-steering mirrors and the photodetection, an accurate recentering vector (blue arrow) is computed. (c) The center of the saccade is updated accordingly. These three steps are repeated continuously (each circular scan takes about one or two milliseconds).



Experimental Setup

Beam Steering mechanism: The beam-steering mechanism represents the critical part of the system, determining both the ultimate tracking performance and the compactness of the whole setup. In our proof-of-principle experiment, a pair of high-performance closed-loop galvano-mirrors with axes perpendicular to each other (GSI Lumonics, Model V500) are used for both generating the saccade and performing the actual tracking. The galvano-mirrors have a typical clear aperture of 5 mm and a maximum optical scan angle of $\pm 50^\circ$. Decoupling the mechanism responsible for the saccade generation from the mechanism used to actually keep track of the object position can lead to optimal tracking speeds (for instance, acousto-optic deflectors or resonant micro-mirrors can be used to generate Lissajou-like saccade patterns, while slower but precise micromirrors are used to position the center of the saccade).

Laser source: The laser source is a focusable (20 mm/inf) diode laser delivering a maximum optical power of 1 mW at a peak wavelength of 633 nm (this power/wavelength corresponds to a Class II laser source, similar to that of a barcode scanner, and in normal conditions does not represent any hazard to the eye). Since the target is likely to represent a small portion of the total field of view, the illuminating energy can be efficiently concentrated, and therefore the system will be extremely power-efficient and work equally well with distant targets. In our experiments, 1 mW was enough to produce a good signal-to-noise contrast up to two meters away from the system. Since the backscattered signal originates by isotropic diffusion/reflection on the skin surface, the necessary optical power scales as the square of the working distance; therefore, for distances less than 10cm (typical of portable devices) the required optical power would drop below 30microW. Visibility of the source greatly simplifies the calibration of the system, but in a final system, an invisible (thus less interfering) eye-safe far-infrared (>1400 nm) laser diode may be preferred.



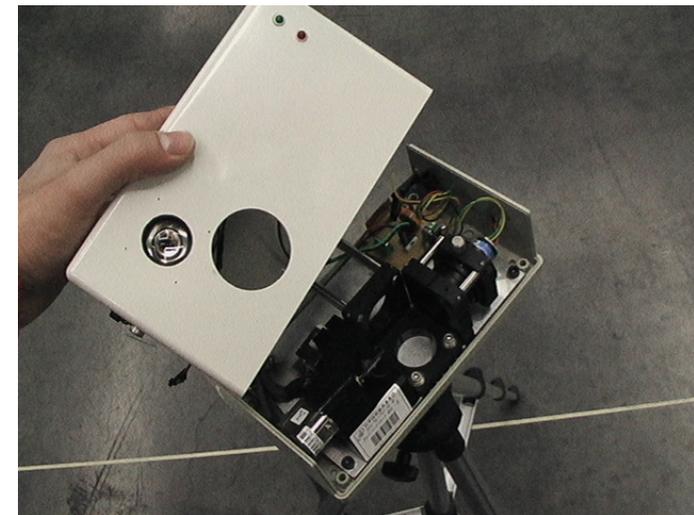
Photodetection: In the present setup, a wide-angle photodetector placed in the neighborhood of the beam-steering mechanism collects all the light from its surroundings. Thanks to synchronous photodetection, this system is insensitive to harsh or changing lighting conditions (a wavelength filter seemed unnecessary). The laser is modulated at around 40kHz (which is the Lock-In amplifier reference TTL-compatible square signal). To avoid saturation as well as large noise when the finger is close to the system, the average laser power is also modulated so as to always have the same optical power on the photodetector (this eliminates the need of a noisy logarithmic amplifier).

It is interesting to note that the overall complexity of the setup is equivalent to that of a portable laser-based barcode reader.

Demonstration and Results

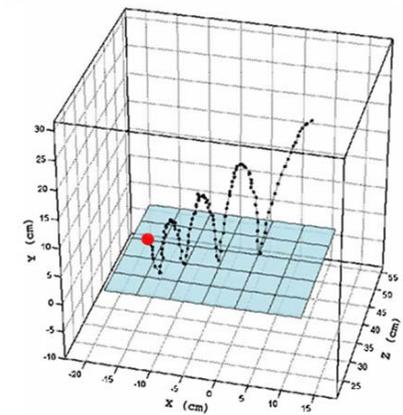
Tracking precision.

Using the simple "closed-loop" tracking mechanism described above, a tracking resolution of about 0.2° (or 3 mm at a distance of one meter) was achieved in a prototype tracking system using off-the-shelf electromagnetic-actuated micro-mirrors (mirror excursion is $\pm 52^\circ$, leading to an optical excursion range of $\pm 50^\circ$ both on X and Y axis). Depth is computed at the end of each saccade from the averaged intensity of the backscattered signal - which evolves roughly as the inverse square of the distance. Discrimination between the tracked object (e.g. fingertip) and the background (including other parts of the hand) is possible as long as the signals present enough contrast. By modeling the various sources of noise, we estimated and later verified that, without resorting to synchronous photodetection nor any signal post-processing, the prototype is already capable of sub-centimeter depth discrimination at a distance of up to 30 cm (with 90% confidence), and has a peak resolution of less than 4 mm at a distance of 5 cm from the system [3]. Using synchronous photodetection, we achieved 1 cm depth resolution at least up to 2 meters [1].



Dynamic Performances.

The dynamic performance of the system was evaluated by measuring the maximum speed an object could move without being lost by the tracking system. The tracked object was a circular piece of white paper, 9 mm in radius, tracing a circular trajectory at different uniform speeds. The distance between the mirrors and the object remained constant. Maximum tracking speed was measured to be around 3 m/s at a distance of about one meter from the mirrors (the typical speed of a finger performing gestures is < 2.5 m/s), and was only limited by the speed of the A/D interface cards [2][3]. It is interesting to note that, since in the present system the angular speed of the mirrors is not a limiting factor, for an object of a given size and assuming that the saccade radius is always properly adjusted to it, its distance from the tracking system has no effect on the tracking speed. The figure shows a real-time tracking of a bouncing ball (two seconds duration), and its corresponding real-time acquired three-dimensional data.



Writing in the air. Assuming that the plane of the gesture, though forming an arbitrary angle with the system axes, remains stable during the acquisition of a single character written on the air, it is possible to project the acquired 3D data over a real-time estimated best fitting hyperplane, thus preparing the data for feeding in a "traditional" 2D character recognition algorithm. Some experiments on character recognition using this system and Hidden Markov Models (HMMs) can be found in [3]. In the future, the reliability of a character/word based recognition algorithm can be enhanced by fusing the three-dimensional data with that of an accelerometer bound up with the acquisition system, so to compensate in real time for motion and twisting artifacts during character acquisition. The ultimate goal is to build a constraint-free input interface for character recognition. Such an interface would enable the user to write freely on the air, without putting much attention on the relative position of the writing hand with respect to the acquisition system (e.g. on a wrist watch on the opposite hand, a portable phone or a PDA for instance), nor caring about the stability of the former. If these conditions are fulfilled, it would be possible for the user to input data while walking on the street for instance.



Of course, the use of the third dimension could allow switching between different modes, such as from character input mode to a mouse-like mode (where a pointer is moved according to the finger movement). Alternatively, in drawing application, switching from a pencil tool to an eraser tool, for example, is possible. More interestingly, the third dimension can be used for more complex applications that fully make use of this feature of the system. One application could be to define a complete 3D alphabet in order to allow a more natural and richer interaction language with a machine. Another could be to allow the user to input, interact and modify 3D data in CAD applications. Lastly, 3D tracking of fingers and hands can simplify or completely redefine the way some typical GUI tasks are performed, such as moving around or controlling the zooming of a map (see video demo - the system works with bare fingers, but white balls are used as scaled-up models of the fingertip, to enable tracking at longer distances).

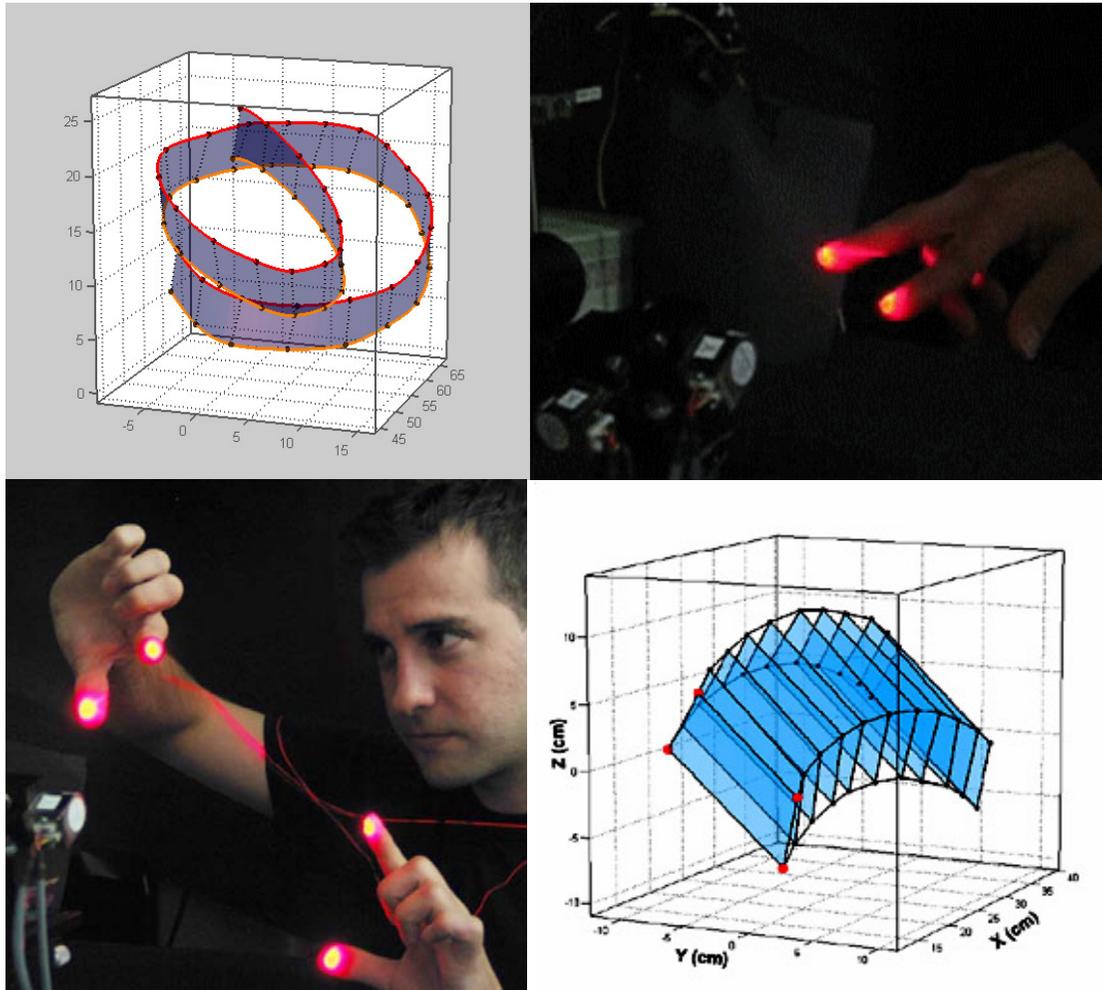


Multiple target tracking.

Applications of a multiple target tracking system are countless. Such a configuration allows, for instance, multiple users to interact on the same virtual space; or a single user to control several virtual tools at the same time, resize windows and control information screens, as imagined in Spielberg's film "Minority Report" -but without the need to wear special gloves nor markers.



Tracking of multiple targets without replicating any part of the system was successfully demonstrated [1][2]. Targets are considered sequentially. The price to pay is a tracking slowdown for each individual target, roughly equal to the inverse of the number of targets. Additional time is required to position the mirrors from one target position to another. A particularly interesting case were tracking performance seems unaffected for all purposes, is when two fingers belonging to the same hand are simultaneously tracked; the maximum tracking speed is only roughly divided by two, which leads to 1.5 m/s tracking speed - enough for most CAD applications (see figure).



Laser Annotation. Last, a very interesting characteristic of the proposed 3D laser-based locator, is that it also can be used as an output device: indeed, the laser scanner can be used to write information back to the user, by projecting alphanumeric data onto any available surface, like the palm of the hand [1]. This has been successfully demonstrated, without having to stop the tracking.

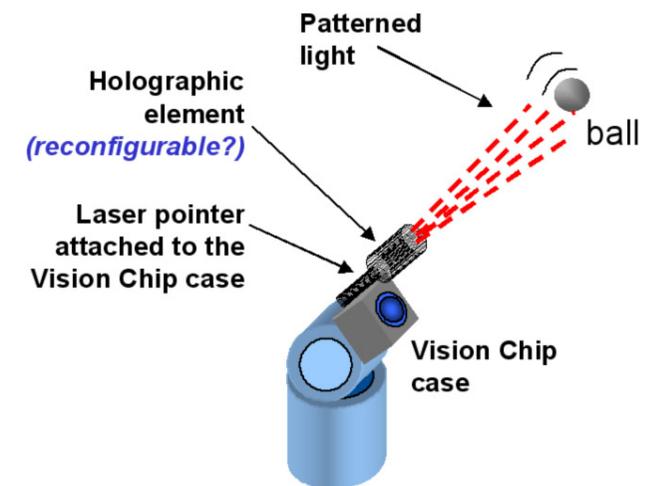
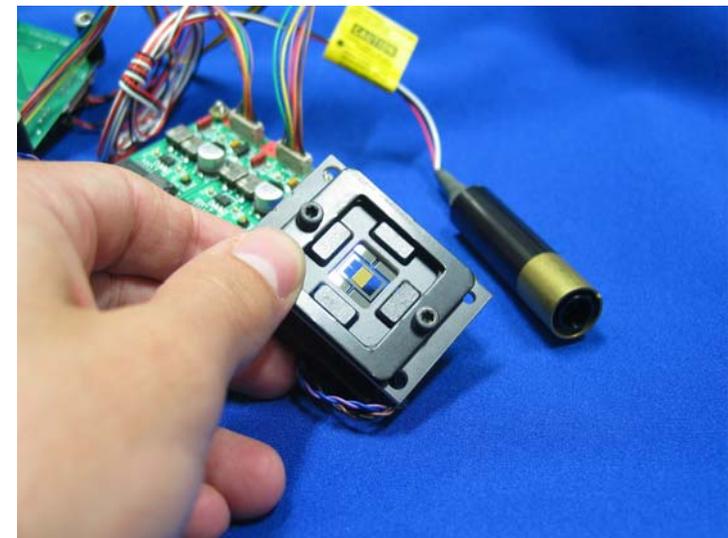
Future works

System on-chip integration. The simplicity of the system is such that, using state-of-the-art Micro-Opto-Electro-Mechanical-System (MOEMS) technology, it would be possible to integrate the whole system on a chip, making it an interesting input interface for portable computing devices. One or two-axis micro-electro-mechanical micromirrors (packaged on a chip 1 cm square in size) are already commercially available. Very similar MOEMS have been successfully built in the past, and challenges related to the integration of an optical source along with its beam-steering mechanism have been addressed both theoretically and experimentally in fields as diverse as free-space optical switching, single-chip scanners, and automatically aligned free-space optical links. We believe that once integrated, it will be readily accepted as a versatile input/output interface, freeing actual physical space on mobile devices. The whole PDA could fit in a keyboardless, screenless chest pin, that when activated will follow hand gestures, and write back information on the palm of the same hand [1].

Incidentally, a point that can significantly add interest to the system, is that it can also be used to create an adaptive free-space optical link between portable computing devices in motion. This way a -mobile- optical network could be easily created between a team of users whenever required.

Laser-enhanced Vision Chip system. **Tracking using Vision Chip** is also based on a closed-loop mechanism relying both on real-time image processing and steering of the camera viewing field. However, given the extreme frame-processing rate, in its current version Vision Chip needs a relatively powerful illumination source. Superluminescent LEDs bound up with the steerable camera module performs well in most short-range experiments. A diode laser attached to the Vision Chip module, or steered by micro-mirrors can be used even when targets are several meters away from the system. Special light patterns can be also generated (using scanning mirrors or **reconfigurable holograms**) opening the door to more powerful, active, Vision Chip-based applications.

Laser-Tracking system without mobile parts. Research is underway on a version of the laser tracking system without any mobile parts (such as micromirrors). This is possible using using advanced (integrable) optoelectronic components. Although tracking resolution is expected to be somehow reduced, it will presumably be more than enough for tracking gestures at a short distance from the system. Results will be published soon.



7.1. Physical Motion Interface

The project is to create a physical, motion based interface using cheap, common electronics and parts, and readily available software. The idea was to show that the implicit notion of human-computer interaction based upon mouse, screen keyboard is just that, an implicit which can be undermined and then improved using simple technologies

The Problem with Mice

The idea of a physical mouse based interface is more than 50 years old. Having being conceived for use with the simplest of graphical user interfaces, such as early radar screens, the mouse was not adopted widely until Apple implemented a mouse on their Lisa machine in the early 1980s. Since then, the format of the personal computer has changed little, and is still to a large extent made up of a computer, screen, mouse and keyboard. The massive advances in technology have not led to a more intuitive input device or scheme to translate physical movement into a computer interface. Despite the seemingly exponential growth of computer power, as demonstrated by Moore's "law", computer input devices, and indeed graphical user interfaces remain locked to the computer, screen, mouse, keyboard format.

This project is an experiment in not what can be done with computer interfaces, but what *should* be done. With computers becoming more pervasive in our lives, the need for a more natural interface will become necessary as industrial, medical and domestic applications become more complex, and even more integrated with computer technology. Tablet PCs have introduced the first realistic and useful implementations of technologies such as handwriting-recognition, touch screen capabilities, but these are still examples of the computer as tool, as a thing to be acted upon, rather than an integral part of interaction between humans and the real world. We have sought to break away from the existing confines of user interface devices, and build an experimental prototype which while not perfect, could demonstrate how it has become possible, with simple and cheap technologies, to fundamentally change the manner in which we interact with technology.



Alternative Systems

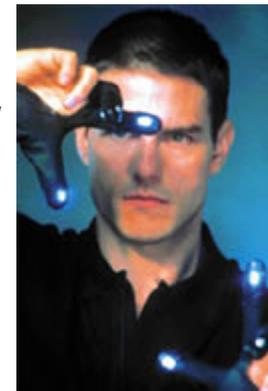
Overview

Science Fiction writers and filmmakers have had the imagination to demonstrate interfaces outside the constraints of current technologies. We have taken many of these interfaces as our own inspiration, such as the "PreCog Neural Viewer" of *Minority Report*, the retinal scanner of *Blade Runner*, and the piloting system of the ships in *Earth: Final Conflict*.

These systems are examples of how discreet physical actions can, in theory, be translated by machines into action. By removing the interface from a two-dimensional plane, and by focusing on the desired tasks and physicalities of the user, a remarkable amount of functionality can be achieved with startlingly simple technologies.

Mouse Gestures

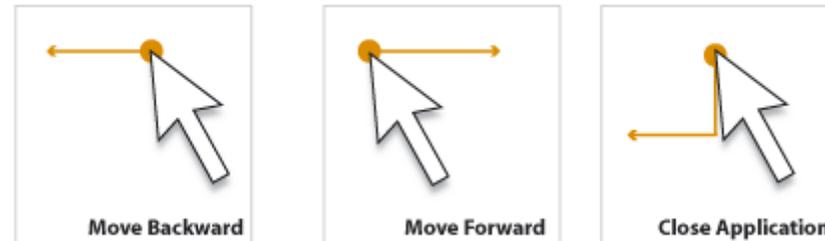
Mouse Gestures are an evolution of the classic mouse based user interface. Typically, mouse gestures are performed when the user holds down a modifier key (e.g. right mouse button) and makes a movement, or gesture with the mouse,



such as a swipe left to move back in a web browser, or a circular motion to reload the page. Mouse gestures are based on the principle that it's easier to make a simple gesture anywhere on the screen than it is to navigate the mouse pointer to a specific area of the screen and hit a button to perform an action. Gestures can be performed wherever the mouse is located, can vary in size, position and speed, and as such are more flexible actions than simple mouse clicks on designated screen areas.

In recent times, the *Opera* web browser has integrated mouse gesture technology into its navigation features, while other software such as *Stroke-It* and *WayV* offer solutions for the entire operating system, allowing users to control their computers using gesture technology. While mouse gestures could never replace menus and other typical UI elements because of the complexity of assigning, and indeed remembering, gestures for every function on a computer, they offer an effective tool in technological systems where repetitive and predictable functions are found, such as the navigation elements of a web browser, or the transport controls of a media player.

Mouse Gestures for Navigation



Tangible Interfaces

Many of the problems people have when interfacing with technologies stem from the disembodiment of the devices, where huge amounts of data, and the virtual tools we use are nothing more than symbols on the screen to be manipulated crudely. Tangible computing is concerned with engaging with the users' sense of touch, by providing an actual physical control over on-screen elements, and in some cases related sensory feedback through force-feedback devices.

Many new "Tablet PCs" engage directly with the user's body using touch areas of the screen, and while this is a far removed and very rudimentary when compared to interaction with real world objects, the added tactility of the display is a small step towards interacting with computers in the way a person would interact with everyday objects. When we consider the introduction of touch screens in consumer PCs; the challenge of moving towards representations of technology and computing which take advantage of physical ergonomics and more comfortable ways of working seems to be close at hand.

MIT's Tangible Media Group are pioneers of this type of interface, and have developed the *Senseable* project, a table that detects object placed on it and feeds input into the computer. Pictured above is one application of it, called *AudioPad*, an computer-based composing and performance instrument for electronic music. The combination of a tangible interface with the flexibility of a display that can change, combines the ease of use of a traditional instrument, everyday objects, and the power of modern computing software.



Ambient Devices

Ambient devices are an alternative form of feedback device. They are used to give the user complex information through our peripheral senses. We notice when food is burning even if we're not paying full attention to it, and tangible devices take advantage of a similar idea by providing information on situations such as the weather, stock markets through colours and sounds transmitted through discreet objects;

through . Shown on the right is the Stock Orb, a device which glows at different colours based on the performance of the stock market. Information is transmitted to the device through radio waves, meaning it can be placed anywhere as an ornament. Ambient devices represent the concept of *subtle computing*, a movement away from classic incarnations of the computer to a schema where information is located directly within our environment, not set apart on a screen.

Voice Recognition

The idea of commanding an autonomous computer using nothing more than our voice is extremely attractive. Along with physical communication, our voice is the primary means of communication between people. The goal of achieving natural language based interaction with a computer could break down many barriers between the user and technological tasks. By walking into a room and saying "Chopin", a computer could choose a piece of music, or by saying "send this file to Paul", a computer could initiate and send an e-mail to a recipient. Building voice recognition systems has proved to be difficult, and many incarnations of the technology have been problematic. IBM's ViaVoice, one of the vanguard in commercial voice-recognition packages, required the user to 'train' the computer for several hours in order for the computer to discern accent, intonation, meter of speech; quite a way behind the idea of *the computer* in *Star Trek*; a wonderous machine that can easily complete complex tasks when given instructions by any user. The target of universal, natural voice-recognition software is achievable, but it may only emerge in 'everyday' situations to complement another interface system such as physical motion. The vagaries of human physical language, body language are minute when compared to the variety of languages, accents, and rhythms of speech found worldwide.

For simple tasks such as retrieving basic information and turning on and off lights, voice recognition could be very useful. For specific tasks however, such as programming, drawing, or quickly navigating large amount of data it may be problematic for quite some time.

Physical Motion Based Systems

Body language is crucial in everyday communication, some would even argue that it is more important than conscious verbal communication. Physical gestures are a huge part of that communication process. Humans are uniquely evolved to interpret not only complex verbal communication, but also to sense mood, temperament and implicit schema of communication through body posture, arm and hand movement, eye and facial expressions. The challenge in creating a physical motion based system is to observe and decode physical movement, conscious or otherwise, and translate it into machine action.

The film *Minority Report* saw Tom Cruise's character using one such system while manipulating the 'PreCog Neural Viewer', a system which allowed him to see premonitions of 'pre-crimes' as video, and manipulate data drawn from population and criminal records, scrubbing back and forth through video, and throwing unwanted files away using hand gestures. This direct manipulation of data bridges the gap between computers and the physical world, removing the abstraction of user interfaces, replacing it with body movements designed around the tasks at hand.

The Project

Our project was to create a physical, motion based interface using cheap, common electronics and parts, and readily available software. The idea was to show that the implicit notion of human-computer interaction based upon mouse, screen keyboard is just that, an implicit which can be undermined using simple technologies.

Goals

Our primary goal was to expand our awareness of user interface technologies, with a view to implementing a prototype system which could form the basis of a physical interface for use in our undergraduate thesis. We set ourselves these conceptual goals:

- To explore & research alternative user interface devices.
- To conceptualize, and design a physical, gesture based, system capable of carrying out simple interface tasks.
- To investigate and evaluate for our purposes various methods of motion tracking that could be used in a limited timeframe.
- To build our system at minimal cost.
- To evaluate the potential use of this system in future projects.

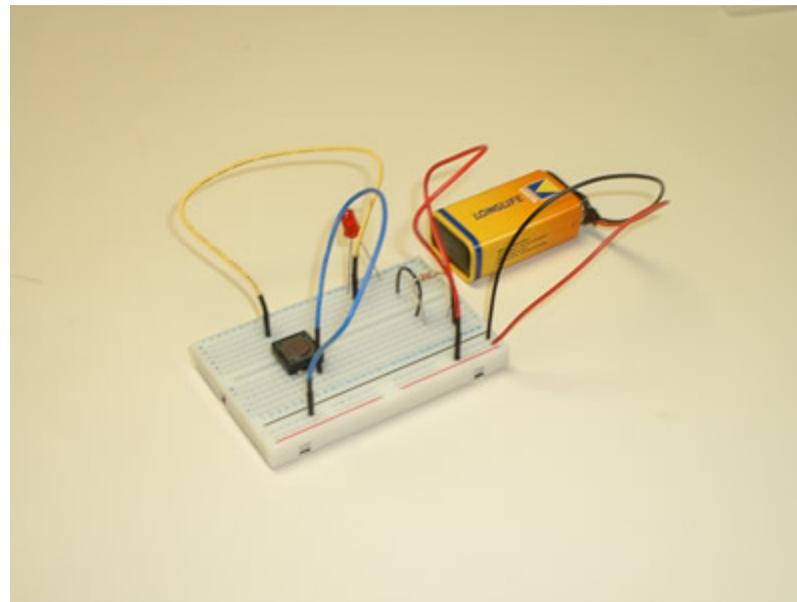
Construction

After researching and discussing several ways of implementing the physical interface, including motion and tilt sensors directly connected to the computer, we conceived a basic system which would be based upon a vision-tracking/target method. We felt that this method was advantageous because of its separation from the actual computer, and because it represented a 'discrete' interface (one which could be worn and not used, or used and then left in place). Our vision based system would be based on three key components:

1. A PC running our software system
2. A device with coloured LEDs bright enough to be discerned from background objects
3. A camera connected to the PC to follow the LED device.

Having evaluated several programming environments, including Java and Processing, we decided to use Macromedia Director for our application. Java presented a high barrier to entry, as we would have been required to learn, or relearn the language, in a short space of time. The experimental Processing environment was extremely attractive, as it represented a middle ground between Java and a multimedia environment, but after tests the robustness and speed of the environment on anything but the fastest hardware was questioned, and we abandoned it in favour of a proven environment; Director. In future implementations of the system, a more robust environment would be used to directly communicate with, and interpret camera input.

Director, while familiar presented several challenges. We had to settle for a third party Xtra to expand its capabilities, and were therefore constrained by its functionality. In a high-level language such as Java or C++ we would have been able, although with difficulty, to manipulate



information coming via the USB camera, but with Director we had to use the functions and information being interpreted by the Xtra. We chose to use the freely available WebCam Xtra by [Josh Nimoy](#) to handle the interface between the camera and the Director application, and some of the basic colour tracking functionality.

Tracking Modules

Initial prototypes involved a basic 'tracking module', nicknamed "the orb". The module would be an everyday object which could be picked up used, and left down again. Similar to a remote control, this 'orb' would be a type of wand interface, which would be followed by the camera. This system was rejected as a general purpose device, as making it feel like a part of the users body, beyond a mere remote control was difficult, and the movement it produced was less natural than we had expected. We think that it has potential usage in simple applications, particularly art installations where limited features are required, but we wanted a system that could be used in a wide range of industrial, domestic and commercial implementations, and which would translate physical movement naturally, beyond the idea of a ruggedised remote control.

Glove

After exploring these alternative ideas we returned to the idea of a glove interface, resolving that it was offered more potential, and created a more natural and flexible interface environment. A number of prototypes were made including, functional mockups for use in software testing, and then a final version which enclosed the wires and switches as much as possible to create a comfortable, attractive device. We tested many different types of LEDs to find the ones that offered the most colour contrast without flooding the camera.

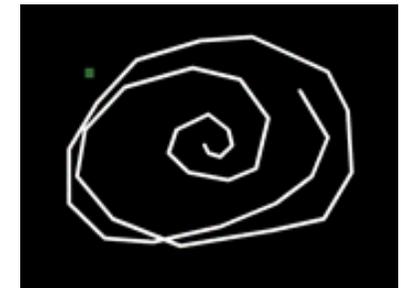
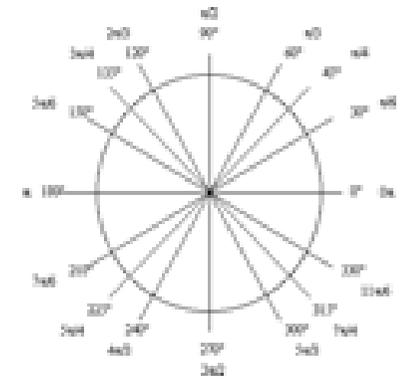
Software

The software was custom made in Director, and utilized Josh Nimoy's [webcam_xtra](#) for the basic tracking technology. Various version of the software were made taking basic The final revision of the prototype system, which interprets the user's gestures and translates them into audiovisual controls. Gestures to the compass points of the screen produce the following actions:

- Right** = Fast Forward
- Left** = Rewind
- Top Right** = Play
- Top Left** = Pause
- Up** = Increase Volume
- Down** = Decrease Volume
- Down Right** = Skip
- Down Left** = Restart

Results

Through our tests we achieved a solid motion tracking technique, which works well in controlled conditions. Performance was heavily related to lighting conditions & background, due to the extremely limited pixel resolution of the camera, and the broad nature of the software used. Best results are achieved in a dimly lit environment with a neutral background.



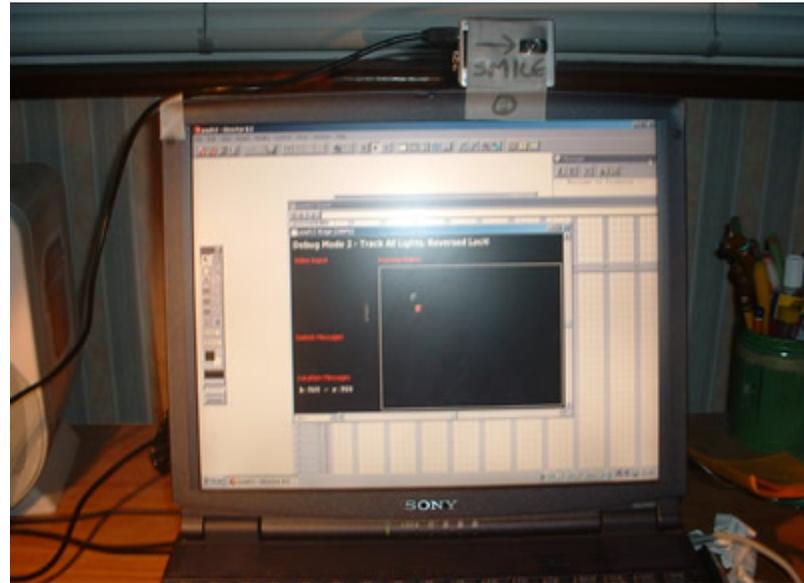
For industrial usage, more 'intelligent' image tracking code would be needed. Motion tracking software is an industry in itself, and we achieved quite decent results considering the time and resources available. Use of alternative tracking solutions could be used such as tracking Infrared LEDs. Although that was out of the scope of our budget for this project, we strongly feel that it could reap significant benefits over a traditional camera and light source set up.

The gesture engine recognises and interprets actions, and while functionality is limited, it carries out definite tasks well based upon user motion. The goal of controlling the computer based on physical motion captured through a camera works well, as demonstrated in pmi21 where video controls are tied to specific gestures.

Future/Lessons Learned

This project represented a valuable learning experience, and gave us an insight into user interfaces, alternative user interfaces, hardware and electronics and software implementations of gesture systems.

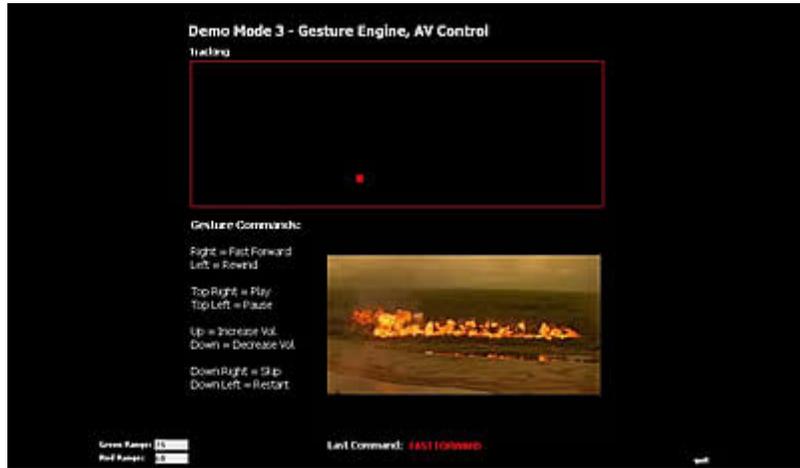
Director does not represent a robust, scalable system sufficient for commercial development of the system. The challenges presented by the Xtra were great, but it did allow us to work at a high programming level, which was advantageous in the timeframe. For future development a more low-level approach would be needed where we could directly access data from the camera. Ideally, a tracking and gesture system would be hardware based for deployment in everyday applications.



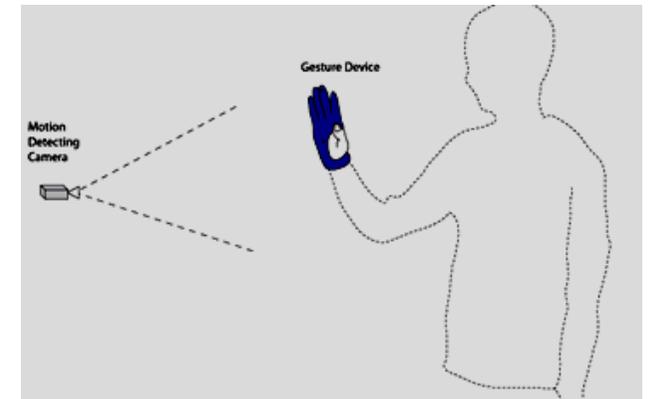
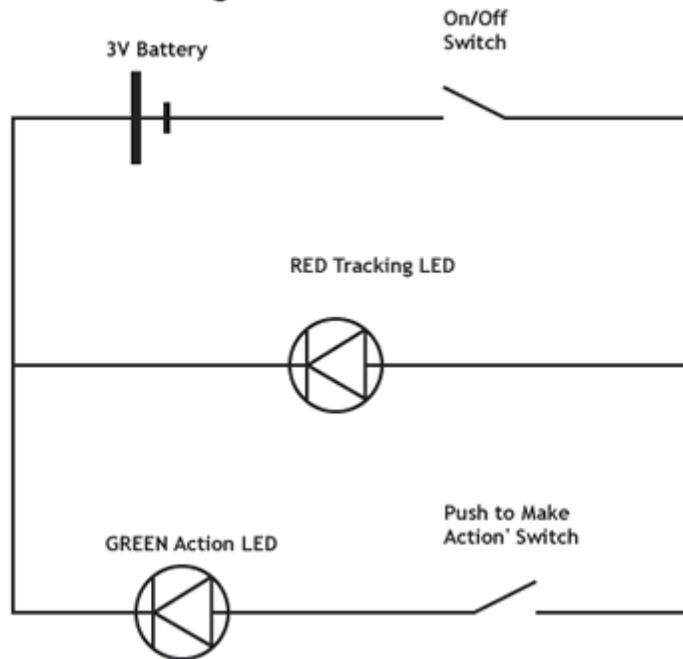
Conclusions

This project was an exploration of the potential for computer technology. By breaking away from the traditional user interface paradigm we feel that there may be great benefit in alternative interfaces, and using computers could become a simple pleasure rather than a frustrating tool. By looking at the way people communicate with each other and manipulate their environment we can adapt computer tools to better enable us to use technology.

Our project sought to build an experimental prototype physical motion interface. We created a basic application which can easily be extended further by us, or by others. We have demonstrated that basic interfaces can be achieved using simple and affordable technology and that interfaces of the future could be fundamentally different than current technologies. The project was not meant as an alternative to the mouse and keyboard, but rather as a challenge to the implicit ways we use technology, and a proposition of one alternative system.



Circuit Diagram



7.2. Sensorizzazione di una protesi d'arto inferiore

I sensori più adatti allo scopo sono:

- Un accelerometro montato sull'invaso per rilevare l'accelerazione a cui è sottoposta la protesi. In questa posizione l'accelerometro rileva anche la contrazione del muscolo;
- Estensimetri montati sulla tibia per capire quale parte del piede è in contatto con il terreno (punta, tallone, pieno);
- Un estensimetro montato su una lamina estensimetrica per rilevare la rotazione del ginocchio o in alternativa un potenziometro a filo oppure un encoder a filo realizzato in laboratorio

Le specifiche di progetto sono:

- utilizzare solo materiale commerciale
- tutti i sensori devono essere montati sulla parte centrale della protesi, quindi non sull'invaso, né sulla tibia o al di sotto di questa
- i costi devono essere contenuti

In prima analisi si può dire che verranno mantenuti gli estensimetri perché affidabili, economici, robusti e precisi;

Specifiche di progetto impongono che l'accelerometro sia installato non più sull'invaso ma sul ginocchio pertanto occorrerà verificare che l'informazione rilevata non sia sostanzialmente diversa da quella rilevata sull'invaso. Sicuramente si perderà l'informazione relativa alla contrazione del muscolo;

La protesi monta un ginocchio commerciale anziché uno sperimentale per cui occorrerà scegliere il sensore più adatto a misurare la rotazione. Erano stati scartati sia il potenziometro a filo perché troppo costoso, sia l'encoder incrementale perché troppo ingombrante. La soluzione era quella di utilizzare la lamina estensimetrica che aveva il solo inconveniente di non rilevare rotazioni superiori a 90°. Il ginocchio commerciale ha una struttura completamente diversa e occorre quindi studiare dove collocare il sensore. Da una prima analisi non sembra possibile montare la lamina estensimetrica per problemi di spazio.

Prima di dedicarsi alla parte sperimentale (montaggio e taratura dei sensori) si approfondiscono le conoscenze sui sensori impiegati, nella specie si riporteranno le caratteristiche generali e non soltanto quelle specifiche al tipo di applicazione.

7.2.1. Estensimetri

Principio di funzionamento

L'estensimetro consiste di una griglia di metallo laminato incollata su un supporto di materiale plastico. Esistono migliaia di configurazioni per adattarsi alla più svariate applicazioni. L'estensimetro è incollato sulla superficie della struttura di cui si vuole misurare la sollecitazione. La deformazione della struttura provoca una variazione della resistenza dell'estensimetro che può essere rilevata da una strumentazione apposita con elevata precisione (ordine micrometro su metro). Per determinare sollecitazioni nelle varie direzioni della struttura si usano estensimetri costituiti da griglie montate in diverse direzioni (rosette). La risposta degli estensimetri è praticamente istantanea e permette di rilevare deformazioni dinamiche con frequenze che arrivano a migliaia di Hertz.

Selezione dell'estensimetro

L'installazione e il funzionamento dell'estensimetro sono influenzati dai seguenti parametri:

- lega metallica sensibile alla deformazione
- supporto
- lunghezza griglia
- configurazione griglia
- numero di auto-compensazione termica
- resistenza della griglia

Nella scelta dell'estensimetro occorre tenere conto di varie specifiche che devono essere soddisfatte:

- precisione
- stabilità
- temperatura
- allungamento
- durata della prova
- resistenza ciclica
- facilità d'installazione
- ambiente

Leghe utilizzate

- *costantana auto-compensata in temperatura*: è la più anziana ma ancora la più usata in virtù della sua buona sensibilità alla deformazione. Ha una elevata resistività (anche con griglie piccole si possono avere resistenze elevate) e un coefficiente di temperatura non eccessivo. Ha una buona durata e una buona capacità di allungamento. Difetti: Per temperature superiori a 65°C si ha una variazione permanente della resistenza ad ogni ciclo (deriva dello zero). Pregi: può essere trattata per l'autocompensazione termica;

- *costantana trattata*: utilizzata per misurare deformazioni superiori al 5%; è duttile in griglie di 3mm, può essere allungata del 20% o più. Difetti: ha una variazione della resistenza ad ogni ciclo (zero non stabile); tendenza alla rottura se sottoposta a deformazioni ripetute. Mal si adatta all'impiego per deformazioni cicliche;

- *lega isoelastica*: si usa per misure puramente dinamiche perché ha una vita a fatica superiore rispetto alla lega in costantana e un migliore rapporto di taratura che migliora il rapporto segnale rumore. Difetti: non ha l'autocompensazione termica; ha una deriva dello zero; la sua risposta alla deformazione non è lineare;

- *lega Karma*: ha una buona vita a fatica e un'eccellente stabilità. E' usata in misure statiche con un range di temperature da -260°C a 260°C circa. Se incapsulata in materiale inerte può essere fino a 400°C. Ha una curva termica più piatta quindi le correzioni di uscita termica sono più precisi. Offre l'autocompensazione termica.

Materiale di supporto

La fabbricazione convenzionale di estensimetri consiste di un foglio metallico fotoinciso su un supporto plastico. Il supporto ha delle funzioni importanti:

- consente la manipolazione durante l'installazione
- fornisce una superficie incollabile per aderire l'estensimetro al materiale
- fornisce isolamento elettrico tra il foglio metallico e l'oggetto in prova

Lunghezza estensimetro

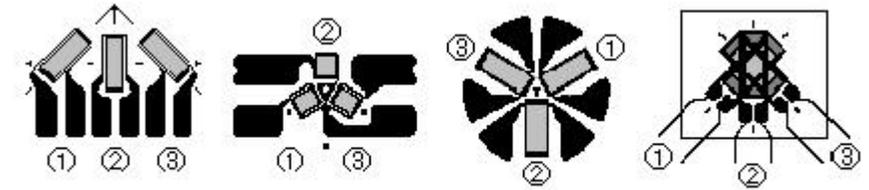
La lunghezza di misura di un estensimetro è la lunghezza attiva o sensibile della griglia. Generalmente si vuole fare una misura estensimetrica dove il gradiente della sollecitazione è più elevato; solitamente l'area di deformazione è molto piccola quindi poiché l'estensimetro media le deformazioni misurate da tutta la griglia, una griglia troppo grande potrebbe dare una misura troppo bassa. Estensimetri troppo piccoli hanno però svantaggi in quanto sono meno stabili se sottoposto a sollecitazioni statiche, hanno un minore allungamento e una minore resistenza a fatica se sottoposti a deformazioni cicliche. Estensimetri più grandi sono più facili da maneggiare ed hanno una maggiore dissipazione termica. Come regola pratica, salvo applicazioni particolari, si usano estensimetri tra i 3 e i 6 mm.

Configurazione dell'estensimetro

Occorre scegliere la forma della griglia, il numero di griglie, la disposizione (se si ha più di una griglia), la configurazione dei terminali e le caratteristiche costruttive.

Singola griglia

- Occorre verificare dimensione e orientamento delle piazzole da saldare
- Una griglia più larga migliora la dissipazione termica, per contro misura anche la sollecitazione perpendicolare alla direzione principale
- E' meglio scegliere estensimetri a elevata resistenza perché meno sensibili al valore di resistenza dei cavi
- Devono essere usati solo quando la sollecitazione è assiale



Rosetta a 2 o 3 elementi

- Sono usate quando si vuole misurare uno stato biassiale delle sollecitazioni.
- Se si conoscono le direzioni principali dello sforzo si possono usare rosette con 2 griglie a 90°
- Se le sollecitazioni non sono note si devono usare rosette a 3 griglie. Di norma si orienta una griglia in una direzione significativa per la struttura. Possono avere configurazione rettangolare a 45° o a delta di 60°.

Bisogna anche considerare le differenze tra rosette piane e sovrapposte.

Quella sovrapposta ha un migliore smaltimento del calore e tutte le griglie sono a contatto con la superficie. Quella piana è più precisa in corrispondenza di forti picchi nella sollecitazione poiché tutte misurano nello stesso punto

Compensazione termica degli estensimetri

Idealmente la resistenza dell'estensimetro non dovrebbe variare con la temperatura, tuttavia questo non avviene, pertanto la relazione tra sforzo e resistenza non varia solo con la deformazione ma anche con la resistenza. Una volta che l'estensimetro è installato in un apparato di misura bilanciato una variazione della temperatura provoca una variazione dell'indicazione dello strumento per effetto della variazione della resistenza. Tale uscita è chiamata uscita termica (thermal output). A seconda dei casi questo effetto può essere trascurato in altri casi bisogna correggere l'errore introdotto.

$$\frac{\Delta R}{R} = \left[\beta_G + F_G \cdot \left[\frac{(1 + K_T)}{1 - \nu_0 \cdot K_T} \right] \cdot (\alpha_s - \alpha_G) \right] \cdot \Delta T$$

$\frac{\Delta R}{R}$ variazione di R rispetto alla R di riferimento

β_G coeff termico della resistenza della griglia del conduttore

F_G gage factor dell'estensimetro

K_T sensibilità trasversale dell'estensimetro

ν_0 coefficiente di Poisson

$\alpha_s - \alpha_G$ variazione tra coeff di espansione termica del substrato e della griglia

ΔT variazione della temperatura

$$\xi = \frac{\Delta R}{R \cdot F_I}$$

ξ thermal output

F_I gage factor di set

Il thermal output è causato da due effetti: variazione della resistenza con la temperatura; differente coefficiente termico tra la griglia del conduttore e il materiale su cui è incollato il conduttore, in quanto se il substrato si espande la griglia è forzata ad espandersi. Un'espressione per il TO è quella a pagina precedente. Se però l'estensimetro e lo strato sono mantenuti alla stessa temperatura durante il test, non esiste il suddetto effetto. Il metodo più semplice di compensazione è quello di collegare due estensimetri su due lati adiacenti del ponte di Wheatstone. Affinchè la compensazione abbia effetto devono valere le seguenti condizioni:

- i due estensimetri devono essere montati su due substrati identici e devono essere sempre sottoposti alla stessa temperatura
- i due estensimetri devono essere identici (stesse dimensioni e caratteristiche)

Un altro metodo è quello dell'autocompensazione, utilizzabile quando si usano estensimetri che hanno un coefficiente di espansione termico molto basso, praticamente costante nell'arco del test. Per compensare occorre inserire sul ponte di Wheatstone, sul lato adiacente a quello dove si è inserito l'estensimetro, una resistenza con un coeff. di temperatura molto basso. Talvolta questa resistenza è già inserita nei morsetti dell'estensimetro stesso. Quando non è possibile fare la compensazione occorre correggere la lettura effettuata. Esistono tabelle sui data-sheet che riportano la correzione da sottrarre alla indicazione dell'estensimetro al variare della temperatura di prova. Senza scendere troppo nei dettagli occorre dire che i dati delle tabelle sono riferiti a specifiche condizioni di prova (ad esempio per un determinato settaggio del gage factor dello strumento)ma che comunque possono essere riportati a qualsiasi condizione usando formule correttive.

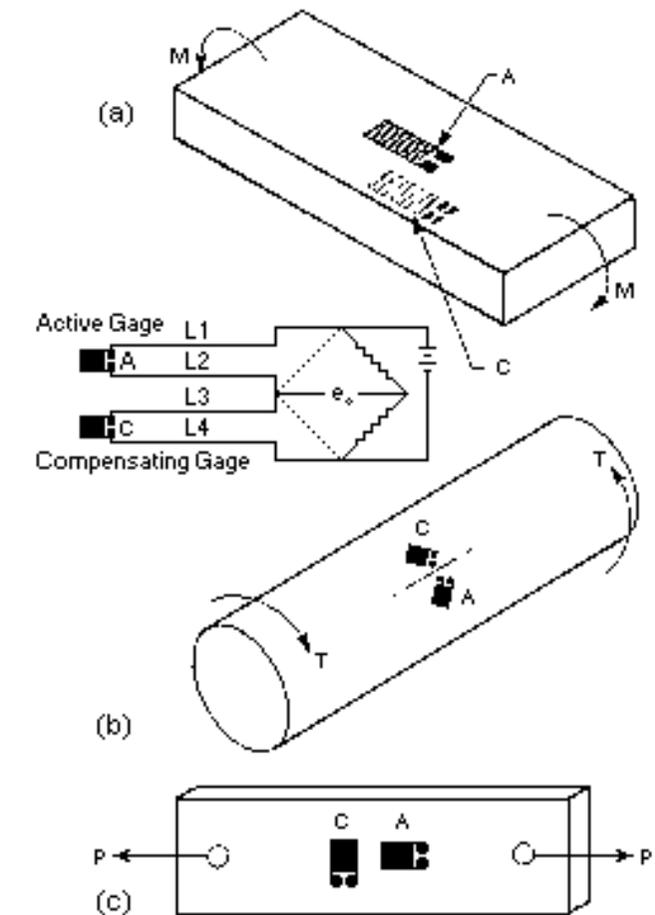
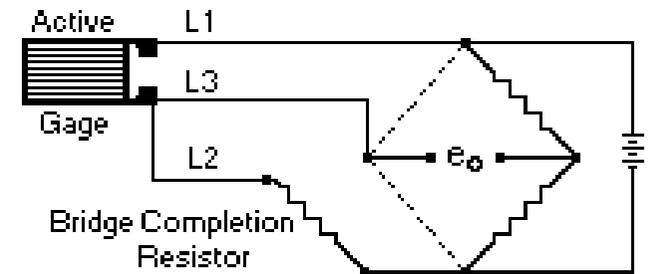
Compensazione termica per gli estensimetri montati sulla tibia della protesi

Nel caso della protesi la compensazione termica si esegue molto facilmente montando i 2 estensimetri in posizione diametralmente opposta cercando di rispettare le condizioni citate sopra. Altri tipi di compensazione non vengono presi in considerazione perché più complicati e perché comunque l'utilizzo di 2 estensimetri è necessario per avere informazioni più precise sulle sollecitazioni della tibia.

Calibrazione della strumentazione degli estensimetri mediante shunt (Shunt calibration)

Esistono due metodi diversi per calibrare la strumentazione di misura degli estensimetri:

Calibrazione diretta: questo metodo è adatto per tarare ad esempio trasduttori (ad esempio quelli di pressione) in cui è possibile applicare un ingresso noto con precisione: lo strumento viene regolato finì a quando non misura il valore corrispondente)



Calibrazione indiretta: viene applicata quando non è possibile usare la calibrazione diretta; nel caso degli estensimetri, supponendo di utilizzare un ponte di Wheatstone come strumento di misura, non è possibile applicare uno sforzo noto con precisione. Pertanto si procede variando la resistenza di un lato del ponte e rilevando l'indicazione dello strumento: concettualmente si potrebbe inserire una microresistenza variabile in serie all'estensimetro, ma dal punto di vista pratico si opera mettendo una resistenza in parallelo alla resistenza di uno dei lati.

Per operare una corretta calibrazione occorre:

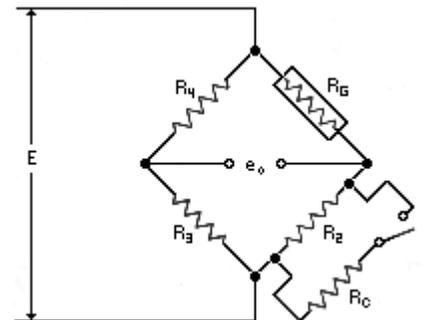
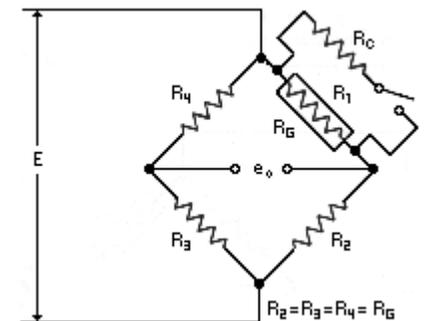
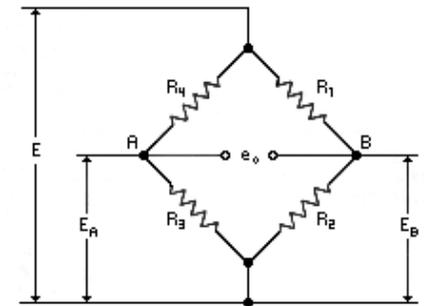
- scegliere il lato del ponte che deve essere shuntato;
- calcolare il valore della resistenza di shunt che simuli un determinato sforzo a cui corrisponda una determinata uscita
- considerare la non linearità del ponte di Wheatstone

Esistono due diversi tipi di calibrazione indiretta, il primo mirato a realizzare una scala dello strumento, il secondo mirato alla verifica della linearità strumento. Nella realizzazione della scala dello strumento si regola il guadagno dello strumento affinché lo sforzo indicato non corrisponda esattamente allo sforzo simulato. Occorre tenere conto ad esempio della resistenza dei cavi di collegamento. Nella verifica della linearità viene fornito un segnale di ingresso allo strumento di misura a cui deve corrispondere una determinata uscita. In caso contrario lo strumento non opera correttamente. Alimentando il ponte con una tensione E il valore della tensione di uscita è dato da:

$$E_A = E \left(1 - \frac{R_4}{R_4 + R_3} \right) \quad E_B = E \left(1 - \frac{R_1}{R_1 + R_2} \right) \quad \frac{e_o}{E} = \frac{R_1/R_2}{R_1/R_2 + 1} - \frac{R_4/R_3}{R_4/R_3 + 1}$$

Si noti che l'uscita dipende solo dal rapporto delle resistenze e quando $R_1/R_2 = R_3/R_4$ l'uscita è zero. Nella pratica si cerca di operare con le quattro resistenze tutte uguali. Supponendo di avere un solo estensimetro su un lato del ponte: shuntando tale estensimetro si ha una variazione negativa della resistenza pari a:

$$\Delta R = \frac{R_1 \cdot R_c}{R_1 + R_c} - R_1$$



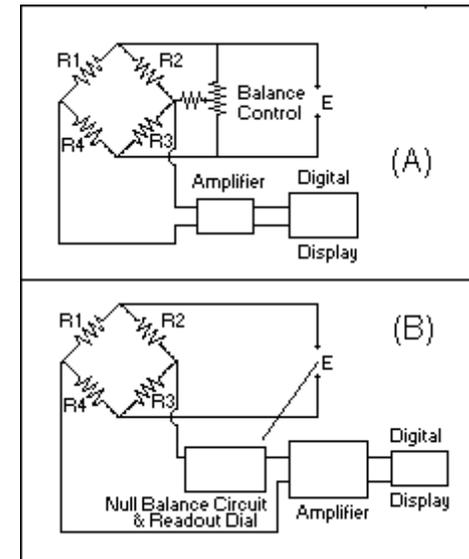
La variazione della resistenza dell'estensimetro è data da :

$$\Delta R = F_G \cdot \xi \cdot R_G$$

dove R_G resistenza nominale estensimetro
 ξ sforzo
 F_G gage factor

Per cui si ottiene:

$$F_G \cdot \xi = \frac{-R_G}{R_G + R_c} \quad R_c = \frac{R_G}{F_G \cdot \xi} - R_G$$



E' quindi possibile dimensionare R_c . Si nota che ad ogni variazione della resistenza corrisponde una variazione lineare dell'uscita (in teoria trascurando le non linearità del ponte). Poiché si ha una variazione negativa di resistenza si possono rilevare sollecitazioni di compressione.

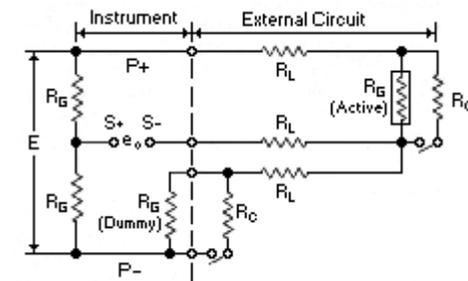
Shuntando il lato adiacente del ponte si ha invece un aumento dell'uscita interpretabile come una sollecitazione a trazione dell'estensimetro. Quando si ha una connessione remota tra strumento ed estensimetro occorre considerare anche la resistenza dei collegamenti. L'unica resistenza che dovrebbe essere trascurabile è quella posta in serie allo strumento che misura e_0 perché si suppone di resistenza infinita.

Questa situazione si verifica quando si usa anziché il classico amplificatore, uno strumento di misura che opera annullando l'uscita del ponte. Nella pratica si shunta il lato adiacente a quello attivo perché più facile operativamente.

Occorre considerare che nel normale funzionamento con i moderni estensimetri il ponte di Wheatstone funziona sbilanciato e questo porta ad avere errori di non linearità del ponte. Infatti supponendo che varino due resistenze adiacenti, l'uscita varierà di (variazione rispetto al caso lineare):

$$\frac{e_0}{E} = \frac{R_1 + \Delta R_1}{R_1 + R_2 + \Delta R_1 + \Delta R_2} - \frac{R_4}{R_4 + R_3}$$

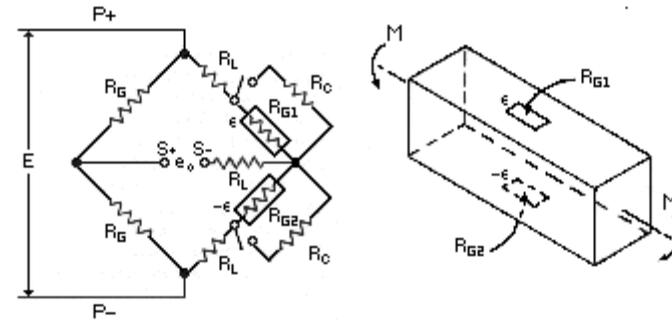
$$\Delta \left(\frac{e_0}{E} \right) = \frac{R_1 + \Delta R_1}{R_1 + R_2 + \Delta R_1 + \Delta R_2} - \frac{R_1}{R_1 + R_2}$$



Come si vede affinché l'uscita del ponte si mantenga lineare deve essere:

$$\Delta R_1 = -\Delta R_2$$

infatti in questo caso la corrente che passa nel ramo non cambia e l'uscita è proporzionale alla sola variazione di resistenza. Ovviamente nel caso generale non si hanno problemi di non linearità del ponte di Wheatstone ogni qualvolta la corrente nel ponte rimane costante. Tornando a parlare della calibrazione la non linearità del ponte di Wheatstone è quindi in linea di principio trascurabile quando si ha a che fare con piccole variazioni della resistenza estensimetrica, mentre devono essere considerate quando la variazione è più ampia.

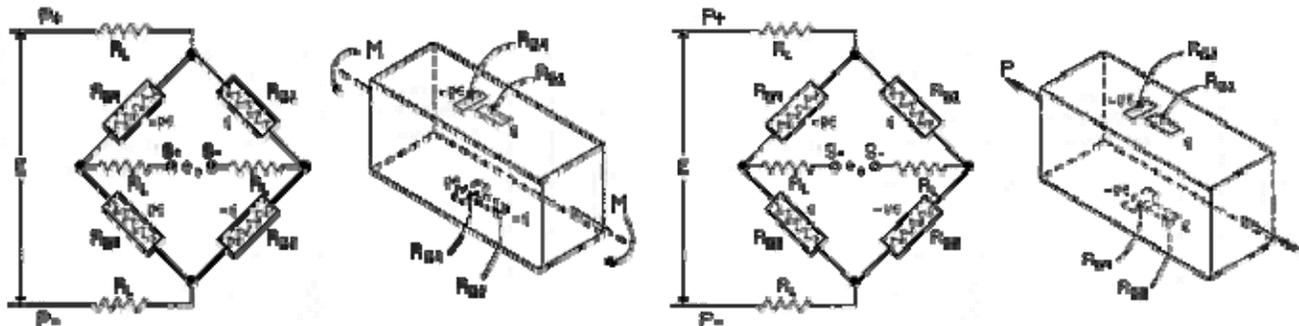


Estensimetri montati sulla gamba

Nel caso della gamba si ha la situazione rappresentata sotto in quanto la tibia è sottoposta ad un momento torcente: di conseguenza i due estensimetri saranno sempre sottoposti uno a trazione e l'altro a compressione. Si può ritenere con buona approssimazione che la resistenza totale di un lato del ponte non vari e che quindi la risposta sia di tipo lineare durante l'esecuzione della taratura. Problemi di non linearità si avranno invece in fase di calibrazione in quanto bisogna far variare una resistenza di shunt alla volta. Il metodo più semplice per ovviare al problema è simulare piccoli allungamenti di modo che si possa considerare l'uscita del sistema lineare. Nel caso fosse necessario operare con elevati sforzi occorre applicare formule correttive per tenere conto della non linearità del ponte. Esistono casi in cui anche in fase di misura occorre considerare la non linearità del ponte. In questi casi, misurato uno sforzo, l'errore commesso è calcolabile come η espresso in microstrain, per cui lo sforzo effettivo è dato da:

$$\eta = \frac{F_G \cdot \xi^2 \cdot 10^{-6}}{2 - F_G \cdot \xi \cdot 10^{-6}}$$

$$\xi = \xi_1 + \eta$$



In alto sono rappresentati alcuni casi in cui si ha la non linearità del ponte.

Nella verifica della linearità dello strumento occorre fornire un segnale di ingresso noto allo strumento di misura e verificare che l'uscita sia esatta e la variazione sia proporzionale alla variazione del segnale. Tale risultato può realizzarsi con uno shunt cercando di eliminare l'incertezza introdotta dalle resistenze di collegamento. Occorre notare che c'è una differenza logica tra i due modi di operare in quanto, quando si va a calibrare la scala, occorre calcolare R affinché l'uscita sia quella desiderata per un determinato sforzo. Il gage factor dello strumento è quindi regolato affinché l'uscita sia quella voluta. Nel secondo caso invece si ha una uscita che deve dipendere solo dalla resistenza dello shunt e non dal gage factor dell'estensimetro e del ponte. L'unico parametro esterno che deve essere noto è la resistenza a riposo dell'estensimetro.

Riassumendo, le operazioni che andranno fatte sugli estensimetri montate sulla gamba saranno la calibrazione del ponte e della strumentazione di misura, al fine di poter rilevare e monitorare gli sforzi in modo preciso. Come già detto non occorre considerare errori di non linearità del ponte ed errori dovuti alla non perfetta compensazione termica.

Una soluzione da evitare è collegare più estensimetri in parallelo allo stesso strumento. Ciò permette di misurare più sollecitazioni provenienti da estensimetri diversi con un unico strumento variando di volta in volta la coppia di morsetti su cui si misura la e_0 .

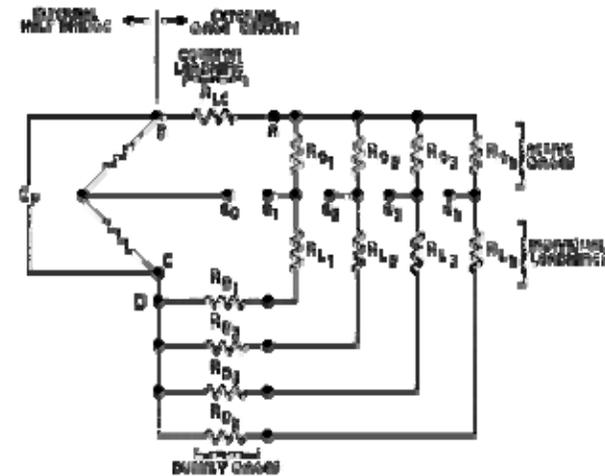
Questo procedimento porta a notevoli errori in quanto la resistenza del conduttore che collega strumento a estensimetro è sempre attraversato dalla somma delle correnti che passano in ciascun estensimetro; non si può quindi eliminare l'errore introdotto attraverso una procedura di calibrazione come si faceva precedentemente. Pertanto a meno di casi estremi questa soluzione è da non applicarsi.

Errori dovuti alla sensibilità trasversale degli estensimetri

Con sensibilità trasversale degli estensimetri si intende la capacità degli estensimetri di rilevare sforzi in direzione perpendicolare a quella dell'asse principale. Ovviamente un estensimetro ideale dovrebbe rilevare sforzi solo lungo la direzione principale ed essere insensibile alle sollecitazioni in tutte le altre direzioni.

Il motivo per cui nella realtà ciò non si verifica è perché una porzione di filo è posta alla fine di ogni giro in direzione perpendicolare; questo comporta che il segno della sensibilità trasversale sia sempre positivo, e poiché, nota la geometria della griglia, è facilmente calcolabile l'intensità di tale grandezza è quindi sottraibile al valore complessivo misurato

Negli estensimetri a foglia la sensibilità trasversale è legata a molti più fattori per cui il coefficiente può essere positivo o negativo a seconda del disegno della griglia, del materiale usato, dello spessore etc..



L'errore dovuto alla sensibilità trasversale è di solito trascurabile, tuttavia a rigore teorico occorre verificare che il coeff di Poisson (coefficiente di contrazione trasversale) del materiale su cui si effettuano le prove, sia lo stesso per cui è stato verificato l'estensimetro. Ogni qualvolta si opera con un coefficiente di Poisson diverso, occorrerebbe considerare anche questo errore.

In generale un estensimetro ha 2 gage factor F_a e F_t , uno riferito alla sollecitazione assiale e l'altro a quella trasversale; vale la relazione:

$$\frac{\Delta R}{R} = F_a \cdot \varepsilon_a + F_t \cdot \varepsilon_t$$

dove

ε sono gli sforzi (in direzione assiale e trasversale)

considerando il coefficiente di Poisson si può scrivere che:

$$\varepsilon_t = -\nu_0 \cdot \varepsilon_a$$

$$\frac{\Delta R}{R} = F_a \cdot (\varepsilon_a - K_t \cdot \nu_0 \cdot \varepsilon_a) \quad \text{dove} \quad K_t = \frac{F_t}{F_a}$$

$$\frac{\Delta R}{R} = F_a \cdot (1 - K_t \cdot \nu_0) \cdot \varepsilon_a$$

$$\frac{\Delta R}{R} = F \cdot \varepsilon_a$$

Pertanto il costruttore fornisce il coefficiente F (gage factor) che tiene conto anche della sensibilità trasversale ovviamente per un dato valore del coeff di Poisson. Nel caso di operare con un materiale con coeff di Poisson diverso da quello specificato si avrebbe un errore pari a:

$$\eta_e = \frac{K_t \cdot \left[\frac{\varepsilon_t}{\varepsilon_a} + \nu_0 \right]}{1 - \nu_0 \cdot K_t}$$

Esistono formule correttive per ogni tipo di rosetta utilizzata. Tali formule non sono riportate in questa trattazione.

Occorre porre molta attenzione nell'incollare l'estensimetro alla superficie su cui si vuole fare una misura; infatti posizionando male l'estensimetro si ha un errore che è sostanzialmente diverso da quello dovuto alla sensibilità trasversale:

infatti si avrà un errore di misura dello sforzo principale dovuto al non perfetto allineamento dell'asse dell'estensimetro con quello dello sforzo nelle due direzioni (uno lungo il piano d'appoggio in quanto l'estensimetro è ruotato rispetto alla direzione principale e uno perpendicolare al piano dovuto al fatto che lo strato di colla non è uniforme)

Taratura degli estensimetri

La prova di taratura degli estensimetri montati sulla protesi consiste nel verificare che lo stato di sollecitazione rilevato dagli stessi in varie situazioni di appoggio della gamba. Rispetto alla taratura fatta sulla precedente protesi è stata modificata la collocazione degli estensimetri che non sono più sulla tibia in titanio ma bensì sulla parte superiore in materiale plastico. Si tratterà di verificare se nonostante questa nuova collocazione la taratura sarà ugualmente efficace. L'esecuzione della prova è di per sé molto semplice. Occorre sollecitare la protesi con pesi differenti e verificare in base alle condizioni di appoggio qual è lo stato di sforzo rilevato e soprattutto se esiste una corrispondenza biunivoca tra sollecitazione rilevata e tipo di appoggio.

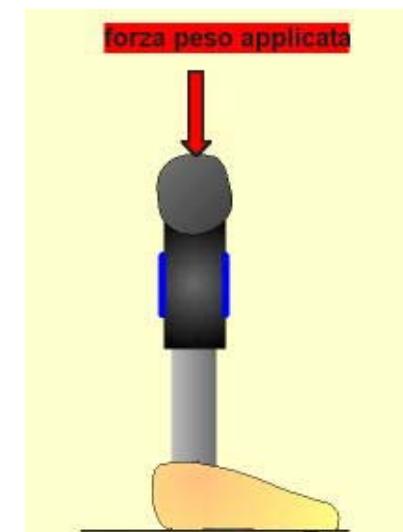
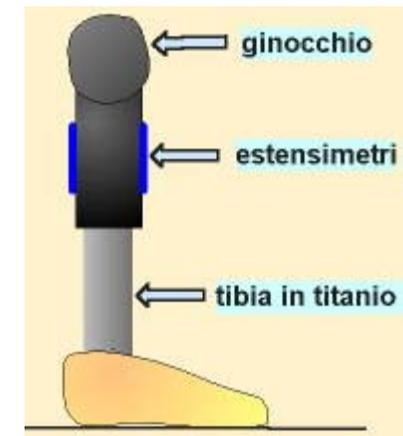
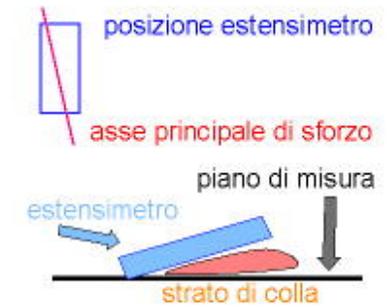
Si applica quindi un peso al ginocchio e si rileva attraverso un ponte di Wheatstone lo sforzo misurato. In particolare la misura viene fatta in queste condizioni di appoggio:

Appoggio completo: Se gli estensimetri sono montati correttamente non dovrebbe essere rilevata nessuna sollecitazione poiché lo sforzo è di sola compressione, mentre per come sono stati montati gli estensimetri rilevano solamente una flessione.

Appoggio del tallone: In questo caso appoggia solo il tallone. Dal punto di vista pratico occorre scegliere l'angolo di inclinazione della protesi rispetto alla verticale ed effettuare la misura. Al fine di avere una maggiore precisione occorre effettuare la prova con diversi angoli di inclinazione per testare meglio il funzionamento degli estensimetri.

Appoggio della punta: Concettualmente la prova è la stessa per l'appoggio del tallone. Anche in questo caso occorre fare più misure con diversi angoli di inclinazione.

Problemi pratici: nell'effettuare la prova bisogna porre particolare attenzione al modo in cui si appoggia il piede al terreno. Infatti per la conformazione del piede non è facile riuscire ad appoggiare il piede sempre nella stessa posizione. Non è garantito che la protesi rimanga sempre a 90° rispetto al terreno.



In particolare al variare dell'appoggio varia l'angolo di carico e questo può portare ad errori di misura e a discordanze. E' da notare che una volta che la protesi è montata sulla gamba del paziente questo problema non si avrà più in quanto l'accoppiamento sarà rigido e scomparirà il grado di libertà che permette la rotazione attorno al terreno.

Numero di prove da eseguire

Per ottenere risultati attendibili occorre ripetere la stessa misura più volte. Pertanto il numero di campioni minimo è dato dalla formula statistica:

$$n = \frac{(\sigma * Z_{\alpha/2})^2}{d^2}$$

essendo:

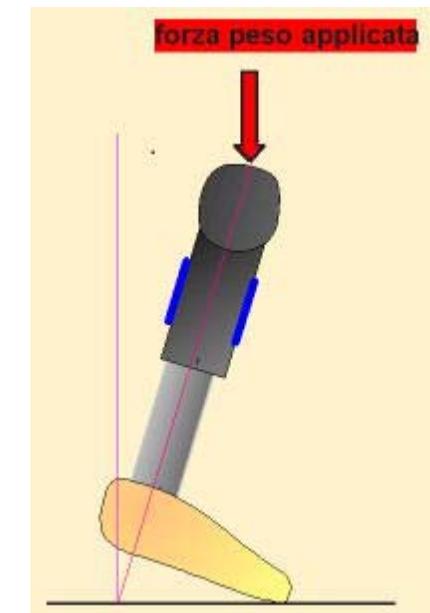
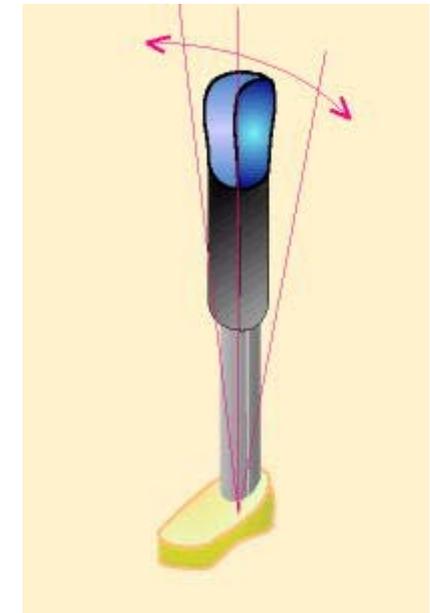
n = dimensione del campione cercata

s = varianza

Z a /2 = pivot sull'intervallo di confidenza

d = scostamento dalla media

Scelto l'intervallo di confidenza e lo scostamento massimo dalla media voluta si ha il numero di campioni per ogni prova.



7.2.2. Accelerometro

L'accelerometro montato sulla protesi è di tipo sperimentale, quindi non si hanno molte informazioni specifiche. In questa parte si illustrerà il principio di funzionamento non avendo informazioni dettagliate in proposito.

L'accelerometro trasduce un'accelerazione in un segnale elettrico. E' pertanto costituito ad 2 componenti, un sensore e una interfaccia elettrica. Gli accelerometri possono essere classificati in base a caratteristiche costruttive, in particolare in base ai materiali impiegati e alla tecnologia utilizzata.

I tipi più comuni sono:

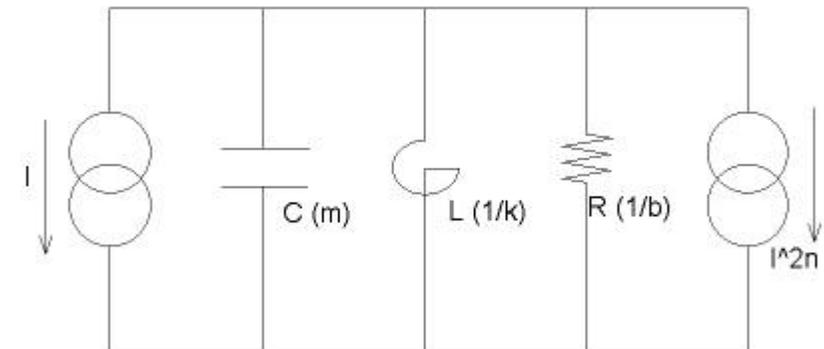
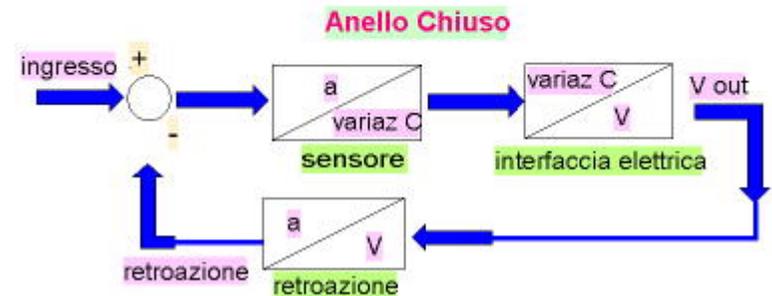
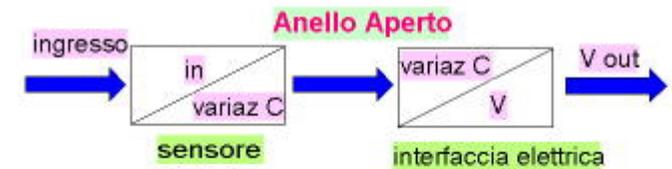
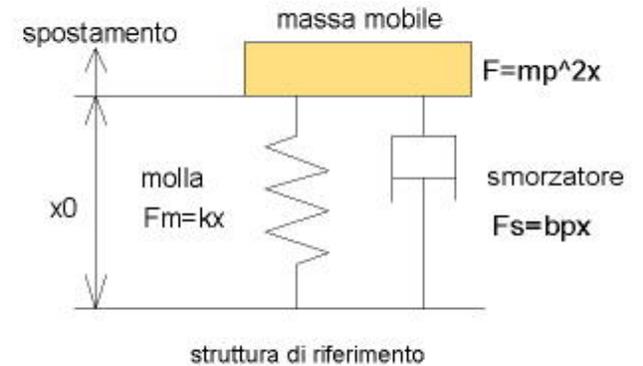
- piezoelettrici;
- piezoresistivi;
- a variazione capacitiva o effetto tunnel.

L'accelerometro montato sulla gamba è del tipo a variazione capacitiva che come dice il nome fornisce una variazione capacitiva se sottoposto ad una accelerazione. Il sensore di natura meccanica è concettualmente strutturato come a destra: in seguito ad una accelerazione si ha movimento della massa mobile rispetto alla struttura di riferimento con conseguente variazione della capacità. Circuitualmente il modello equivalente è del tipo a destra. L'equazione che lega lo spostamento x alla forza di sollecitazione esterna è del tipo:

$$m \cdot \frac{d^2}{dt^2} x(t) + b \cdot \frac{d}{dt} x(t) + k \cdot x(t) = F$$

L'interfaccia elettrica può essere realizzata in due diversi modi:

- si può leggere e amplificare direttamente la variazione capacitiva: in questo modo si ha un approccio di tipo diretto e si ha in uscita una tensione direttamente proporzionale alla variazione della capacità. Tale schema è ad anello aperto
- si può creare un sistema retroazionato dove l'anello di retroazione misura lo spostamento dell'elemento sensibile dalla sua posizione nominale ed applica una forza per tenerlo



centrato. In questo caso quindi la massa sarà sottoposta solo a piccole variazioni e quindi se l'anello ha una notevole larghezza di banda si può estendere la banda del sensore oltre alla frequenza naturale della massa. Questo dispositivo non si può utilizzare in quelle applicazioni dove l'accelerazione d'ingresso è di molti G in quanto l'anello di retroazione non è in grado di realizzare le forze di attrazione necessarie a mantenere l'elemento sensibile nella posizione di riposo. Poiché come già detto non si hanno informazioni dettagliate sull'accelerometro usato si riportano alcuni dati tecnici di accelerometri commerciali prodotte da altre case costruttrici comunque utilizzabili sulla protesi.

Silicon Designs

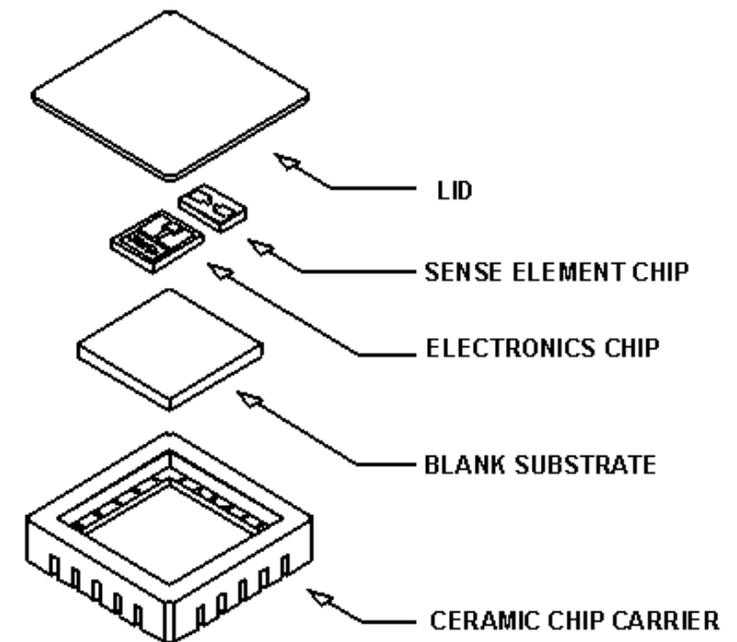
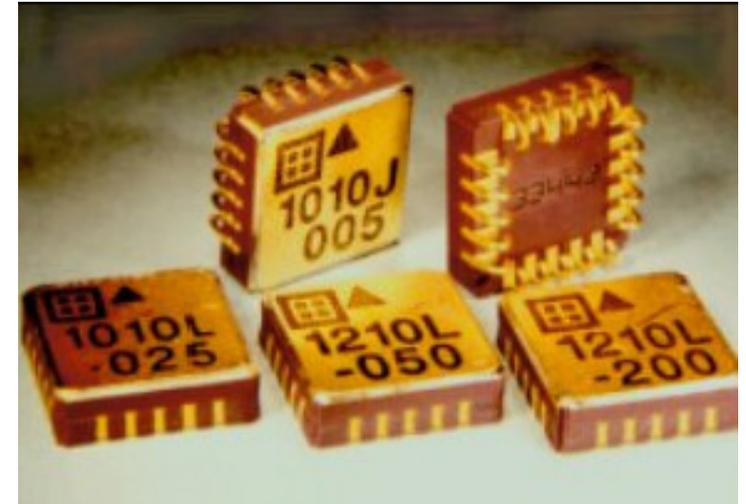
Questa casa costruttrice ha in commercio accelerometri molto simili a quello utilizzato sulla protesi. Di seguito ne vengono presi in considerazione alcuni rimandando al sito della casa per maggiori dettagli

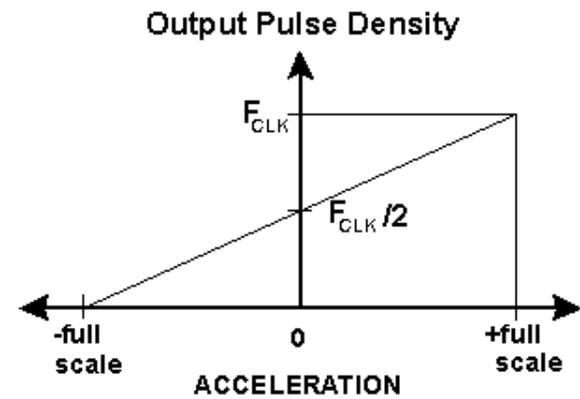
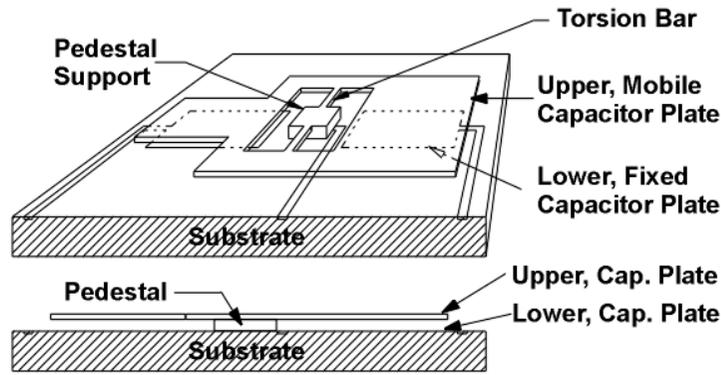
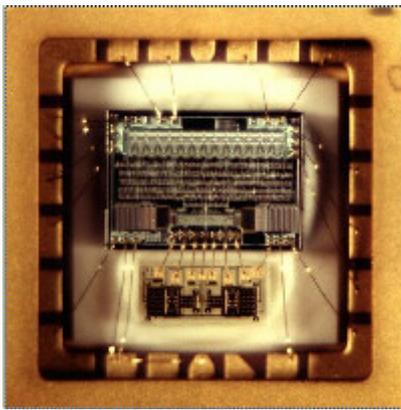
Accelerometri ad un solo asse.

È un accelerometro di dimensioni e peso molto contenuti (0.75 g). E' disponibile sia con uscita di tipo digitale che analogica.

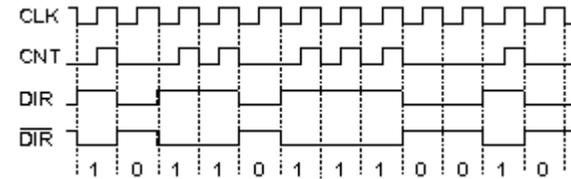
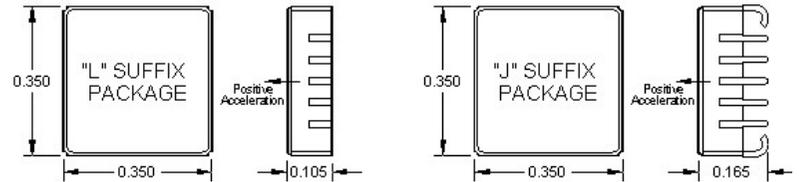
Principio di funzionamento: è un accelerometro a variazione capacitiva. L'elemento sensibile è una lamina in nichel sostenuta da due barrette di torsione fissate ad un piedistallo centrale che sorregge la struttura. La struttura presenta una asimmetria (un lato è più pesante) che è compensata dalla forza di torsione esercitata dalle barre. Quando un'accelerazione produce un momento attorno all'asse di torsione delle barre la lamina superiore può ruotare attorno a tale asse limitata solo dalla reazione esercitata dalle lamine stesse.

Come si può notare dalla figura, la lamina superiore corrisponde al piatto superiore del condensatore. Ogni lato di tale lamina è affacciato ad un piatto inferiore appoggiato al substrato. Di conseguenza si hanno due condensatori con una connessione in comune. Quando la lamina ruota una capacità aumenta (diminuisce la distanza tra le lamine) mentre l'altra diminuisce (aumenta la distanza). Le dimensioni della lamina sono 1000*600* 5 micron (lunghezza*larghezza*spessore) e la capacità di ogni piatto è circa 0.15 pF. La sensibilità del dispositivo varia al variare della massa, della distanza del baricentro della lamina dal centro di torsione, e dal coefficiente di torsione delle barre. Per evitare sollecitazioni eccessive è presente un fermo meccanico a fine corsa. Ogni accelerometro è costituito da due lamine sensibili. Un chip provvede a convertire la variazione di capacità in un treno di impulsi (nel digitale) o in una tensione (nell'analogico).



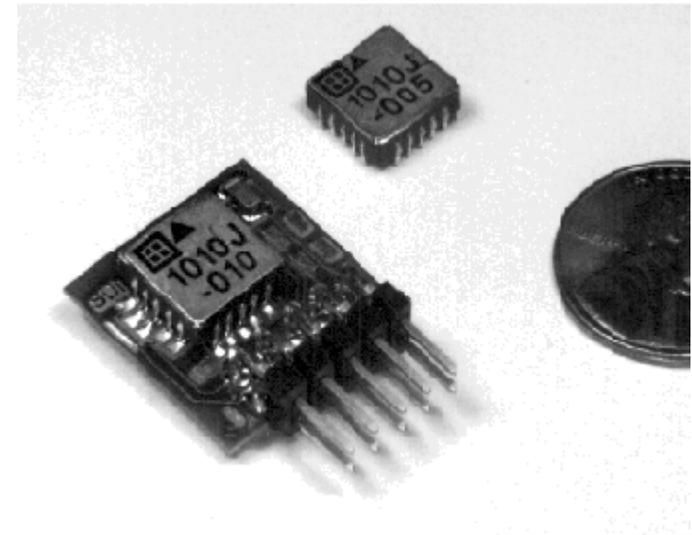


Digitale: ha un circuito integrato che ha un convertitore sigma-delta A/D integrato. E' disponibile in vari modelli per misurare accelerazioni da 50 fino a 200g. Fornisce un treno di impulsi la cui densità (n° impulsi per minuti) è proporzionale all'accelerazione applicata. Richiede un'alimentazione di 5V e un clock di 100 kHz - 1MHz. Features: Digital Pulse Density Output - Low Power Consumption - -55 to +125 °C Operation - Built-in Nitrogen Damping - TTL/CMOS Compatible - +5V DC Power - No External Reference Voltage - Easy Interface to Microprocessors - Good EMI Resistance - Response to DC & AC Acceleration - Non Standard G Ranges Available - Hermetic LCC or J-Lead Surface Mount Package.



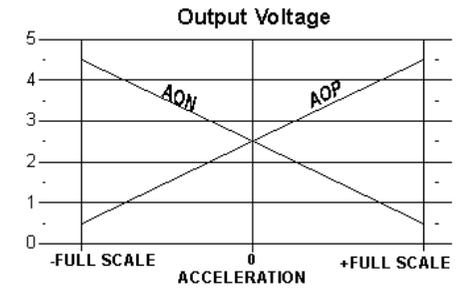
Descrizione dei segnali: clk è il segnale di clock dato in ingresso; cnt treno di impulsi in uscita; dir ad ogni fronte di discesa del clock rimane alto se cnt è alto. Filtrato con un passa basso può fornire un'uscita di tipo digitale.

Analogico: ha le stesse dimensioni di quello digitale ed è anche questo disponibile in vari modelli per accelerazioni da 50 a 200 g. Features: ± 4V Differential or .05 to 4.5V Single Ended Output - Low Power Consumption - -55 to +125 °C Operation - Built-in Nitrogen Damping - Calibrated to 1% Bias and Scale Factor (Typ) - +5 V DC Power - Responds to DC and AC Acceleration - Non standard G Ranges Available - Hermetic LCC or J-Lead Surface Mount Package. Fornisce due uscite di tipo analogico che possono essere usate separatamente oppure in modo differenziale. come si nota quando l'accelerazione è nulla entrambe le uscite sono pari a 2.5 volt. AOP aumenta all'aumentare dell'accelerazione (positiva); AON diminuisce con l'aumentare dell'accelerazione. Le principali differenze tra questi due accelerometri e quello utilizzato nella prova sono essenzialmente: l'accelerometro utilizzato è retroazionato, mentre questi funzionano in anello aperto; la scheda fornita prevede sia una uscita di tipo analogico che digitale. Non si sa se lo stesso accelerometro può funzionare sia in analogico che digitale oppure se la scheda è la stessa per entrambi i tipi di funzionamento. C'è da dire che anche la scheda è di tipo sperimentale quindi non si hanno informazioni dettagliate; l'accelerometro è stato utilizzato con uscita analogica. A differenza di quello presentato qui presenta una sola uscita e non due. E' possibile acquistare

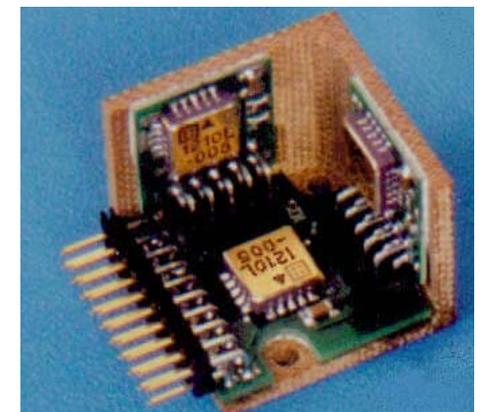
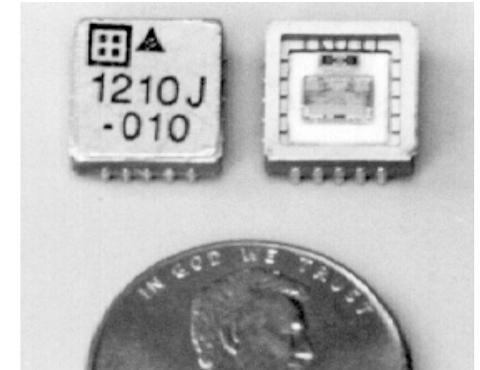


l'accelerometro in appositi moduli che prevedono il cavo di interfaccia verso strumenti registratori. A scopo informativo si riporta anche un accelerometro a 3 assi: è ottenuto semplicemente dall'unione di 3 accelerometri ad un solo asse.

Accelerometri Seika serie B BD BDK: anche questa casa costruttrice produce sensori capacitivi di dimensioni contenute che potrebbero essere utilizzati sulla protesi. Features: very high overload-solidity; different casing-variations; low weight; linear frequency-walk without as well as with very inferior resonance-superelevation at the low low border-frequency (especially BDK..); high signal - intoxication-relationship (especially BD..); enclosed hermetically; low cross-sensitivity; high long-time-stability; integrated sensor-electronics; very low stream-reception (especially BD..).



Type	BD3 / BDK3	BD10 / BDK10	BD100 / BDK100
Accelerazione rilevabile	±3g (ca.±30m/s ²)	±10g (ca.±100m/s ²)	±100g (ca.±1000m/s ²)
Risoluzione	<10 ⁻³ g	<5*10 ⁻³ g	<5*10 ⁻² g
Campo di frequenza BD-Serie	10...300Hz	10...800Hz	10...1500Hz
Campo di frequenza BDK-Serie	1....300Hz	1....800Hz	1....1500Hz
Sensibilità fino U _b = 5Volt	ca.150mV/g	ca.60mV/g	ca.15mV/g
variazione con temperatura all'equilibrio	<0,1mV/K		
Impedenza di uscita	ca. 100 Ohm		
Sollecitazione massima ammissibile	ca.10 000g (ca.100 000m/s ²) !		
Alimentazione	U _{bN} = 5Volt		
tensione ammissibile	U _{bz} =2 ... 16Volt		
Assorbimento U _b = 5V	BD..: ca. 250uA (optional: 30uA) BDK...: ca. 2mA		
Protezione	IP65		
Temperatura di lavoro	-40° C bis +85° C		
Temperatura di deposito	-45° C bis +90° C		
Peso con struttura in acciaio senza cavi	ca. 17Grammi		
Peso con struttura leggera	ca. 7Grammi		



PCB piezotronics: accelerometri miniaturizzati piezoresistivi

DJB è il proprietario del brevetto degli accelerometri "konic", i quali sono realizzati con un elemento piezoelettrico a forma di tronco di cono cavo. Questo sistema di realizzazione offre diversi vantaggi tra cui una spiccata affidabilità, e consente di ottenere alte sensibilità con dimensioni contenute. Anche la tecnologia 'piezo-ceramic shear plate' viene impiegata da DJB, ed offre un alto grado di miniaturizzazione ed un miglioramento delle caratteristiche di sensibilità. Gli accelerometri piezo-tronic integrano un convertitore carica/tensione, e sono alimentati da una sorgente di corrente costante. Questi trasduttori hanno un'uscita in tensione a bassa impedenza e sono immuni al rumore. Sono inoltre disponibili accelerometri industriali, per alte temperature, collaudati sotto pressione, e intrinsecamente sicuri. Sono disponibili anche accelerometri realizzati in titanio, e quindi estremamente leggeri.

Prove eseguite sull'accelerometro

Le prove sull'accelerometro sono mirate a verificare cosa rileva l'accelerometro in varie situazioni, in particolare in camminata normale, in corsa, durante la salita e la discesa dalle scale, nell'atto di sedersi e alzarsi da una sedia.

Le prove non sono state eseguite su un arto artificiale ma su una persona in quanto non era possibile fare diversamente. Pertanto l'accelerometro è stato applicato sul ginocchio di una persona che ha simulato tutte le prove di cui sopra. L'accelerometro del tipo ad asse singolo rileva per come è stato montato sul ginocchio l'accelerazione centripeta.

Il segnale d'uscita è stato monitorato utilizzando una scheda di acquisizione dati ed il software LabView della National Instruments. Si è collegata l'accelerazione rilevata con l'angolo di rotazione del ginocchio, proprio perchè sulla protesi le informazioni che arrivano sono essenzialmente queste due, più l'informazione rilevata dagli estensimetri che dicono quale parte del piede sta appoggiando. Il primo rilevamento è stato effettuato con l'accelerometro in posizione di riposo, appoggiato sul banco per verificare la stabilità dell'uscita. Come si può notare l'uscita si mantiene stabile attorno al valore di 2.115 volt

Camminata: in tutti i grafici il picco iniziale (in questo caso è attorno al campione n 70) è dovuto al fatto che la persona a cui è stato montato l'accelerometro, per segnalare l'inizio del test, batteva il piede per terra. Analizzando i risultati si nota una certa ripetitività del segnale. Si nota il segnale dato dal primo passo, una zona in cui il segnale rimane stabile, e il secondo passo.



High Sensitivity,
100 mV/g, 2 gm



Low Profile, 10
mV/g, 0.5 gm



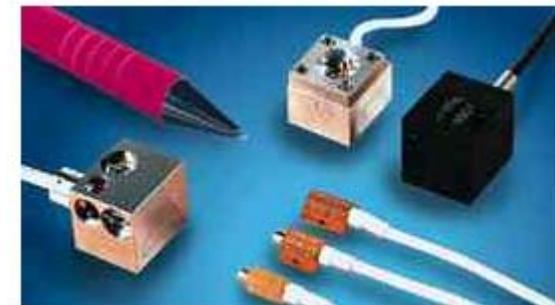
Light Weight,
Hermetic, 10
mV/g, 0.7 gm



Miniature High-Frequency Quartz
ICP®
Accelerometer, 10
mV/g, 1.8 gm



Miniature High-Frequency Quartz
Accelerometer, 10
mV/g, 1.5 gm



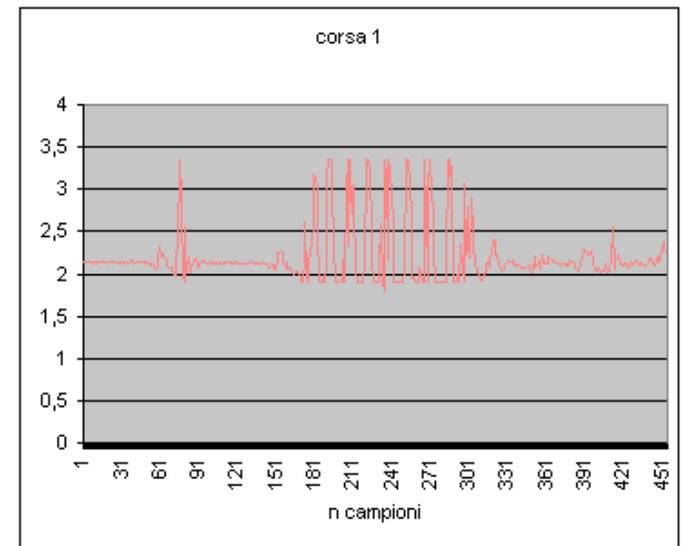
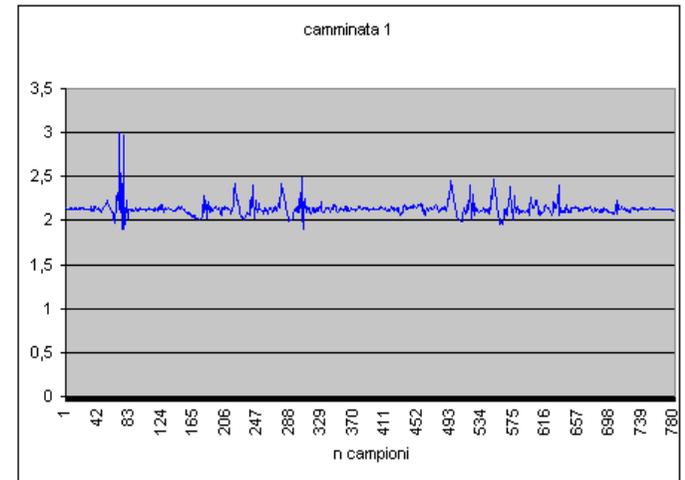
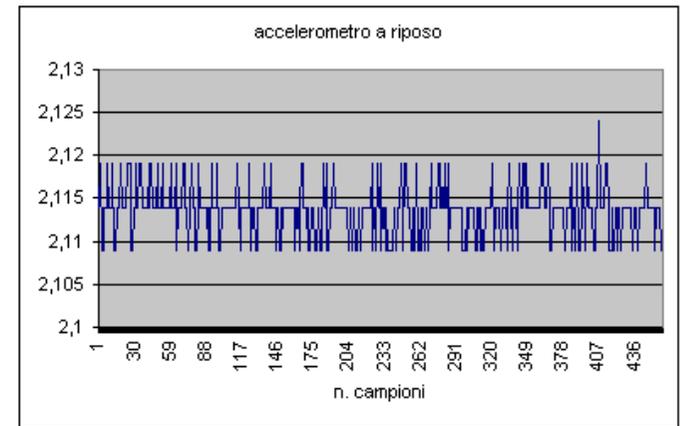
Corsa: si nota un susseguirsi di picchi di ampiezza costante. Il segnale è anche in questo caso periodico. Poiché i picchi sono molto vicini tra loro sono stati rilevati anche dati relativi ad una corsa lenta.

Discesa da scalini: come in tutte le prove anche qui si rileva la periodicità del segnale, in particolare si nota attorno al campione 220 l'appoggio della gamba senza accelerometro, l'appoggio della gamba con accelerometro, e poi il segnale che si ripete ciclico.

Salita delle scale: in questa prova il segnale rilevato non è periodico, anzi si fatica a capire le varie fasi dell'appoggio. Tale prova è stata ripetuta più volte (come le altre del resto , anche se è stato riportato un solo rilevamento) ma i risultati sono stati sempre gli stessi. Il motivo della non periodicità può essere spiegato col fatto che gli scalini sono stati realizzati in laboratorio con materiale di fortuna, quali ad esempio mattoni, uno sgabello ed una sedia. In particolare la struttura non appariva molto solida e stabile: è possibile che la persona durante la fase di discesa non avesse particolari problemi, ma nella fase di salita in cui lo sforzo è maggiore, avesse problemi di equilibrio e qualche tentennamento, che ha alterato la prova. Questo deve comunque far riflettere su quanto sia delicata e complicata la taratura di questo sensore, in quanto piccoli tentennamenti fanno perdere l'aperiodicità del segnale.

Il modo migliore per sfruttare la misura dell'accelerometro è disporre di una memoria per analizzare la periodicità del segnale. Il valore puntuale di per sé non risulta essere molto utile nemmeno se integrato dalle misure degli altri due sensori. Sarebbe utile poter effettuare una analisi armonica del segnale in tempo reale e in base all'armonica principale capire se la persona sta camminando, correndo etc. Non è però possibile montare sulla protesi un processore per questo scopo a causa di problemi di costo.

La gestione del dato rilevato è complicata, in quanto con un valore puntuale non si è assolutamente in grado di capire cosa sta facendo la persona. Nelle idee iniziali appariva possibile dalla rilevazione dell'accelerometro capire se, ad esempio durante la camminata, la persona aveva inciampato. Ci si aspetta una variazione del segnale dell'accelerometro. Simulando un inciampo durante la camminata ci si aspetta di vedere un picco fuori sincronia rispetto al passo; ovviamente guardando il grafico a posteriori, si vede che c'è un inciampo quando il segnale perde periodicità, ma il problema è un altro: come può il software accorgersi dell'inciampo? L'unica possibilità è quella di "prevedere il futuro" cioè dai dati precedenti sapere che tipo di dati aspettarsi in futuro, e se il dato rilevato è diverso da quello atteso agire in modo opportuno perché si è verificato un problema. A questo punto sorge però un altro problema: come può il software capire che la persona ha smesso di correre e ha ripreso a camminare senza interpretare la variazione del segnale come un problema? Come si vede la gestione delle informazioni dei sensori è in generale molto complicata e non si pensa di poter risolvere i problemi in questa trattazione.



L'ultima prova effettuata, se vogliamo un po' diversa dalle altre è la seduta e alzata da una sedia: la prova non è dinamica come negli altri casi, ma si rileva solo una variazione del segnale di uscita in quanto varia, dalla posizione eretta a quella seduta, la componente dell'accelerazione di gravità rilevata. Qui ovviamente appare più semplice realizzare il software in grado di capire se la persona è seduta oppure in piedi.

Come conclusione appare un'impresa ardua riuscire ad utilizzare i dati rilevati dall'accelerometro. A parte questo il sensore appare certamente stabile e preciso.

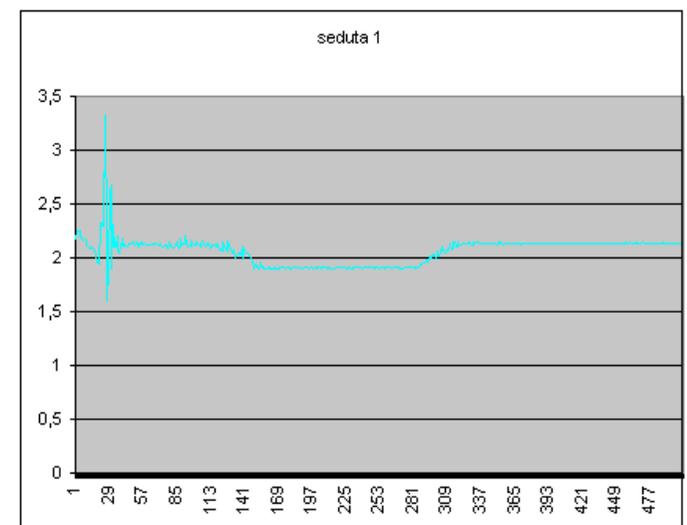
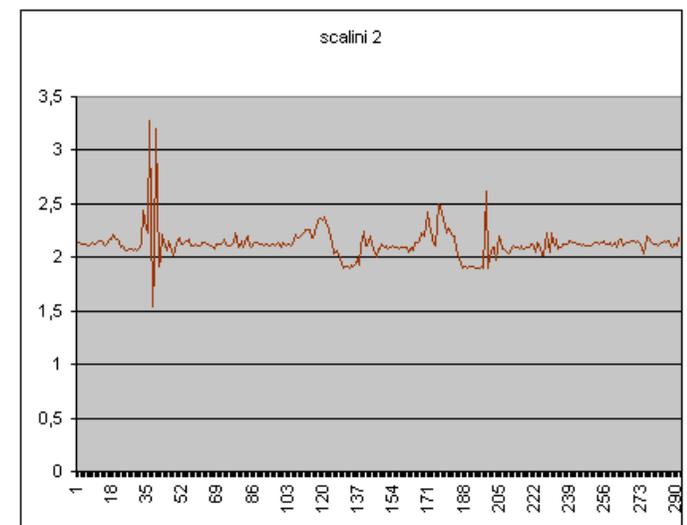
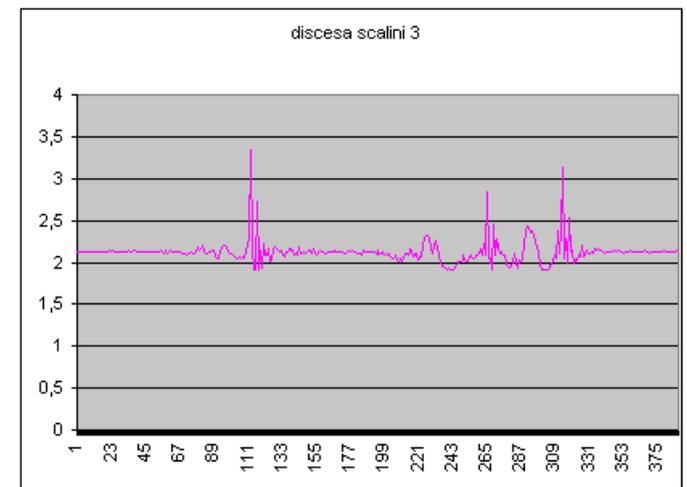
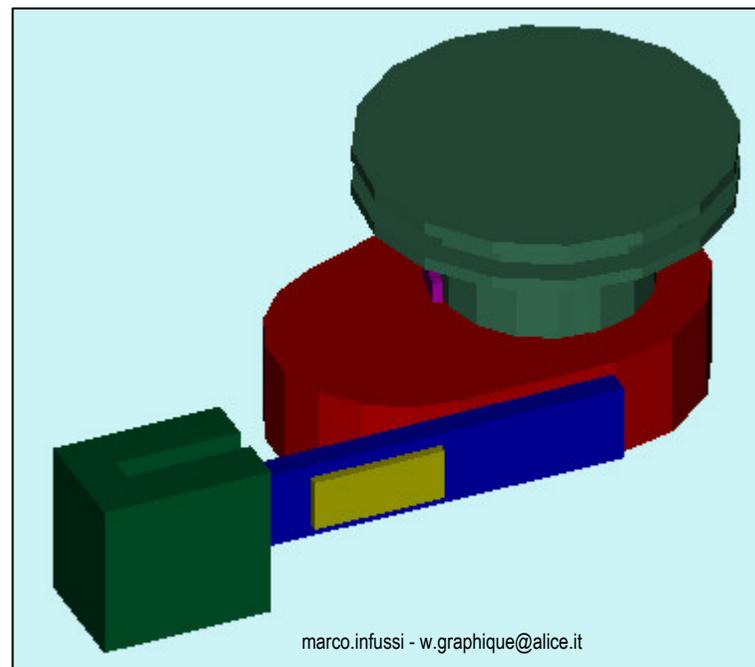
Proposta per un tipo di sensore basato sul principio di funzionamento della lamina estensimetrica

L'unico inconveniente collegato all'utilizzo di una lamina estensimetrica è l'impossibilità di misurare rotazioni superiori a 90°, dovuta alla posizione all'interno del ginocchio. Tra tutti i sensori provati il più affidabile e preciso è risultato essere il potenziometro a filo che presentava come unici svantaggi un certo ingombro ma soprattutto un costo elevato se rapportato al tipo di protesi che si deve realizzare. Si è pensato di proporre come sensore un qualcosa che riunisse i vantaggi offerti dalla lamina estensimetrica, funzionasse dal punto di vista dei collegamenti esterni come il potenziometro a filo ma fosse comunque compatto e soprattutto avesse un basso costo.

Da qui l'idea di realizzare un eccentrico a camma che muovesse durante la sua rotazione una lamina di ferro su cui è installato un estensimetro.

In figura si distinguono:

- in rosso l'eccentrico a camma;
- in blu la lamina di metallo;
- in giallo l'estensimetro;
- in verde chiaro il rocchetto su cui avvolgere il filo;
- in verde scuro il supporto della lamina;
- in viola un semianello.



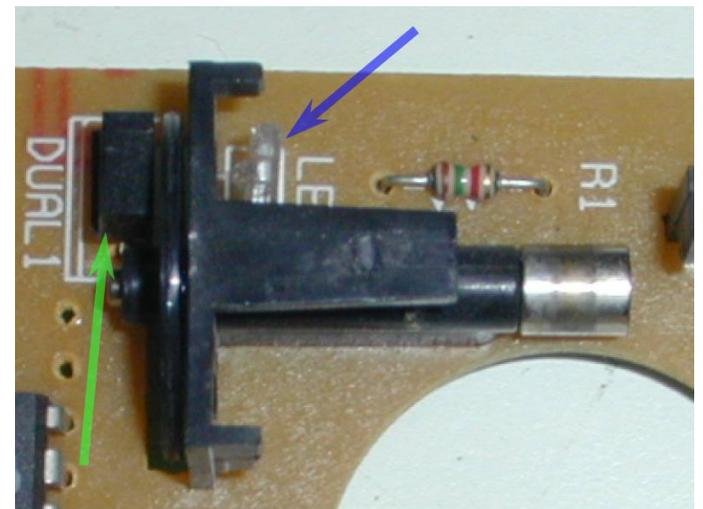
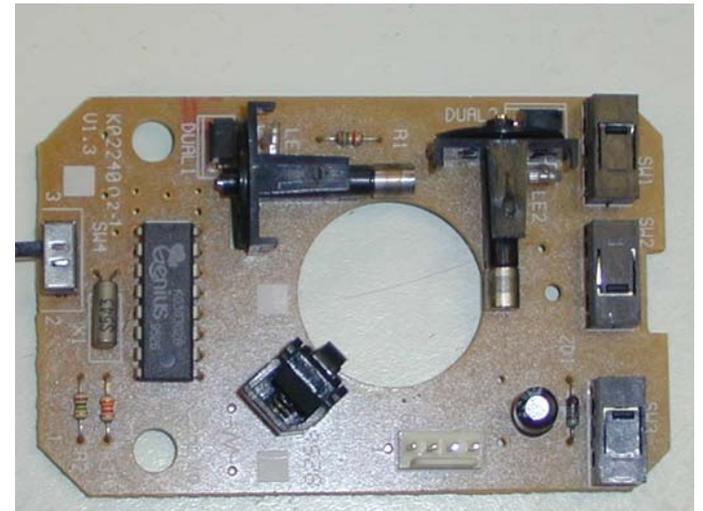
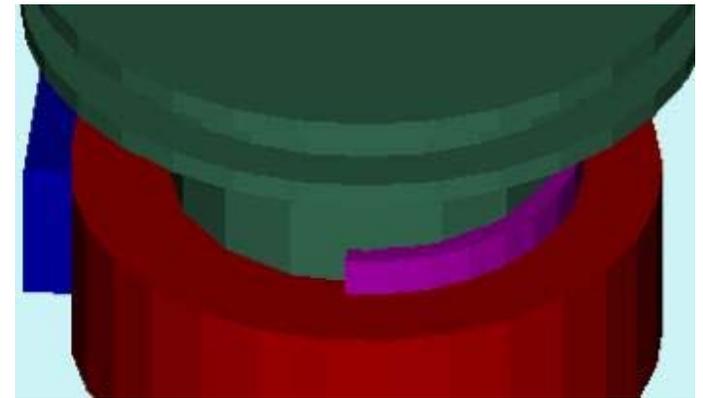
La camma ha un angolo di funzionamento utile pari a 180° . Il primo obiettivo è quello di limitare la rotazione della camma sostanzialmente a mezzo giro. E' per questo motivo che è stato previsto il rocchetto su cui avvolgere il filo. Infatti il diametro di tale rocchetto andrà calcolato in modo che durante la completa estensione del ginocchio la camma non ruoti più di 180° . In questo caso facendo funzionare la camma su tutto l'angolo utile si avrebbe una perdita di biunivocità dell'informazione. Ad ogni indicazione dell'estensimetro potrebbero corrispondere 2 diverse posizioni. E' per questo che è stato previsto un semianello di metallo. A contatto con questo semianello può essere montato un semplice fine-corsa in modo tale che esso rimanga chiuso fino a quando la camma ha ruotato per 90° per poi scattare quando la camma supera tale angolo. L'informazione andrà a settare di un bit di flag sulla scheda di controllo della gamba. In questo modo è quindi possibile usare la camma per tutto il suo angolo utile. Le dimensioni di questo sensore dovrebbero essere compatte in quanto l'unico elemento che potrebbe avere dimensioni più elevate è il rocchetto su cui avvolgere il filo.

Tra i possibili difetti vi potrebbe essere quello dell'utilizzo del fine corsa in quanto l'informazione sulla cima della camma potrebbe non essere precisa per il problema della commutazione del fine corsa. La sensibilità andrebbe verificata dal punto di vista pratico. Dal punto di vista teorico dovrebbe essere paragonabile a quella della lamina estensimetrica montata direttamente sul ginocchio senza avere però un limite dato dalla limitazione dell'angolo massimo di lavoro pari a 90° . Il materiale di costruzione non è specifico. La lamina è di metallo mentre la camma può essere in materiale plastico o anch'essa di metallo; occorrerebbe eventualmente analizzare il rapporto tra peso e usura.

Sensore per la misura della rotazione del ginocchio

Sempre nell'ottica della riduzione dei costi e della semplicità si pensa di utilizzare un sensore che funzioni sullo stesso principio del mouse del computer. Si passa prima di procedere oltre, a spiegare il funzionamento di tale dispositivo. Il principio di funzionamento del mouse è molto semplice: è riportata una visione d'insieme della circuiteria interna di un mouse. Per rilevare il movimento di rotazione si utilizzano due fotodiodi, uno con la funzione di emettitore e l'altro con la funzione di ricevitore. Quando il fascio luminoso tra emettitore e ricevitore è interrotto il segnale logico dell'uscita cambia stato. Nella foto è indicato con una freccia verde il diodo ricevitore e con quella blu il diodo emettitore.

Come mostrato in figura, utilizzando un disco con tanti settori che vanno ad interrompere il raggio luminoso tra emettitore e ricevitore è possibile misurare la rotazione. Nel cerchio blu in figura è evidenziato ancora il ricevitore. Ovviamente l'informazione è discretizzata ed è tanto più precisa quanto maggiore è il diametro del disco e il numero di settori. Partendo da uno stato noto e contando gli impulsi è possibile risalire all'angolo di rotazione. Nel mouse sono presenti due dispositivi di questo tipo, uno che misura la rotazione lungo l'asse x e l'altro che la misura lungo l'asse y. Il principio è lo stesso di quello degli encoder e quindi come negli encoder di tipo incrementale si ha che per poter



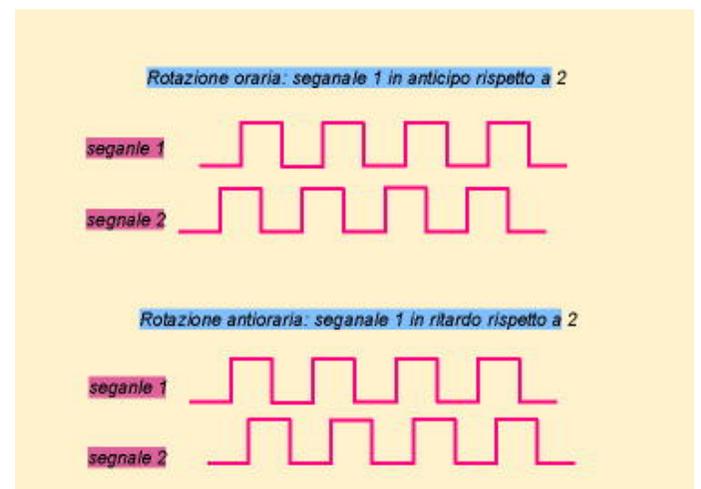
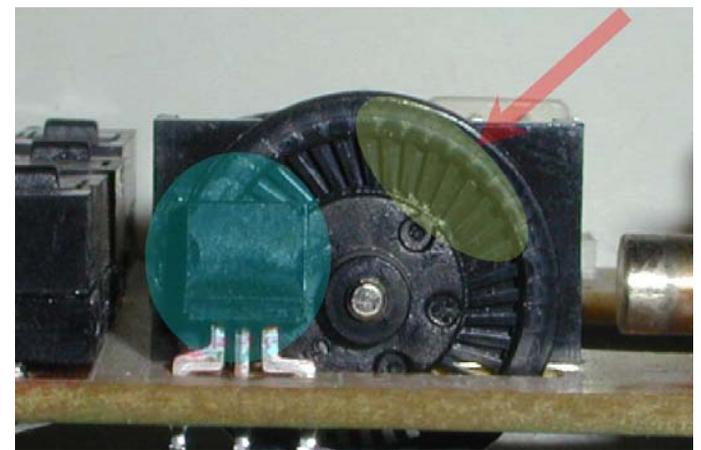
risalire al senso di rotazione occorrono due treni di impulsi sfasati tra di loro di 1/4 di periodo. Vedendo quale dei due è in anticipo è possibile risalire al senso di rotazione.

Un'altro problema comune agli encoder di tipo incrementale è la impossibilità di determinare la posizione al momento dell'accensione ma solo la variazione rispetto a questa posizione. Talvolta gli encoder hanno anche un bit di zero che è usato per avere un riferimento iniziale. Soprattutto negli azionamenti elettrici, il rotore viene fatto ruotare fino a quando non si rileva il bit di zero. A questo punto è possibile fare anche una misura di posizione e non solo di velocità. Esistono encoder di tipo assoluto che hanno una uscita codificata su n bit e pertanto la posizione è univoca in quanto ogni settore ha una diversa codifica. Sono molto più costosi di quelli incrementali. Nell'uso informatico la non conoscenza della posizione iniziale non è un problema, poiché all'accensione del computer il cursore viene posizionato al centro dello schermo. Nell'uso della protesi questo potrebbe essere un limite. Infatti anche utilizzando il bit di zero non potrei risalire alla posizione se non facendo ruotare completamente la gamba alla ricerca del segnale del bit di zero.

Poiché nel mouse non si ha il bit di zero, per ovviare al problema si potrebbe pensare al momento dell'accensione della protesi di portare la gamba in una posizione determinata e settare il sensore in quel punto mediante un pulsante di set. In questo modo a partire da quella posizione sarebbe possibile risalire a tutte le posizioni. Sia in un caso che nell'altro è comunque necessario far ruotare il ginocchio per trovare la posizione iniziale. Un sensore di questo tipo appare semplice ed economico. La precisione non è sicuramente comparabile con quella del potenziometro a filo, ma si ritiene accettabile anche perché il controllo verrà realizzato mediante logica fuzzy, quindi non è richiesta una precisione elevatissima.

***I sensori piezoelettrici offrono ancora prestazioni uniche
articolo tratto da elettronica oggi n 291 giugno 2000***

L'effetto piezoelettrico fu scoperto da Pierre e Jacques Curie nel 1880, ma rimase una mera curiosità scientifica fino al 1940 circa; la proprietà che hanno alcuni cristalli di generare carica elettrica se sottoposti a stress meccanico non ha avuto infatti alcuna utilità pratica fino alla comparsa degli amplificatori ad alta impedenza d'ingresso. Il principio che sta alla base dell'amplificatore di carica è stato brevettato da W.P. Kistler nel 1950 e ha trovato ampio utilizzo negli anni 60. La comparsa dei dispositivi Mosfet a stato solido e lo sviluppo di materiali ad alto isolamento quali il Teflon e il Kapton hanno contribuito a migliorare sensibilmente le prestazioni dei sensori piezoelettrici, che oggi trovano posto virtualmente in tutte le aree della moderna tecnologia. I sensori piezoelettrici sono sistemi elettrici "attivi"; in altre parole, i cristalli producono un'uscita elettrica solo quando si verifica una variazione nel carico (stress) meccanico. Per questa ragione, non sono in grado di effettuare misure statiche nel vero senso della parola; non è comunque corretto affermare che la strumentazione piezoelettrica è in grado di effettuare solo misure dinamiche: trasduttori a quarzo, accoppiati ad adeguati circuiti per il condizionamento del segnale, possono effettuare misure quasi statiche in modo



eccellente e con ottima precisione, su intervalli di minuti e finanche di ore. I sensori piezoelettrici sono utilizzati ampiamente sia nei laboratori che nelle linee di produzione, in tutte quelle applicazioni che richiedono misure accurate dei cambiamenti dinamici di grandezze meccaniche quali la pressione, la forza e l'accelerazione. La lista delle possibili applicazioni continua a crescere, e a tutt'oggi comprende l'aerospazio, la balistica, la biomeccanica, l'ingegneria meccanica e strutturale.

Come funzionano: la maggior parte dei trasduttori piezoelettrici utilizza il quarzo come sensore, in virtù delle sue eccellenti caratteristiche: elevata resistenza allo stress meccanico (20.000 psi), resistenza a temperature fino a 500 C, alta rigidità, alta linearità, isteresi trascurabile, sensibilità costante in un ampio range di temperature e bassissima conducibilità. Sono comunque in uso anche elementi piezo-ceramici e micromachine in silicio. I trasduttori in quarzo consistono essenzialmente in sottili lastre di cristallo tagliate in modo opportuno (rispetto agli assi del cristallo), a seconda della specifica applicazione; il cristallo genera un segnale (una carica da pochi picroCoulomb) che è proporzionale alla forza applicata. Il meccanismo piezoelettrico è semplice: quando al cristallo viene applicata una forza meccanica, le cariche elettriche si spostano e si accumulano sulle facce opposte. La forma e le dimensioni del cristallo dipendono dall'applicazione specifica; gli accelerometri in genere dispongono di una massa aggiuntiva (detta "sismica").

Esistono due diverse soluzioni: a pressione e a compressione. La configurazione a compressione ha il vantaggio di un'alta rigidità, che la rende adatta per la rilevazione di pressioni e forze ad alta frequenza. Alla semplicità del sistema a flessione si contrappone il limitato range di frequenze di funzionamento e la bassa tolleranza ai sovraccarichi meccanici. La configurazione "shear" (o a taglio) è quella tipicamente adottata negli accelerometri in quanto offre il giusto equilibrio tra range di frequenze utilizzabili, bassa sensibilità alle sollecitazioni fuori asse, bassa sensibilità alle sollecitazioni della base e bassa deriva termica. Il taglio del cristallo è spesso oggetto di brevetto; la maggior parte dei sensori Kistler, ad esempio, includono un elemento che è sensibile a carichi sia di compressione che shear; altri tagli "specializzati" includono il taglio trasversale (per alcuni trasduttori di pressione) e il taglio "polistabile" per trasduttori di pressione ad alta temperatura.

La struttura tipica di sensori di forza, pressione e accelerazione è la seguente: è presente un contenitore del sensore, il cristallo piezoelettrico e l'elettrodo dove si localizza la carica generata per effetto piezoelettrico prima che sia condizionata dall'amplificatore. L'accelerometro incorpora anche una massa. Non c'è molta differenza nella configurazione interna tra i diversi tipi di sensore. Negli accelerometri, che misurano la variazione di moto, la massa sismica viene costretta dal cristallo a seguire i movimenti della base e della struttura a cui è attaccata. Dalla forza risultante è possibile ottenere l'accelerazione tramite la legge di Newton $F=m \cdot a$. I sensori di pressione e forza sono quasi identici, e in entrambi i casi la forza viene applicata direttamente sul cristallo; differiscono principalmente nel fatto che i sensori di pressione utilizzano un diaframma per convertire la pressione (che è forza per unità di area) in forza. A causa delle suddette analogie, i sensori progettati per misurare una grandezza spesso sono sensibili anche ad altre sollecitazioni. Questa sensibilità è spesso indesiderata, e può essere ridotta con una progettazione accurata: nei sensori di pressione, ad esempio, può essere introdotto un elemento di compensazione per ridurre la sensibilità all'accelerazione; detto elemento altro non è che un accelerometro opportunamente calibrato connesso in serie al sensore di pressione, con polarità opposta. Un altro problema è la deriva termica del cristallo; questa può essere compensata mediante l'utilizzo di amplificatori con caratteristica termica opposta.

L'amplificatore: il segnale (una carica elettrica) generato dal cristallo va amplificato utilizzando un amplificatore ad elevata impedenza d'ingresso. Il dispositivo attivo è un amplificatore di tensione ad alto guadagno (in configurazione invertente) con ingresso a Fet o Mosfet. In sostanza, l'amplificatore agisce da integratore di carica, compensando la carica del trasduttore con una carica di valore uguale ma polarità opposta, producendo così una tensione ai capi del condensatore Cr. In linea di principio, più che di amplificatore si dovrebbe quindi parlare di convertitore carica-tensione. I due parametri più importanti di cui tenere conto nell'utilizzo pratico degli amplificatori di carica sono la costante di tempo e il drift. La costante di tempo è definita come il tempo di scarica di un circuito accoppiato in AC: applicando in ingresso un gradino,

trascorso un intervallo pari alla costante di tempo la tensione in uscita si riduce al 37% del suo valore iniziale. Il drift è definito come la variazione (indesiderata) del livello di uscita nel tempo, quando il segnale applicato in ingresso è costante. Dei due effetti uno sarà dominante: con ingresso "nullo" (forza sul cristallo costante) l'uscita dell'amplificatore di carica saturerà alla tensione di alimentazione (se a dominare sarà il drift) o decadrà a zero (alla velocità della costante di tempo). La costante di tempo ha inoltre un impatto importante sulle prestazioni del sensore nel suo complesso; in generale, più lunga è la costante di tempo, migliore è la risposta del sensore in bassa frequenza. Nella misura delle vibrazioni, l'amplificatore si comporta sostanzialmente come un filtro passa alto a singolo polo: se si vogliono misurare eventi impulsivi, la costante di tempo deve essere almeno 100 volte più grande della durata totale dell'evento stesso; in caso contrario, la componente in continua del segnale d'uscita va a zero prima che l'evento sia terminato.

Alta o bassa impedenza? Il progredire della microelettronica ha reso possibile l'integrazione dell'amplificatore nel sensore; i dispositivi di questo tipo sono classificati come sensori a "bassa impedenza", ed offrono una serie di vantaggi tra i quali ricordiamo: 1) la disponibilità in uscita di un segnale di tensione a bassa impedenza, compatibile con la maggior parte della strumentazione di misura; 2) la possibilità di trasmettere il segnale in cavi di notevole lunghezza senza degrado delle prestazioni; 3) la possibilità di caratterizzare in modo molto preciso il sensore, le cui prestazioni non dipendono più dall'elettronica cui è collegato. Dal canto loro i sensori cosiddetti ad "alta impedenza" - quelli cioè che non integrano l'amplificatore - sono di utilizzo più complesso, in quanto: 1) hanno bisogno di un circuito esterno di condizionamento e conversione del segnale; 2) a causa dell'elevata impedenza d'uscita sono più sensibili ai disturbi ambientali quali accoppiamenti parassiti, interferenze elettromagnetiche, ecc.; 3) richiedono speciali cablaggi a basso rumore e bassa perdita, che necessariamente non possono essere molto lunghi. Ma i sensori ad alta impedenza offrono anche alcuni vantaggi rispetto a quelli a bassa impedenza, che possono consigliarne l'impiego in applicazioni particolari: non integrando elettronica, sono utilizzabili anche a temperature molto elevate (fino a 500 gradi centigradi); inoltre, le caratteristiche in termini di sensibilità e risposta in frequenza possono essere modificate agendo sull'amplificatore esterno. In conclusione, i sensori piezoelettrici possono offrire prestazioni uniche rispetto ai sensori realizzati con altre tecnologie; la convenienza del loro utilizzo va valutata tenendo conto dei vantaggi (l'ampio range di frequenze e di temperature di lavoro) e degli svantaggi (non effettuano misure statiche) alla luce della specifica applicazione.

Sensori commerciali: Measurement Specialties, Inc

Una prima applicazione dell'effetto piezoelettrico si ebbe nel 1920 da Langevin che realizzò un emettitore ed un ricevitore funzionante sott'acqua, il primo sonar. Durante la guerra si scoprì che alcune ceramiche potevano avere proprietà piezoelettriche se sottoposte ad una elevata tensione di polarizzazione con un processo simile alla magnetizzazione dei materiali ferromagnetici. Più tardi sono stati scoperti effetti piezoelettrici nelle ossa delle balene e nei tendini. Questo ha portato la ricerca ad orientarsi verso materiali organici che potessero avere proprietà piezoelettriche. Nel 1969 è stata scoperta una elevata attività piezoelettrica nel PVDF (polarized fluoropolymer polyvinylidene fluoride) superiore a quella di altri materiali come il nylon e il PVC. Il PVDF è altamente piroelettrico e produce una carica elettrica come risposta alla variazione della temperatura. Assorbe molto l'energia ad infrarosso ed è quindi utilizzato per la costruzione di sensori all'infrarosso in sistemi di visione quali telecamere.

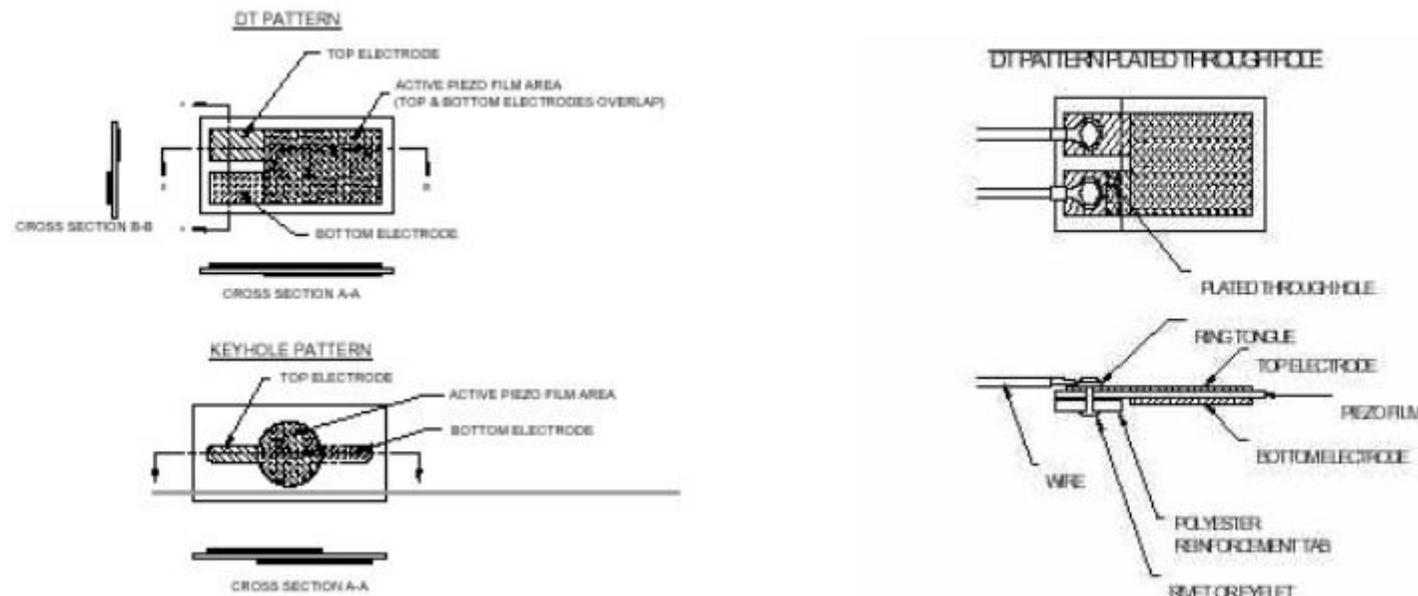
Proprietà della pellicola piezoelettrica: è molto flessibile e leggera. Le principali caratteristiche sono:

ampio spettro di frequenza da 0.001 a 10^9 Hz;
range di utilizzo molto ampio (da 10^{-8} a 10^6 psi);
bassa impedenza acustica;
elevata tensione di uscita;

alta capacità dielettrica;
elevata resistenza meccanica;
elevata stabilità ;
può essere fabbricato in forme particolari.

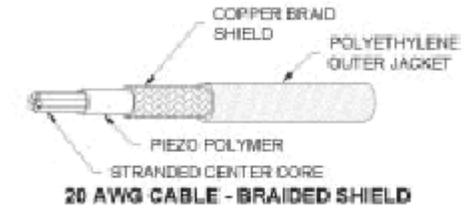
Uno dei principali vantaggi è la bassa impedenza acustica simile a quella dell'acqua, dei tessuti umani e di altri componenti organici. Questo permette ad esempio una ottima trasduzione dei segnali nell'acqua. Il principale limite è una minore capacità di trasmissione elettromeccanica se comparata con la ceramica soprattutto a frequenze basse e di risonanza. Tali dispositivi possono essere utilizzati per convertire energia elettrica in meccanica, ad esempio sono impiegati nella costruzione di altoparlanti. Maggiore interesse (a fini puramente sensoristici) desta la capacità di convertire energia meccanica in energia elettrica. Concettualmente la pellicola piezoelettrica funziona come un estensimetro con la differenza che non ha bisogno di essere alimentato e fornisce un segnale in uscita grande quanto quello dell'estensimetro amplificato. La risposta in frequenza non è limitata dalla necessità di avere elevati guadagni. Uno dei vantaggi è la possibilità di trasdurre elevate forze e pressioni anche utilizzando pellicole di bassissimo spessore. Inoltre a differenza degli estensimetri le dimensioni della pellicola non sono limitate e si può monitorare un'area maggiore. Analizzando la caratteristica piroelettrica si può dire che per notare una variazione del segnale di uscita la variazione di temperatura deve essere rapida ed elevata. Normalmente tale effetto può essere trascurato. Elettricamente la pellicola può essere vista come un carico capacitivo la cui frequenza di taglio è proporzionale a $1/RC$. Elevate capacità aumentano la banda passante ma ritardano la risposta. Tale caratteristica va considerata anche quando si usa la pellicola come sensore in quanto una pellicola di elevato spessore genera un'elevata tensione ma ha una capacità piccola. Si preferisce usare più pellicole di piccolo spessore anziché una di spessore elevato. Quando una pellicola non è sollecitata funziona da carico capacitivo per l'altra.

Connessione elettrica alla pellicola piezoelettrica: di seguito sono riportate diverse tecniche di connessione alla pellicola.

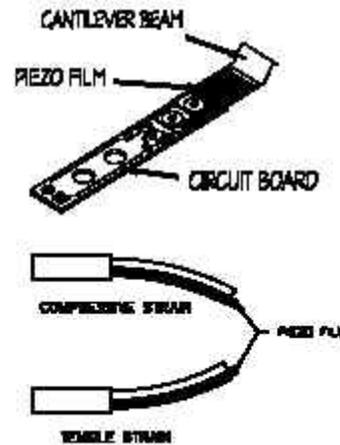


Presentazione di alcuni sensori piezoelettrici: si analizzano alcuni sensori piezoelettrici costruiti e commercializzati dalla MSI.

Piezo cable. Concettualmente si basa sullo stesso principio della pellicola piezoelettrica. In apparenza sembra un comune cavo coassiale ma come si vede dalla figura ha al suo interno uno strato di materiale piezoelettrici avvolto su anima centrale e protetto da treccia di protezione. Il tutto è ricoperto da una guaina in polietilene. Le applicazioni sono varie e vanno dalla monitoraggio delle condizioni del traffico su strada, all'utilizzo come sistema di sicurezza fino all'uso nei touch panel. La risposta del cavo è lineare con la sollecitazione applicata. Uno dei suoi principali vantaggi è la possibilità di avere elevate lunghezze e quindi di poter monitorare elevate aree. Inoltre in virtù della sua forma può essere collocato in zone impervie.



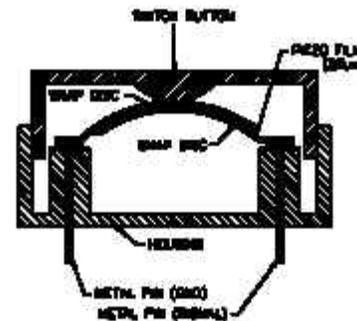
Interruttore. Praticamente ha le caratteristiche di un interruttore mosfet che non assorbe energia durante la commutazione. Ha una tensione di blocco fino a 50 V. E' stato testato fino a 10 milioni di cicli. E' utilizzato come contattore sulle linee di assemblaggio, interruttore, sensore di impatto.



Beam Switch. E' costituito da due lamine di piezo film. Grazie alla sua struttura può misurare sia l'ampiezza che la direzione della sollecitazione. Si può notare che quando la sollecitazione arriva lungo l'asse neutro il segnale dovuto alla compressione di una lamina è uguale

a quello dell'estensione dell'altra.

Sensori di impatto. Devono avere elevata velocità di risposta.



Sensori di traffico. Si utilizzano cavi piezoelettrici che in base alla frequenza delle sollecitazioni misurate permettono di risalire alle condizioni del traffico.

Parameter	Units	Value
Capacitance @ 1kHz	pF/m	600
Tensile Strength	MPa	60
Young's Modulus	GPa	2.3
Density	kg/m ³	1890
Acoustic Impedance	MRayl	4.0
Relative Permittivity @1kHz		9
tan δ ₁ @1kHz		0.017
Hydrostatic Piezo Coefficient	pC/N	15
Longitudinal Piezo Coefficient	Vm/N	250 x 10 ⁻⁶
Hydrostatic Piezo Coefficient	Vm/N	150 x 10 ⁻⁶
Electromechanical Coupling	%	20
Energy Output	mJ/Strain (%)	10
Voltage Output	kV/Strain (%)	5

Dati tecnici

Figure 27a. Sensitivity vs. load

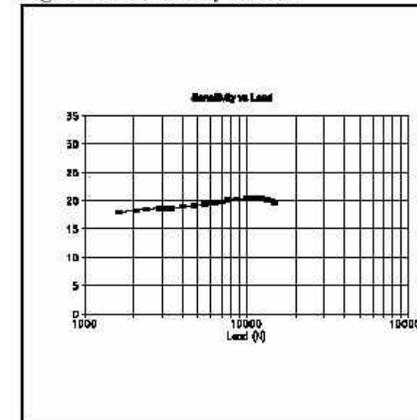
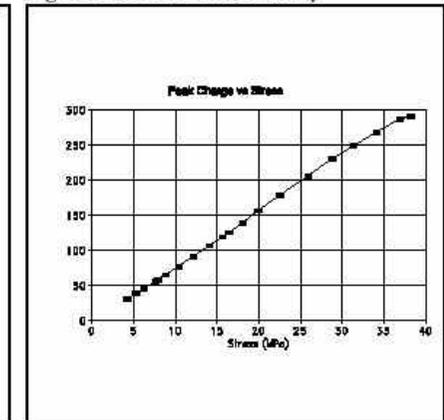


Figure 27b. Piezo cable linearity



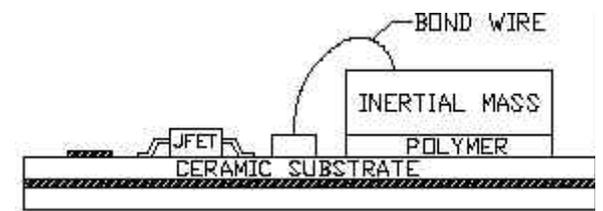
Risposta lineare alle sollecitazioni

Accelerometro. Sono progettati sia usando un materiale tradizionale della ceramica, sia polimeri piezoelettrici. La MSI produce sia accelerometri a compressione che a torsione. Si differenziano per il fatto che quelli a compressione hanno una frequenza di risonanza più elevata, mentre quelli a torsione sono in grado grazie a una particolare tecnologia della MSI a misurare contemporaneamente l'accelerazione lungo le tre direzioni. Tra le principali caratteristiche si segnalano: elevata risposta in frequenza; bassa sensibilità trasversale; eccellente risposta di fase; ampio range di temperatura; ampio range dinamico; bassa dipendenza dalla temperatura; segnale di uscita da 3V a 40V; bassissima impedenza di uscita; frequenza di risonanza molto alta; eccellente linearità.

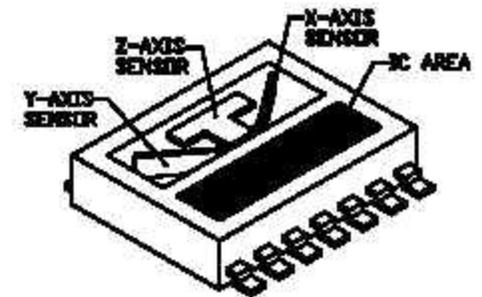
Sensori piezoelettrici utilizzabili sulla gamba

Per quanto riguarda l'applicazione dei sensori piezoelettrici sulla protesi è chiaro che si potrebbero utilizzare accelerometri di tipo piezoelettrico al posto di quello capacitivo. Dai dati tecnici non si evidenziano particolari vantaggi nell'utilizzo dell'uno o dell'altro tipo di sensore. Si ritiene che entrambi i sensori possano soddisfare le specifiche e che solo una questione di prezzo e non tanto di prestazioni a fare la differenza. Sarebbe molto utile collocare una lamina di materiale piezoelettrico sotto il piede della protesi. In questo caso si potrebbe rilevare quale parte del piede è effettivamente in appoggio. L'informazione dovrebbe essere più precisa di quella degli estensimetri perché si può dire che il rilevamento è più diretto. Nell'uso degli estensimetri infatti si risale alla parte del piede appoggiato dalla deformazione della tibia, mentre con il film piezoelettrico si rileva direttamente quale parte del piede è in appoggio. Lo svantaggio più evidente nell'uso di questo tipo di sensore è che non rispetta le specifiche di progetto in quanto andrebbe collocato sul piede (mentre le specifiche impongono che tutta la parte sensoristica sia montata sulla parte centrale della protesi tra ginocchio e tibia). Tuttavia viste le caratteristiche del film piezoelettrico non è errato pensare di collocare questo sensore tra il piede della protesi e la scarpa, come un sottopiede. Sarebbe adatto a tutti i pazienti. Bisogna ricordare il motivo per cui si vuole raggruppare tutta la sensoristica sulla parte centrale: tale parte è più o meno simile per tutti i pazienti, mentre l'invaso, la tibia, il piede possono variare di volta in volta e possono essere modificate in funzione delle esigenze del paziente. esistono ad infatti piedi più adatti per fare sport, ad esempio per correre rispetto ad altri studiati per altri scopi. Inoltre l'invaso è realizzato su misura e quindi ogni volta che si cambia l'invaso andrebbe ricambiato l'accelerometro.

Inoltre accorpate tutto in un unico blocco permette in caso di malfunzionamento di mandare in riparazione solo quella parte e non a tutta la protesi completa. L'utilizzo di un sottopiede realizzato con piezo film non cambierebbe sostanzialmente la realtà delle cose in quanto non dipende dal tipo di piede utilizzato e per la riparazione può essere facilmente rimosso. Bisogna anche considerare che attualmente i sensori piezo film sono di difficile taratura e solo una buona messa a punto darebbe risultati validi. Sarebbe interessante poter confrontare i dati rilevabili dal piezo film con quelli rilevati con l'estensimetro. Non è escluso che la soluzione migliore si avrebbe proprio utilizzando entrambi i sensori e integrando le informazioni.



accelerometro a compressione



accelerometro a torsione

sottopiede con piezo film



pellicola di materiale piezoelettrico

7.3. Total internal reflection

Total internal reflection is an optical phenomenon that occurs when light is refracted (bent) at a medium boundary enough to send it backwards, effectively reflecting all of the light. The critical angle is the angle of incidence above which the total internal reflection occurs.

When light crosses materials with different refractive indices, the light beam will be partially refracted at the boundary surface, and partially reflected. However, if the angle of incidence is shallower (closer to the boundary) than the critical angle, the angle of incidence where light is refracted so that it travels along the boundary, then the light will stop crossing the boundary altogether and instead totally reflect back internally. This can only occur where light travels from a medium with a higher refractive index to one with a lower refractive index. For example, it will occur when passing from glass to air, but not when passing from air to glass.

Optical description

Total internal reflection can be demonstrated using a semi-circular glass block. A "ray box" shines a narrow beam of light (a "ray") onto the glass. The semi-circular shape ensures that a ray pointing towards the center of the flat face will hit the surface at right angles. This prevents refraction at the air/glass boundary. At the glass/air boundary what happens will depend on the angle. Where θ_c is the critical angle:

If $\theta < \theta_c$, as with the red ray, the ray will split. Some of the ray will reflect off the boundary, and some will refract as it passes through. If $\theta > \theta_c$, as with the blue ray, all of the ray reflects from the boundary. None passes through.

The second situation is total internal reflection.

This physical property makes optical fibres useful, and rainbows and prismatic binoculars possible. It is also what gives diamonds their distinctive sparkle, as diamond has an extremely high refractive index.

An important side effect of total internal reflection is the propagation of an evanescent wave across the boundary surface. This wave may lead to a phenomenon known as frustrated total internal reflection.

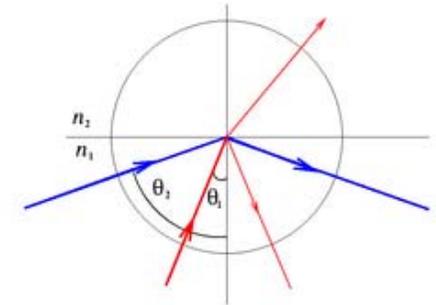
Critical angle

The critical angle is the minimum angle of incidence at which total internal reflection occurs. The angle of incidence is measured with respect to the normal at the refractive boundary. The critical angle θ_c is given by:

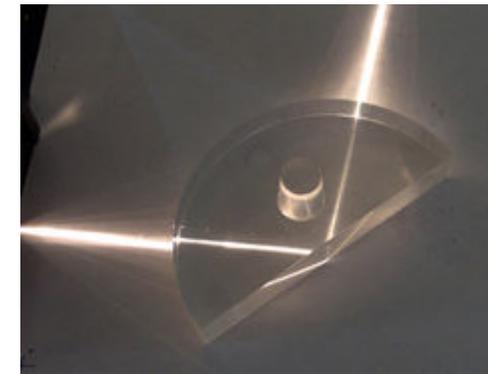
$$\theta_c = \arcsin\left(\frac{n_2}{n_1}\right),$$

where n_2 is the refractive index of the less dense medium, and n_1 is the refractive index of the denser medium. This equation is a simple application of Snell's law where the angle of refraction is 90° .

If the incident ray is precisely at the critical angle, the refracted ray is tangent to the boundary at the point of incidence. For visible light travelling from glass into air (or vacuum), the critical angle is approximately 41.5° . The critical angle for diamond is about 24.4° , which means that light is much more likely to be internally reflected within a diamond. Diamonds for jewelry are cut to take advantage of this; in particular the brilliant cut is designed to achieve high total reflection of light entering the diamond, and high dispersion of the reflected light (known to jewelers as fire).



The larger the angle to the normal, the smaller is the fraction of light transmitted, until the angle when total internal reflection occurs. (The colour of the rays is to help distinguish the rays, and is not meant to indicate any colour dependence.)



Total internal reflection



Frustrated total internal reflection

Under "ordinary conditions" it is true that the creation of an evanescent wave does not affect the conservation of energy, i.e. the evanescent wave transmits zero net energy. However, if a third medium with a higher refractive index than the second medium is placed within less than several wavelengths distance from the interface between the first medium and the second medium, the evanescent wave will be different from the one under "ordinary conditions" and it will pass energy across the second into the third medium.

A common example in everyday use is a beam splitter. A transparent, low refractive index material is sandwiched between two prisms of another material. This allows the beam to "tunnel" through from one prism to the next in a process very similar to quantum tunneling while at the same time altering the direction of the incoming ray.

Applications

Optical fibres, which are used in endoscopes and telecommunications, operate based on total internal reflection of light. Mirages are also formed by total reflection of light traveling from a denser medium (cool air) to a less dense medium (warm air near the ground).

7.3.1. Apple's Frustrated Total Internal Reflection Multi-touch Screen

Interview with Jeff Han at E-Tech, 7 Mar 2006

Consulting research scientist at NYU's department of Computer Science. This stuff is literally just coming out of the lab right now. You're amongst the first to see it out of the lab. I think this is going to change the way we interact with computers. Rear-projected drafting table, equipped with multitouch sensors, ATMs, smart whiteboards, etc. can only register 1 point of contact at the time. Multitouch sensor lets you register multiple touch points, use all fingers, both hands. Multitouch itself isn't a new concept. Played around with multitouch in 80s, but this is very low cost, very high resolution, and very important.

Technology isn't the real exciting thing, more the interactions you can do on top of it once you're given this precise information. For instance, can have nice fluid simulation running. Induce vortice here with one hand, inject fluid with another. Device is pressure sensitive, can use clicker instead of hand. Can invent simple gestures. This application is neat, developed in lab. Started as screen saver, but hacked so it's multitouch enabled. Can use both fingers to play with the lava. Take two balls, merge them, inject heat into the system, pull them apart. This obviously can't be done with single point interaction, whether touch screen or mouse.

It does the right thing, there's no interface. Can do exactly what you'd expect if this were a real thing. Inherently multiuser. I can work in an area over here, and he can be playing with another area at the same time. It immediately enables multiple users to interact with a shared display, the interface simply disappears. Here's a lightbox application: dragging photos around. Two fingers at once, I can start



The Panasonic booth was a sure draw with its Minority Report-like finger-gesture controlled screen. Just point and click into the air, and everything shows up on the screen a meter away.



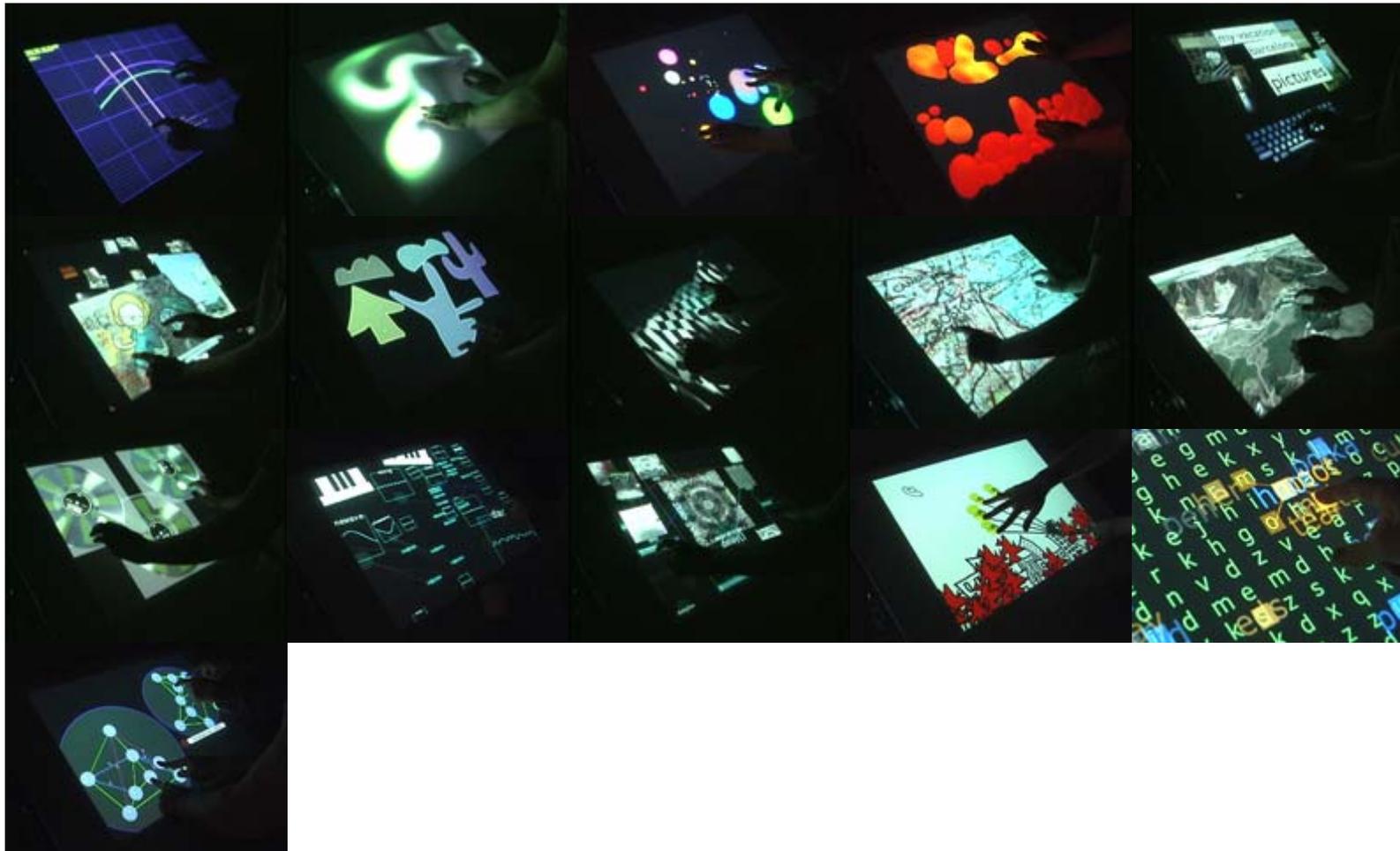
zooming, rotating, all in one really seamless motion. It's neat because it's exactly what you expect would happen if you grabbed this virtual photo here. All very seamless and fluid. Someone who's new to computing culture can use this. Could be important as we introduce computers to a whole new group of people. I cringe at the \$100 laptop with its WIMP interface.

Really simple and elegant technique for detecting touch point, scattered light by deformation caused by touch on screen. Kinaesthetic memory, the visual memory where you left things. Ability to quickly zoom, get a bigger work area if you run out of space, etc. changes things. More of an infinite desktop than standard fixed area. Now, of course, can do the same thing with videos as with photos. All 186 channels of TW cable. Inevitably there'll be comparisons with Minority Report. Minority Report and other gestural interfaces aren't touch based. Can't differentiate between slight hover and actual touch. Disconcerting to user if they have action happen without tactile feedback. I argue that touch is more intuitive than gross gestural things. Also gestural is very imprecise.

Ability to zoom in and out quickly lets you find new ways to explore information. What's interesting is that we're excited about potential for this in information visualization applications. Can easily drill down or get bigger picture. Having a lot of fun exploring what we can do with it. Another application we put together is mapping. This is NASA WorldWind, like Google Earth but Open Source. We hacked it up to use the two fingered gestural interface to zoom in. Can change datasets in NASA Worldwind. They also collect pseudocolour data, to make a hypertext map interface. Three dimensional information, so how do you navigate in that direction. Use three points to define an axis of tilt. Could be right or wrong interface, but example of kind of possibilities once you think outside the box.

Virtual keyboard, rescalable. No reason to conform to physical devices. Brings promise of a truly dynamic user interface, possibility to minimize RSI. Probably not the right thing to do, to launch in and emulate things from the real world. But lot of possibilities, we're really excited. Lots of entertainment applications, multiuser with many people playing in parallel. Here's a simple drawing tool. Can add constraints and have multiple constraints, to make a really asy virtual puppeteer tool. Lot of math under the surface to do what's physically plausible (algorithm published last year at SIGGRAPH).

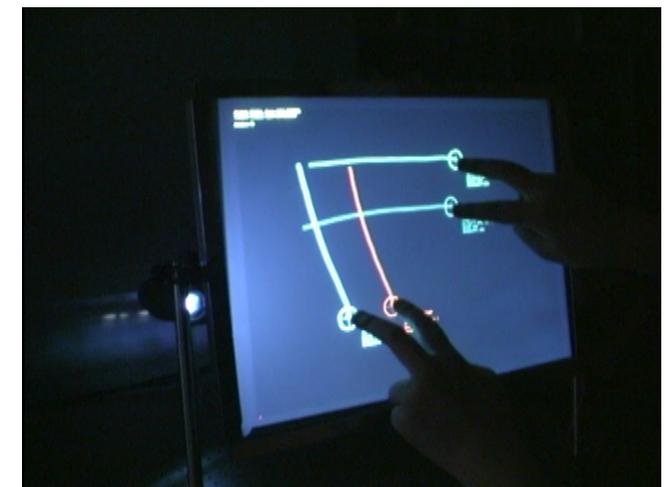
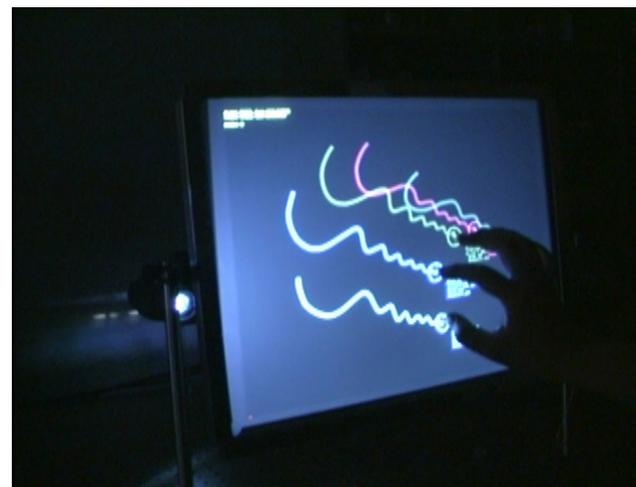
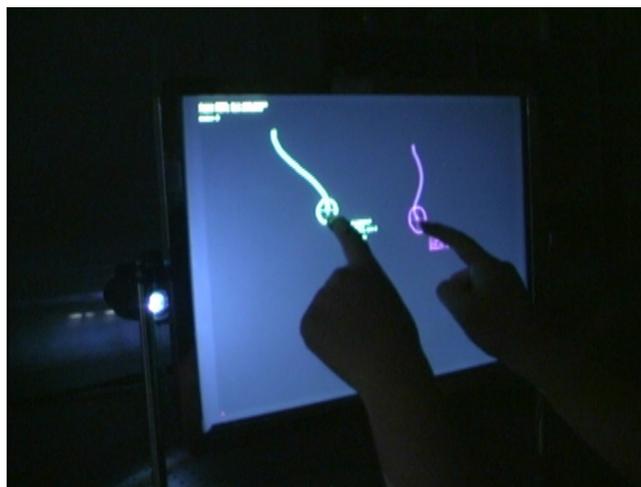
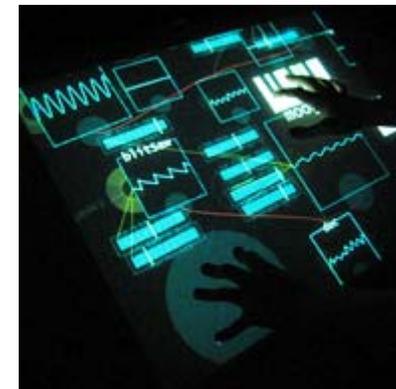
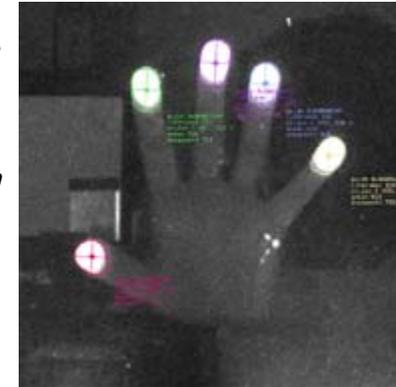
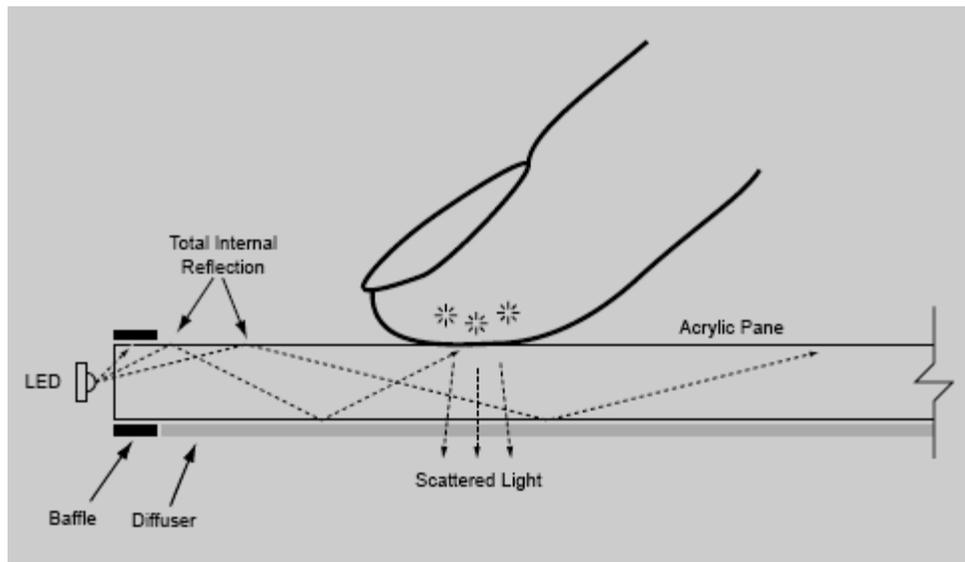
While touch sensing is commonplace for single points of contact, *multi-touch* sensing enables a user to interact with a system with more than one finger at a time, as in chording and bi-manual operations. Such sensing devices are inherently also able to accommodate *multiple users* simultaneously, which is especially useful for larger interaction scenarios such as interactive walls and tabletops. Since refining the *FTIR (frustrated total internal reflection)* sensing technique, we've been experimenting with a wide variety of application scenarios and interaction modalities that utilize multi-touch input information. These go far beyond the "poking" actions you get with a typical touchscreen, or the gross gesturing found in video-based interactive interfaces. It is a rich area for research, and we are extremely excited by its potential for advances in efficiency, usability, and intuitiveness. It's also just so much fun! Our technique is force-sensitive, and provides unprecedented resolution and scalability, allowing us to create sophisticated multi-point widgets for applications large enough to accommodate both hands and multiple users. The drafting table style implementation shown here measures 36"x27", is rear-projected, and has a sensing resolution of better than 0.1" at 50Hz. Stroke event information is sent to applications using the lightweight OSC protocol over UDP. Further development is related to recognizing also the finger used to manipulate objects.



We introduce a simple technique that enables robust multi-touch sensing at a minimum of engineering effort and expense. It relies on frustrated total internal reflection (FTIR), a technique familiar to the biometrics community where it is used for fingerprint image acquisition. It acquires true touch information at high spatial and temporal resolutions, and is scalable to very large installations.

Han, J. Y. 2005. Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*

Bi-manual, multi-point, and multi-user input on graphical interaction surfaces.



7.3.2. Apple applies for patents for touchpad technology

System and method for processing raw data of track pad device

Filed: September 21, 2005

An input device and system are described that acquires (measures) raw track pad sensor data and transmits this data to a host computer where it is analyzed by an application executing on one or more host computer CPUs. The resulting input processing architecture provides a track pad input device that is both lower in cost to manufacture and more flexible than prior art track pad input devices. Also details of gestures for a touchpad.

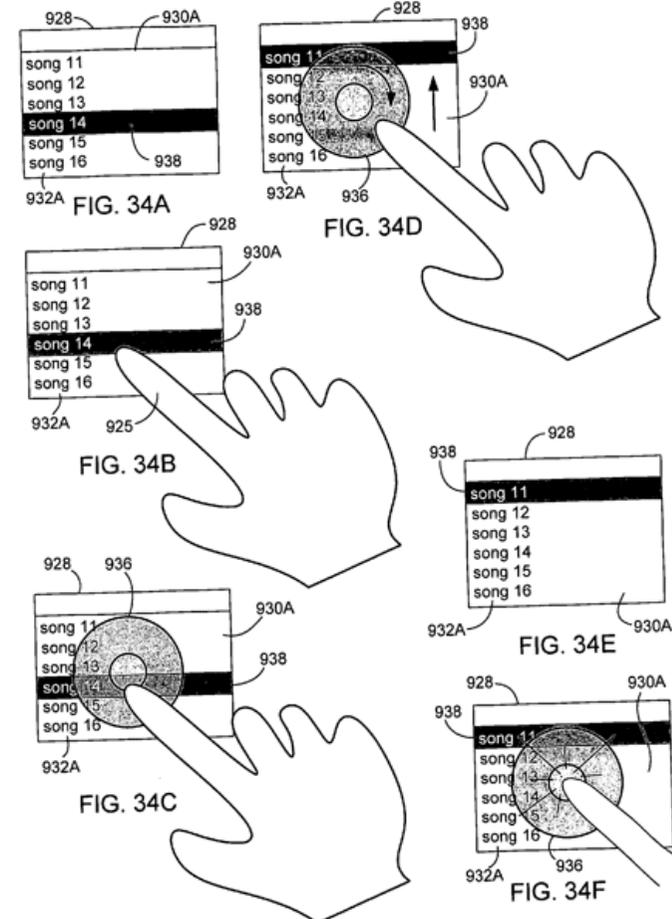
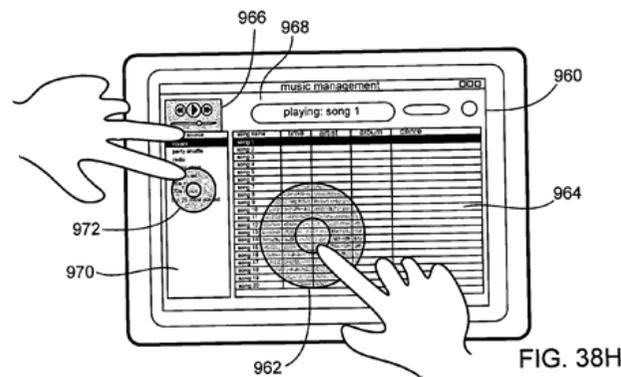
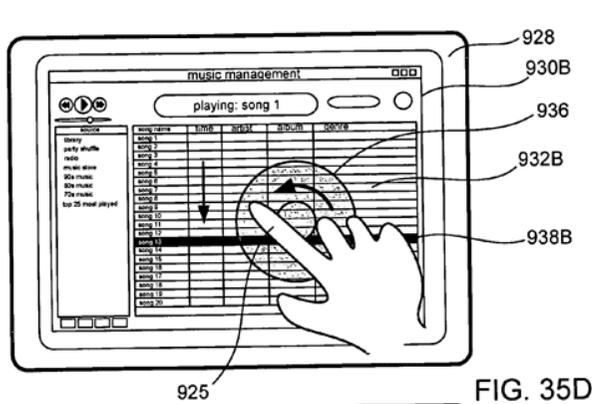
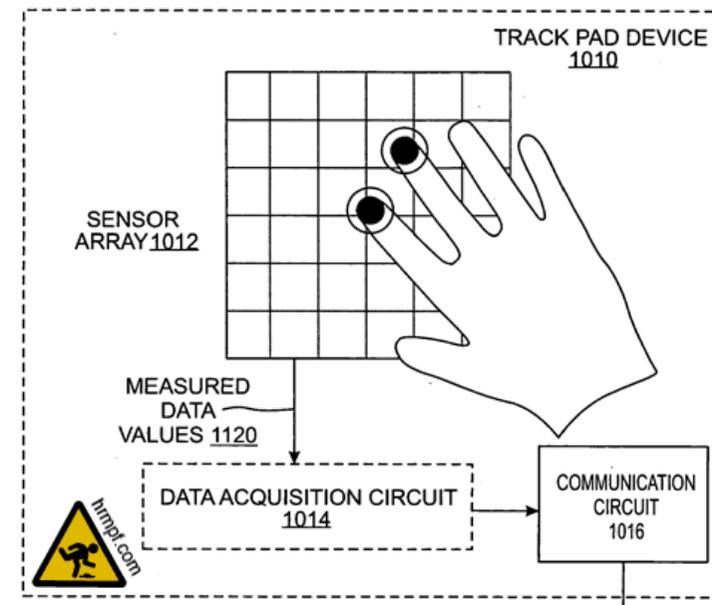
Low EMI capacitive trackpad

Filed: September 24, 2004

A trackpad with resistance to electromagnetic interference - advantageous because it lessens the requirements for shielding and increases the signal:noise. Inventors: Benjamin Lyon, Stephanie Cinereski, Chad Bronstein and Steven Hotelling (not all on all patent applications)

Touch sensitive I-pod

“Here two virtual touch wheels appear on the screen allowing independent action- the left hand has one finger on the volume and one on the smaller virtual scroll wheel allowing volume control and playlist selection, at the same time the right hand is scrolling through available songs in the iTunes right hand pane.”



602  A point operation can be implemented with a first gesture involving touching and moving any two adjacent fingers on the track pad device.

604  A click operation can be implemented with a second gesture involving a tap (momentary touch) of any two adjacent fingers on the track pad device.

606  A double click operation can be implemented with a third gesture involving a tap (momentary touch) of any three adjacent fingers on the track pad device.

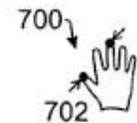
608  A right click operation can be implemented with a fourth gesture involving a tap of the thumb, middle finger, and ring finger on the track pad device.

610  A drag/select operation can be implemented with a fifth gesture involving a touch and movement of three fingers on the track pad device.

612  A scroll operation can be implemented with a sixth gesture involving the touch and slide up/down of four fingers on the track pad device.

614  The system can also be used to emulate operation of a three-button mouse. A click of a middle button can be implemented by a seventh gesture involving a tap of the thumb, index, and middle fingertips on the track pad device.

616  In addition, a right click in a three-button emulation can be implemented by an eighth gesture involving the spreading of the hand and then the tap of the thumb, ring, and pinky fingers on the track pad device.

700  A cut operation can be implemented with a first gesture involving a touch and pinch of the thumb and middle finger on the track pad device.

704  A copy operation can be implemented with a second gesture involving a tap of the thumb and middle finger on the track pad device.

706  A paste operation can be implemented with a third gesture involving a touch and expansion of the thumb and middle finger on the track pad device.

708  An Undo/Redo operation can be implemented with a fourth gesture involving a touch and slide of the thumb and middle finger up/down on the track pad device. Undoing just one operational step can be implemented by sliding quickly, or multiple steps can be implemented by sliding gradually.

710  A Tab/BackTab operation can be implemented with a fifth gesture involving a touch and slide of the thumb and middle finger right/left on the track pad device. Just one tab can be implemented by sliding quickly, and repetitive tabs can be implemented by sliding gradually.

712  An arrows operation can be implemented with a sixth gesture involving a touch with a slide of any finger on the track pad device to move the text cursor.

714  A Text Select operation can be implemented with a seventh gesture involving a touch with a slide up/down of three spread fingers on the track pad device.

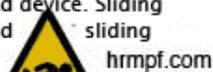
716  A Tab/Back Tab operation can be implemented with an eighth gesture involving a touch of the thumb and middle finger and a slide left/right on the track pad device.

718  A Home/End operation can be implemented with a ninth gesture involving a touch of four spread fingers with slide left/right on the track pad device.

720  Finally, a Page-Up/Page-Down operation can be implemented with a tenth gesture involving a touch of four spread fingers with a slide up/down on the track pad device.

- 802  An open file operation can be implemented with a first gesture involving a touch with counter-clockwise rotation of the thumb and three fingers on the track pad device.
- 804  A close file operation can be implemented with a second gesture involving a touch with clockwise rotation of the thumb and three fingers on the track pad device.
- 806  A save file operation can be implemented with a third gesture involving a touch with contraction of the thumb and three fingers on the track pad device.
- 808  A new file operation can be implemented with a fourth gesture involving a touch with expansion of the thumb and three inner fingers on the track pad device.
- 810  A print file operation can be implemented with a fifth gesture involving pre-spreading of the hand and then a touch with further expansion of the thumb and three outer fingers on the track pad device.
- 812  A next file operation can be implemented with a sixth gesture involving a touch with a slide of the thumb and three fingertips to the left on the track pad device.
- 814  A previous file operation can be implemented with a seventh gesture involving a touch with a slide of the thumb and three fingertips to the right on the track pad device.
- 816  An operation for showing the desktop can be implemented with a eighth gesture involving touching the track pad with spread thumb and three fingers and then sliding to the left on the track pad device.
- 818  An exit application operation can be implemented with a ninth gesture involving a touch with clockwise rotation of spread thumb and three fingers on the track pad device.
- 820  A switch application operation can be implemented with a tenth gesture involving touching the track pad with spread three fingers and thumb and then sliding left or right on the track pad device. Sliding crisply can be used to advance just one window, and gradually can be used to scroll through whole list.

- 900  A back operation can be implemented with a first gesture involving a touch with a slide of the thumb and three fingertips to the left on the track pad device.
- 902  A forward operation can be implemented with a second gesture involving a touch with a slide of the thumb and three fingertips to the right on the track pad device.
- 906  A scroll operation can be implemented with a third gesture involving a touch with a slide of four fingers up/down on the track pad device. If desired, the thumb can be rested on the track pad device after starting the slide on the track pad device.
- 908  A zoom-in operation can be implemented with a fourth gesture involving a touch with expansion of the thumb and four fingers on the track pad device.
- 910  A zoom-out operation can be implemented with a fifth gesture involving a touch with a contraction of the thumb and four fingers on the track pad device.
- 912  A find-in-page operation can be implemented with a sixth gesture involving a touch with a pinch of the thumb and two fingertips on the track pad device.
- 914  A fine scroll operation can be implemented with a seventh gesture involving the "roll" of fingers instead of sliding them on the track pad device. If desired, the thumb can be rested on the track pad device once the scroll starts.
- 916  A pointing operation can be implemented with an eighth gesture involving dropping other fingers to point with all five fingers on the track pad device once pointing starts.
- 918  Additional gestures can include a ninth gesture involving moving all five fingers on the track pad device.
- 920  A tenth gesture 920 involving a touch and movement of a portion (e.g., palm) of a hand on the track pad device.



7.3.3. Other Innovative Interfaces

SIGGRAPH conference showcased a number of innovative human-computer interface solutions. Though the general trend was toward natural interfaces that respond to intuitive human gestures and motions, even Aibo was not discounted as a potential interface solution. Natural interfaces are likely to give the mouse a run for its money sometime soon, but they require robust sensors to accurately track position in real time and respond accordingly. Some examples presented at SIGGRAPH are below:

CirculaFloor. CirculaFloor is a locomotion interface using a group of movable floors. The movable floors employ a holonomic mechanism that achieves omni-directional motion. Circulation of the floors enables users to walk in arbitrary directions in a virtual environment while their positions are maintained.

GelForce. GelForce is a novel interface that measures the distribution of both the magnitude and direction of force. The sensor is composed of a transparent elastic body, two layers of blue and red markers, and a CCD camera. Force vectors are calculated from the captured movement of the markers.

Lumisight Table: Interactive View-Dependent Display Table Surrounded by Multiple Users. Lumisight Table is a novel interactive table display that can display information in each required direction on a shared screen and capture multiple users' gestures simultaneously.

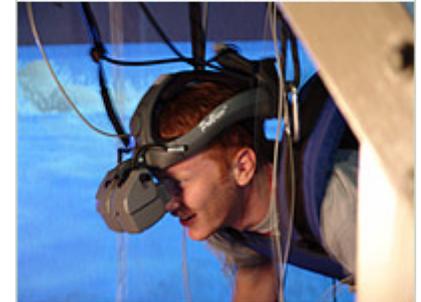
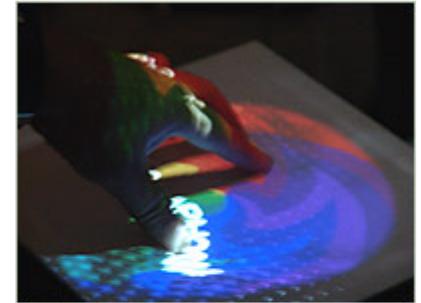
Novel Infrared Touch-Screen Technology and Associated Artwork. This is a novel technique for a touch screen using front-projected infrared with a rear-projected interactive video display. The technology is demonstrated with three art pieces. Two of the art pieces allow participants to create works in the forms of Calder (MPG) and Mondrian. The third is a visualization of radio interferometry.

Untethered Force Feedback Interface That Uses Air Jets. This is an untethered interface that eliminates the annoyance of wires and other connectors by using air jets to establish force feedback. Attendees experience interaction with a virtual object that responds to being "touched." The sense of touch is provided by air jets, while visual clues are provided by a projection-based stereo display.

Phantom Omni Haptic Device. The SensAble Technologies PHANTOM product line of haptic devices makes it possible for users to touch and manipulate virtual objects.

Swimming Across the Pacific. In this virtual swimming apparatus, a swimmer is suspended in a swimming cage wearing a hang-gliding harness, a head-mounted-display, and eight tracking sensors. When the apparatus is installed in an airplane flying across the Pacific Ocean, a performer will swim for the entire duration of the flight, and the plane, a bubble of non-time and non-space, will become a collaborative artistic space.

NetAIBO Project. An unsimultaneous-communication and 3D-space-sharing system using web3D techniques and AIBO as the interface and a proposal for new entertainment content extended from the original function of AIBO and other stand-alone robots.



3D TV

A system for real-time acquisition, transmission, and 3D display of high-definition 1300 x 1030 dynamic content. Image acquisition consists of an array of hardware-synchronized cameras that capture multiple-perspective views of the scene. To deal with the high processing and bandwidth requirements, the system uses a fully distributed architecture with clusters of PCs for sending and receiving. A multi-projector 3D display with horizontal parallax only achieves large-scale, high-resolution output images. The system is scalable in the number of acquired, transmitted, and displayed real-time video streams. The hardware is relatively inexpensive and consists mostly of commodity components that will further decrease in price. The system architecture is flexible to enable a broad range of research in 3D TV. *Life Enhancement.* This system demonstrates that 3D TV offers a richer, more immersive experience than regular TV. It increases entertainment value and realism without the encumbrance of special glasses. *Vision.* Three-dimensional TV is expected to be the next revolution in the history of television. It has only recently become feasible to deal with the high processing and bandwidth requirements for real-time acquisition, transmission, and display of high-resolution 3D TV content. Our system uses existing broadband protocols and compression standards for immediate, real-world 3D TV experiments and market studies. The decoders and receivers are well established and widely available. This system can plug into today's digital TV broadcast infrastructure and co-exist in perfect harmony with 2D TV. Today, digital broadcast networks carry hundreds of channels and presumably a thousand or more channels after the introduction of MPEG-4. This makes it plausible that a number of them (perhaps 16) will be dedicated to 3D TV. Similar to HDTV, the introduction of 3D TV can proceed gradually, with one 3D channel at first and more to follow, depending on market demand. Our system uses 16 projectors to produce a view-dependent display with 12.5 million pixels. We believe that new, high-resolution display technologies -- such as organic LEDs (OLED), field emission devices (FEDs), and color nanotube displays -- will achieve similar or higher pixel resolution in the near future. In conjunction with lenticular screens this will lead to inexpensive, high-resolution 3D displays for consumer devices. Avenues for future work include: high-dynamic-range cameras and displays, precise color reproduction of natural scenes, 3D video conferencing, and deformable 3D displays. *Goals.* The main goal of this work is to show that 3D TV systems can be built today from off-the-shelf components.



Innovations

1. Distributed architecture. In contrast to previous work in multi-view video coding, this system uses distributed processing for acquisition, compression, transmission, and image display.
2. Multi-view video rendering. A new algorithm efficiently renders novel views from multiple dynamic video streams.
3. Scalable system. The system is completely scalable in the number of acquired, transmitted, and displayed views.
4. High-resolution 3D display. This is the first display that provides horizontal parallax with 16 independent perspective views per pixel.
5. Computational Display. The system can render new views to improve the 3D display based on feedback from cameras on both sides of the display.

CirculaFloor

CirculaFloor is a locomotion interface using a group of movable tiles. The movable tiles employ holonomic mechanism that achieves omni-directional motion. Circulation of the tiles enables the user to walk in virtual environment while his/her position is maintained. The user can walk in arbitrary direction in virtual environment. This project is a joint research with ATR Media Information Science Labs.

Basic Design of the CirculaFloor

Locomotion interfaces often require bulky hardware, because they have to carry whole body of the user. Also, the hardware is not easy to reconfigure to improve its performance or add new functions. Considering these issues, the goals of the CirculaFloor project are:

- (1) To develop compact hardware for creation of infinite surface for walking.
- (2) To develop scalable hardware architecture for future improvement of the system.



In order to achieve these goals, we designed a new configuration of locomotion interface by the use of a group of omni-directional movable tiles. Each tile is equipped with a holonomic mechanism that achieves omni-directional motion. Infinite surface is simulated by circulation of the movable tiles. The motion of the feet is measured by position sensors. The tile moves opposite to the measured direction of the walker, so that motion of the step is canceled. The position of the walker is fixed in the real world by this computer-controlled motion of the tiles. The circulation of the tiles has an ability to cancel the displacement of the walker in arbitrary direction. Thus, the walker can freely change direction while walking.

Implementation of a prototype system

The hardware configuration and overview of the CirculaFloor system are shown in Figure 1 and 2. For the movable tiles, we use four holonomic omni-directional vehicles known as the Vmax Carrier made by RITECHS. The size of the movable tile is 568mm (W)x568mm (D)x92mm (H). Each tile weighs 16.2kg, and has ability to carry 80kg payload in 1200mm/s. To avoid getting tangled up the communication lines between the movable tiles and PC for controlling the tiles, the movable tiles and the PC are equipped with wireless RS232c modules. Also for position sensing of the tiles, we use a wireless ultrasonic position sensor known as the IS-600 mark2 made by InterSense. Two position sensors put on each movable tile, so that its position and orientation are measured. The positions of the walker is measured with a Laser range finder, the LMS200, made by SICK so that the user can walk with no sensor attachment. The operating system of the PC is Windows XP, and we developed all of the software using C language with VC++.

Creation omni-directional infinite surface

The method of circulation of the movable tiles can be divided into 2 phases; (1) detection of walking, (2) circulation of movable tiles.

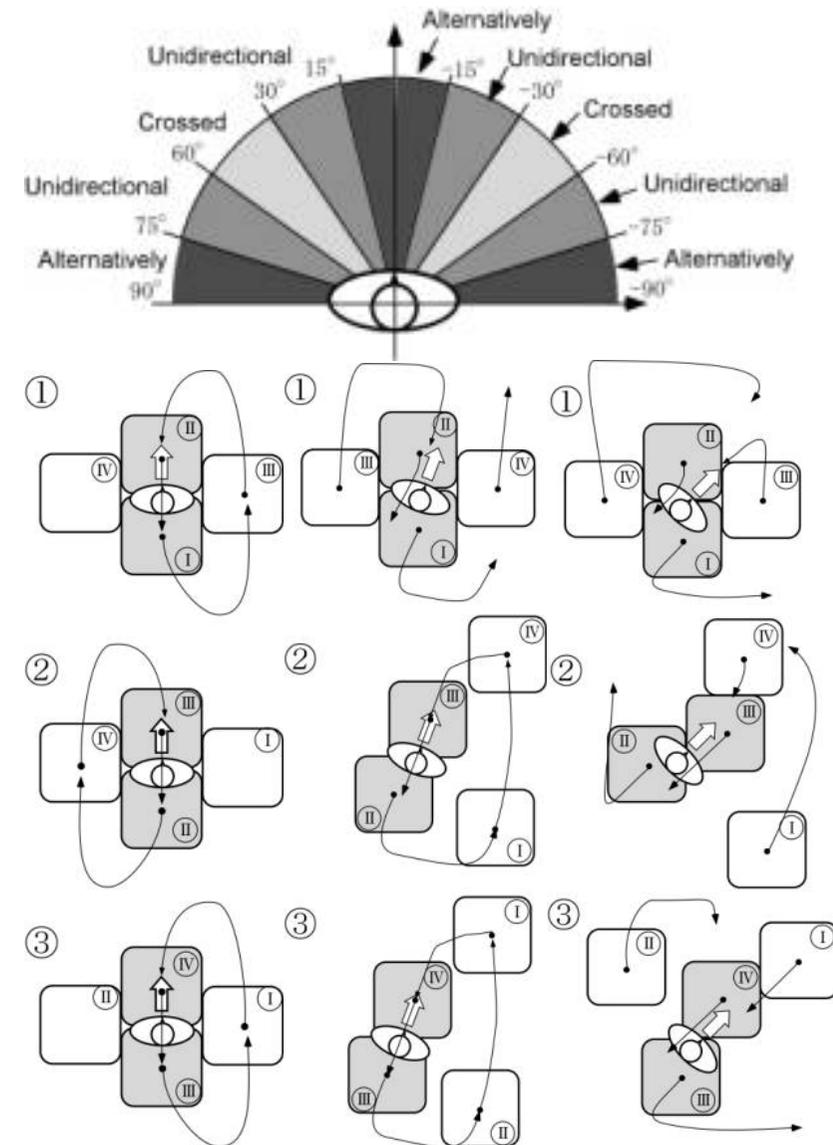
Detection of walking

The user's position in virtual space is updated corresponding to the results of motion tracking of the feet.

Figure 3 illustrates basic idea of the detection of walking. A circular dead zone is placed in the center of the walking area. In this study, the dead zone is 200 mm diameter circle. The positions of the walker's knees are measured by position sensors. The midpoint between these sensors can be assumed to be a central point below the body. Point G in Figure 3 represents the projection of the central position of the walker. The movable tiles don't move while the point G is inside the dead zone. If the point G leaves the dead zone, the tiles move so that the point G returns to this area. The direction and distance between the point G and the circle determines the pulling back direction and velocity of the tiles respectively.

Circulation of movable tiles

Current circulation method of the movable tiles is designed to satisfy following conditions; (1) two of the movable tiles are used for pulling back the user to center of the dead zone. (2) The rest of the movable tiles are used for creation of a new front surface. (3) These tiles move the shortest distances to the next destination, while they avoid colliding other tiles. (4) Control program allocates all destinations of the tiles, when the tiles reach their destinations. (5) The tiles don't rotate corresponding to walking direction for simplification of the algorithm. (6) The walker can change walking direction anytime.



Considering above conditions, the circulation method has to be varied corresponding to the walking direction. Three modes, “alternative circulation”, “unidirectional circulation”, and “cross circulation” are designed corresponding to the direction (Figure 4). Typical motion of each mode is illustrated in Figure 5-7. The velocity of pulling back is proportion to the distance between the dead zone and the point G. Alternative circulation (Figure 5): This mode is adopted to the directions between ± 15 deg and ± 75 - 105 deg. The tiles for creating new front surface (white colored tiles in Figure 5) are sneaking around to the front of the tiles for pulling back (in Figure 5, gray colored tiles) from right /left side alternatively. Unidirectional circulation (Figure 6): This mode is adopted to the directions between ± 15 - 30 deg and ± 60 - 75 deg. The tiles for creating new front surface are sneaking around to the right/left front of the tiles for pulling back with a unidirectional circulation. Cross circulation (Figure 7): This mode is adopted to the directions between $+30$ to $+60$, -30 - -60 deg. The tiles for creating new front surface are sneaking around to the right/left front or the right/left side of the tiles for pulling back. When a user of the CirculaFloor switches over the walking direction, the control program calculates the nearest phase of each tile by using template-matching technique corresponding to the new direction. Then the tiles take the shortest way to their destinations. At this time, the maximum walking velocities are 330 mm/s in alternative circulation mode and unidirectional circulation mode, and 210 mm/s in crossed circulation mode.

Applications of the CirculaFloor

It has often been suggested that the best locomotion mechanism for virtual worlds would be walking. It is well known that the sense of distance or orientation while walking is much better than that while riding in a vehicle. However, the proprioceptive feedback of walking is not provided in most applications of virtual environments. The CirculaFloor is a new locomotion device which provides such a sense of walking. It will make revolution in entertainment or training simulators. One of the serious applications will be an "evacuation simulator." Analysis of evacuation of people in disasters is important in social safety. However, it is impossible to carry out experiments with human subjects during an actual disaster. Virtual environment is inevitable for such experiments. Since evacuation is done by walking or running, the CirculaFloor will be an indispensable interface device for the experiments. Combination of the CirculaFloor and an immersive projection display may provide ultimate sense of presence. The integrated system can greatly contribute to teleoperation or virtual travel.

Feelex

The FEELEX is a combination of an object-oriented-type force display and real-world graphics. A user of a real-world graphics system, such as the DigitalDesk or the Luminous Room space, interacts with rigid objects. On the other hand, the FEELEX system presents deformable objects, just like living creatures. This function provides a new interaction style compared to the object-oriented-type force display and real-world graphics.

In 1995, we started to develop a preliminary implementation of the FEELEX. A rubber screen was put on top of five linear actuators. An image was projected onto the screen and was deformed by the motion of the linear actuators. The basic function of the FEELEX was confirmed by this prototype. We developed further prototypes using down-sized actuators that improved the resolution of the haptic surface. We developed the FEELEX 1 in 1997. It was designed to enable double-handed interaction using the whole of the palms. Therefore, the optimum size of the screen was determined to be 24cm X 24cm. The screen is connected to a linear actuator array that deforms its shape. Each linear actuator is composed of a screw mechanism driven by a DC motor. The screw mechanism converts the rotation of an axis of the motor to the linear motion of a rod. The motor must generate both motion and a reaction force on the screen. The diameter of the smallest motor that can drive the screen is 4cm. Therefore, a 6X6 linear actuator array can be set under the screen. The deformable screen is made of a rubber plate and a white nylon cloth. The thickness of the rubber is 3mm. Fig.1 shows an overall view of the device. The screw mechanism of the linear actuator has a self-lock function that maintains its position while the motor power is off. We learned of the difficulty involved in representing a hard virtual wall from our experiences with the tool-handling-type force display. Considerable motor power is required to generate the reaction force from the virtual wall, which often leads to uncomfortable vibrations. The screw mechanism is free from this problem. A soft wall can be represented by the computer-controlled motion of the



linear actuators based on the data from the force sensors. A force sensor is set at the top of each linear actuator. Two strain gauges are used as a force sensor. The strain gauge detects small displacements of the top end of the linear actuator caused by the force applied by the user. The position of the top end of the linear actuator is measured by an optical encoder connected to the axis of the DC motor. The maximum stroke of the linear actuator is 80mm, and the maximum speed is 100mm/s.

The FEELEX 2 is designed to improve the resolution of the haptic surface. In order to determine the resolution of the linear actuators, we considered the situation where a medical doctor palpates a patient. Considering the above condition, the distance is set to be 8mm. The size of the screen is 50mm X 50mm, which allows the user to touch the surface using three fingers.

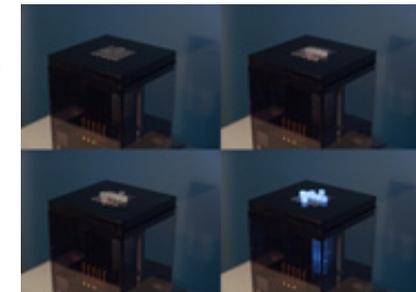
In order to realize 8mm resolution, a piston-crank mechanism is employed for the linear actuator. The size of the motor is much larger than 8mm, so the motor should be placed at a position offset from the rod. The piston-crank mechanism can easily achieve this offset position. Figure 4 illustrates the mechanical configuration of the linear actuator. A servo-motor from a radio-controlled car is selected as the actuator. The rotation of the axis of the servo-motor is converted to the linear motion of the rod by a crank-shaft and a linkage. The stroke of the rod is 18mm, and the maximum speed is 250mm/s. The maximum torque of the servo-motor is 3.2Kg-cm, which applies a 1.1Kgf force at the top of each rod.

This force is sufficient for palpation using the fingers. The flexible screen is supported by twenty-three rods, and the servo-motors are set remotely from the rods. Fig.2 shows an overall view of the FEELEX 2. The twenty-three separate sets of piston-crank mechanisms can be seen in the picture.

The diameter of each rod is 6mm. We cannot put a strain gauge on the top of the rod because of its small size, so we therefore measure the electric current going to each servo-motor to sense the force. The servo-motor generates a force to maintain the position of the crank-shaft. When the user applies a force to the rod, the electric current on the motor increases to balance the force. We measured the relationship between the applied force and the electric current. The applied force at the top of the rods is calculated using data from the electric current sensor. The resolution of the force sensing capability is 40gf.

Lumen: Interactive Visual and Shape Display for Calm Computing

Lumen is an interactive display that presents visual images and physical, moving shapes, both controlled independently. The smooth, organic physical motions provide aesthetically pleasing, calm displays for ambient computing environments. Users interact with Lumen directly, forming shapes and images with their hands. *Life Enhancement*. There is something magical and emotionally engaging in smooth, slow physical motions, such as waves of the sea or the motion of grass in the wind. The word "organic" perhaps most closely defines them. Lumen delivers such slow, organic animations, creating calm, emotionally pleasing shape and image displays for future living environments. Lumen's ambient, calm computing environment is ideal for presenting minimal information unobtrusively. Though the current system is small, Lumen can be easily scaled up to create larger surfaces embedded into walls, furniture, and ceilings, and it can become invisible when it is not needed. Physical shape can be very expressive in communicating between human and machine: using the same image but different physical motions can create a very different "feel of the data." Users interact with Lumen by simply moving their fingers over its surface. They can also interact with each other by touching hands in Lumen's asynchronous image and shape display. *Vision*. As new actuators and "artificial muscle" technologies are developed, Lumen could be used to create wall-size, low-power, inexpensive displays for home and office interiors or architecture. It will allow users to present information in a simple, aesthetically pleasing, yet efficient way, creating a new layer of information devices for ubiquitous computing environments. Lumen technology could also be used to create virtual controls: buttons and sliders that are formed dynamically. *Goals*. This project has three main goals:

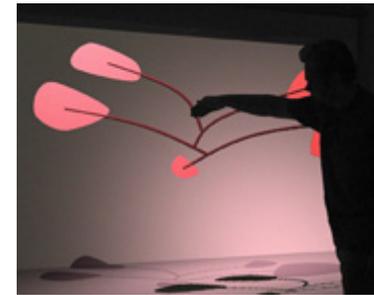


1. Alternative information displays. For many applications, a small amount of information is sufficient, and Lumen-type devices can be very effective.
2. New communication experiences. Controlled physical motion is a very attractive technique for creating communication links. It can allow two people to touch each other over the network, to see and feel the shapes of each other's hands.
3. Novel human-machine interaction. Physical shape can be very expressive in communicating between human and machine. Using the same image but different physical motions allows creation of a very different "feel of the data" with very low-resolution images.

Innovations. The entire Lumen device is custom made; it cannot be simply assembled from the off-the-shelf components. Designing Lumen was a challenge because it is intended for home use, so it must be compact and noiseless, actuation should be smooth and continuous, and users must be able to interact with the device directly, without using any external input devices. Adding interactivity was another innovation. A custom-made Smart Skin sensor was built into the surface of the Lumen to recognize shapes and fingers. It can also feel the user's motions. It calculates the distance between the hand and the surface by using capacitive sensing and a mesh-shaped antenna made of simple copper wire. Smart Skin is perfect for Lumen because all sensing elements can be integrated within the surface of the device, and unlike camera techniques, it does not suffer from lighting and occlusion.

Novel Infrared Touch-Screen Technology and Associated Artwork

This is a novel technique for a touch screen using front-projected infrared with a rear-projected interactive video display. The technology is demonstrated with three art pieces. Two of the art pieces allow participants to create works in the forms of Calder and Mondrian. The third is a visualization of radio interferometry. *Life Enhancement.* This technology is generally applicable to numerous user-interface applications. However, I have chosen to use it for interactive artwork. In my opinion, artwork by its very nature enhances the life of both creator and viewer. In this and all artwork I create, I strive to give participants a playful environment in which to interact. I believe that all play is a form of learning, and I hope that my projects create at least intuitive, if not deeper, understanding of some aspect of life or the universe without necessarily being overtly didactic. *Vision.* Unlike my previous shadow-sampling techniques, I believe that this system has very broad usefulness beyond interactive artwork (my primary concern). In particular, I can imagine it being used in situations where one wants to permit interaction without any exposed devices such as shopping windows or gallery fronts. Artistically, I believe that interactive artworks such as these and others that I have created have a broad appeal to an enormously wide range of audiences. I have rarely displayed any interactive work that did not elicit smiles, laughter, and a childlike sense of playfulness in the participants. The power of interactive art lies in harnessing play, as is manifest in games of all kinds, from traditional sports and board games to advanced video games. Play is a powerful, if poorly understood, human instinct and is a powerful tool which, when harnessed properly, delights and teaches in unexpected ways. *Goals.* Technologically: Proof of concept that infrared shadow-sensing works. Artistically: To permit participants to work in the form of two famous abstractionists (Calder and Mondrian) so they can explore their work as a creator instead of as a viewer. In the case of "Interference," to build intuition for the beautiful mathematics of wave mechanics. *Innovations.* All hardware components are off-the-shelf consumer items. The key innovation is the idea that multiple, diffuse light sources cast from oblique angles can be used as an image-detection system. Many well-known image-detection and filtering algorithms are exploited, and coded in custom C/C++. Also necessary are the calibration algorithms that correlate camera space to screen space, but these algorithms are not novel, as I used identical code in the shadow detection pieces displayed at SIGGRAPH 2002. The artistic components include several innovative uses of commonly available algorithms. In particular, the Calder piece relies heavily on an open-source physics engine (ODE).



Pop Up!: A Novel Technology for Shape Display of 3D Objects

Pop Up! is a new visual and tactile display for presenting 3D shapes using a pin-rod matrix that conveys visual depth information and is designed to be touchable. Due to its broad range of movement, it is capable of displaying large-scale, dense objects such as human faces, geographical features and landscapes. *Life Enhancement.* When they create digital artwork with this Pop Up! display, artists can see not only a 2D perspective, but also a complete, touchable 3D shape. The form can be directly modified with hand motions, just like sculpting with clay. CG artists share a common dream: to directly manipulate 3D CG works with their own hands, but only classical and bulky force-feedback devices have even partially achieved this objective. Pop Up! can also be used to create powerful advertising. Almost all visual advertising displays 2D images, on paper, in video, or on billboards. With Pop Up! advertisers could display large objects and change their shapes in real time. *Vision.* In the near future, 3D information will be shown not only in 2D flat displays, but also in true 3D systems. Pop Up! is an infant example of this next-generation display technology. The system

solves three practical problems: speed, density, and pixel volume. In a few years, it might be hard to imagine life without Pop-Up! *Goals.* One goal is realization of a "face phone" that can display dynamic range-finder data obtained from a human face. Conventional videoconference technology can convey vocal and visual information, but a "face phone" could also convey a sense of existence, which would enhance human communication. The ultimate goal is for Pop Up! to become a substitute or supplement to the traditional LCD or CRT display, to enhance its quality and modality as a 3D interface. *Innovations.* The most important innovation of this display is its use of a coil-formed shape memory alloy (SMA) as an actuator. Ordinary SMA does not have a long stroke, and the movement is quite slow, which is the main reason why it is generally avoided as an actuator. To improve its response rate, the SMA was formed into a coil, which enables much faster cooling than previous SMAs. Thanks to its thinness, speed, and simplicity, it is an ideal actuator for Pop Up! Activating multiple actuators at the same time is another technical challenge. Because the goal of this project is to drive 32 X 32 pin rods, it uses a "matrix drive" method. In most matrix-drive systems, such as those used in LCDs, each element reaches its target luminance instantly. But the SMA reaches its target position with some delay, which is a function of its previous position. Therefore, the challenge is to optimize time scheduling, to determine how much time should be assigned to each pin rod. Pop Up! features a custom-optimized controlling algorithm that enables real-time display.

Untethered Force Feedback Interface That Uses Air Jets

This is an untethered interface that eliminates the annoyance of wires and other connectors by using air jets to establish force feedback. Attendees experience interaction with a virtual object that responds to being "touched." The sense of touch is provided by air jets, while visual clues are provided by a projection-based stereo display. *Life Enhancement.* Many devices created to realize force feedback demand the use of connecting wires and/or demand that the participant wear a heavy device. Both of these requirements disturb the user's free movement and raise the level of annoyance. Compared to the tools used in daily life, existing VR systems leave a lot to be desired. We focus on acceptance of VR systems, so we concentrate on eliminating anything that restrains the user. A tool will not become part of daily life if it is seen as an encumbrance. Our keyword is "untethered." *Vision.* Provision of unencumbered touch will augment the sense of vision to make virtual environments an everyday occurrence. The most exciting new possibility provided by our system is its freedom in range of experience. For instance, we can imagine this system being installed in large areas or buildings to enrich user experience and establish interaction over extended distances. Multiple users can participate in the same game when nozzles are placed in strategic locations. Users can enjoy interaction with a virtual object while walking along a hallway. The area of a single array is not limited; some applications may require only one air jet, while others may need several hundred. This system is still quite primitive, and more research should yield an interface with better feel and less noise. Even in its present form, it provides an excellent indication of what is possible and VR's potential to enter daily life. *Goals.* The prime goal is a force-feedback interface that dispenses with all tethers. Users need only red-and-blue 3D glasses and a lightweight paddle. This system does not constrain the user's activity, and it is comfortable. *Innovations.* The major component is the force-feedback interface. The basic idea of this method is that air released from a jet impacts the "air receiver" (paddle) held by the user to provide force feedback. The demonstration system uses a flat, 2D array of air jets that point upward to create the feel of touching three-dimensional virtual objects. The system detects the air receiver's position, determines if it is "contacting" the surface of a virtual object in the virtual space, and if contact exists, releases air from the nozzle immediately under the air receiver. The release of air continues until the air receiver is removed from the surface or is moved to a new position on the surface. In the latter case, the system releases air from the next appropriate nozzle. The second component is the projection-based stereo display system. Stereo images are projected onto the top surface of the desk. The black air-jet holes on the surface are covered with simple flaps so that the stereo images projected on the desk are not degraded. The third component is the optical position-tracking system. The user's 3D glasses and paddle have visual markers, and their positions and orientations are detected by two cameras. These three components realize a completely wire-free interface. The user simply puts on the lightweight 3D glasses and holds the lightweight paddle. The user is not prevented from moving freely within the system, and is never bothered by system operation. The system is completely devoid of user restraints.

7.4. Wii Remote

The Wii Remote (also called the "Wii-mote", or "Wii-Pointer") is the controller for Nintendo's next generation console, named Wii. It has received attention for being unlike the controllers for any other popular gaming console. Nintendo reps have noted that the controller is still considered a prototype, as the possibility for major and minor changes exist.

Remote control design

The primary controller for Wii is not based on the traditional gamepad controller design of current gaming consoles, instead assuming a one-handed remote control-based design more familiar to the non-gaming public. This was done in part to appeal the console to a broader audience. Named the "Wii Remote" (often referred to as the "Wii-mote"), the controller communicates wirelessly with the console via short-range Bluetooth radio, with which it is possible to operate up to 4 controllers as far as 10 meters (approx. 33 ft.) from the console.[1][2] The controller's symmetrical design allows it to be used left-handed or right-handed. The Wii Remote can also be turned on its side and used like a NES controller. The current design of the Wii Remote uses AA batteries as a power source; a direct recharging option has not been revealed. According to an interview with Nintendo industrial designer Lance Barr, limitations of the Wii Remote's expansion port make it unlikely that it will be used for internal battery charging.[3] Nintendo of America's Senior Director of Public Relations, Beth Llewelyn, has stated that Nintendo "hasn't quite figured out what to do about power for the Remotes".[4] Barr has stated that the battery life for the Wii Remote will certainly be less than that of the Wavebird (which has an estimated life of 100 hours), explaining that the Wii Remote uses a greater amount of bandwidth than the Wavebird.[3] Color variations at launch time are currently unknown; Nintendo has said that they are waiting for feedback before they confirm any of the colors.[3] At E3 2006, Nintendo displayed white, black, and blue controllers.

Sensing

The Wii Remote is also able to sense movement and orientation. Sensors in the Wii Remote allow it to sense linear motion along three axes, as well as tilt.[1][5]

The sensor bar

In addition, a sensor bar placed near and parallel to the display screen allows the Remote to be used as a pointer up to 5 meters (approx. 16.5 ft.) away.[2]

This allows players to mimic actual game actions, such as swinging a sword or using a flashlight, instead of simply pushing buttons. An early marketing video showed actors miming such actions as fishing, cooking, drumming, conducting a string quartet, shooting a gun, sword fighting, and performing dental surgery.[6]

Despite the controller's similarity to light guns, which are only compatible with standard-definition cathode ray tube televisions, Nintendo has stated that Wii and its controller will be compatible with all televisions including digital projectors.

Controller feedback

The Wii Remote also provides basic audio and force feedback functionality. At the 2006 E3 press conference, it was revealed that the Wii Remote has its own independent speaker on the face of the unit. This was demonstrated by a developer as he strung and shot a bow in *The Legend of Zelda: Twilight Princess*. The sound from both the Wii Remote and television was altered as the bow shot to give the impression of the arrow travelling away from the player.



Personalization

In an interview with Kotaku, Shigeru Miyamoto mentioned that, as it is hoped that every player in the household will ultimately have their own Wii Remote, one of Nintendo's ideas for the Wii Remote is the creation of personalized user profiles which are tied to individual Remotes.[7]

Controller expansions

The Wii Remote also features an expansion port at the bottom which allows various optional attachments to be added to the controller. The

Nunchuk

The first attachment Nintendo revealed (in tandem with the controller itself, at the 2005 Tokyo Game Show) is a unit featuring an analog stick and two trigger buttons. It connects to main Wii Remote via a short cord, and its appearance while connected to the main controller has led it to become dubbed the "Nunchuk" controller. Like the Wii Remote, the Nunchuk controller also provides three-axis motion-sensing.[2][8] Nintendo has confirmed that a Nunchuk controller will be bundled with the Wii console.[9][10] The two shoulder buttons, named Z1 and Z2 respectively, had been reshaped and renamed since the Game Developers Conference. The top shoulder button, now called C, is a good deal smaller than the lower shoulder button, now called Z. The C button was oval shaped, while the Z button was a more traditional (and slightly larger) square. [11]

Classic Controller

Nintendo had previously announced a controller "shell" which will resemble a traditional game controller, often referred to as a "classic-style expansion controller".[12] As described at the time, the Wii Remote would slot inside this shell, allowing gamers to play games using a traditional-style gamepad, while leveraging the remote's motion sensing capability. According to Satoru Iwata, it would be meant for playing "the existing games, virtual console games, and multi-platform games".[13]

During E3 2006 Nintendo introduced a Classic Controller, which plugs into the Wii Remote via a cord, similar to the Nunchuk controller.[1] The classic controller showcased contains no space for the Wii Remote to slot inside, as previously described, but does contain notches on the backside, presumably for clipping the controller to something else. The purpose for these notches remains undisclosed.[14]

The Classic Controller features two analog sticks, a D-pad, A, B, X, and Y buttons, L and R analog shoulder buttons and a Z button located next to the R shoulder button. Despite press material which indicates a Z button next to the L button (referred to as 'ZL'; the Z button on the right side referred to as 'ZR'),[1] hands-on accounts from E3 have reported that the most recent design has no Z button on the left side.[14] It also has a set of -, Home, and + buttons like those on the Wii-mote, with the - and + buttons labelled 'Select' and 'Start', respectively.

British magazine Edge reports in the July 2006 issue that Nintendo had privately shown prototype replicas of classic console controllers for Virtual Console games.[18] Included in the article was a photograph showing a briefcase with what appears to be a Wii Remote surrounded by controllers for the NES, Super NES, N64, Genesis, and TurboGrafx-16.[1] The appearance of a cord leading from the Wii Remote to the Super NES Controller has lead to Internet speculation that Nintendo will release similar controllers which connect to the Wii Remote in the same fashion as the Classic Controller.[19]



References

- ^ a b c d (Japanese) Wiiの概要 [コントローラ]. Nintendo Company, Ltd.. Retrieved on 2006-05-09.
- ^ a b c Nintendo Wii - Hardware Information. Nintendo of America. Retrieved on 2006-05-09.
- ^ a b c Wii Controllers: No Recharging Yet. The Wiire. Retrieved on 2006-05-11.
- ^ Crecente, Brian. Rechargeable Wii-Motes?. Kotaku. Retrieved on 2006-06-08.
- ^ Wisniowski, Howard (2006-05-09). Analog Devices And Nintendo Collaboration Drives Video Game Innovation With iMEMS Motion Signal Processing Technology. Analog Devices, Inc.. Retrieved on 2006-05-10.
- ^ TGS 2005: Revolution Teaser Video. Retrieved on 2006-03-16.
- ^ Crecente, Brian (2006-05-19). E306 Video: Shigeru Miyamoto Interview. Kotaku. Retrieved on 2006-05-31.
- ^ STMicroelectronics Drives Gaming Revolution with Nintendo's Wii. (2006-05-09). Retrieved on 2006-05-12.
- ^ Wales, Matt (2006-05-22). Reports claim Wii to slap down 16 at launch (in English). Computer and Video Games. Retrieved on 2006-05-25.
- ^ Berghammer, Billy (2006-06-02). The Ultimate in PR Spin: The Perrin Kaplan Interview: Part Four (in English). (WMV) Game Informer. Retrieved on 2006-06-08.
- ^ Hands-On with the Wii Controller. (2006-05-12). Retrieved on 2006-05-12.
- ^ Gibson, Ellie (2005-09-16). Jim Merrick Takes Control. Eurogamer. Retrieved on 2006-05-09.
- ^ Satoru Iwata. TOKYO GAME SHOW 2005 : TGS FORUM Keynote Speech. Retrieved on 2006-03-15.
- ^ a b Sklens, Mike (2006-05-10). News Article: Wii 'Classic Controller' Revealed. Planet GameCube. Retrieved on 2006-05-11.
- ^ Miller, Ross (2006-05-10). E3: The Wii Zapper prototype revealed. Joystiq. Retrieved on 2006-05-11.
- ^ E3 2006: Light Gun Shell Revealed!. IGN Wii. Retrieved on 2006-05-10.
- ^ Daiker, Brandon (2006-05-11). "Duck Hunt" Wii Pointer Shooting Gallery Impressions. N-Sider. Retrieved on 2006-05-10.
- ^ (July, 2006). Edge #164: p. 10. Future Publishing.
- ^ EDGE has pictures of "new" Wii controller add-ons!. NeoGAF forum. (2006-06-13). Retrieved on 2006-06-13.

7.4.1. Rasseqna stampa

MEMS Plays in Nintendo's Wii

Staff Reporter - Electronic News, 5/10/2006

Nintendo has revealed major technology-forward elements of its previously enigmatic Nintendo Wii (formerly known as Revolution) -- controllers that operate cable-free in a three-dimensional landscape using MEMS technology. STMicroelectronics and Analog Devices Inc. are both providing micro electro-mechanical systems (MEMS) technology to the Wii (pronounced "we") controllers. MEMS technology, micron-sized structures that interact with the physical world, allows the Wii controller to respond to changes in direction, speed and acceleration, down to the most delicate movements of its players, Nintendo said from the E3 gaming show this week. Controller sensors can detect the motion and tilt of a player's hand in all three dimensions and convert it into immediate game action. Specifically, the controllers use embedded acceleration sensors to enable players' wrist, arm and hand movements to interact with the games. Tilt measurements allow users to move characters, while the three-axis acceleration sensing easily transforms the controller that can be held in one hand into a virtual sword, gearshift or musical instrument.

Central to the Wii console's design is multi-axis linear acceleration sensing. ADI's ADXL330 has a three-axis motion signal processing performance to allow the gamer's body motion to control his actions in the game in real time. When the new controller is picked up and manipulated, it provides a quick element of interaction, sensing motion, depth and positioning



Wii's controller uses MEMS sensors and wireless technology to detect the motion and tilt of a player's hand in all three dimensions and convert it into game action.

dictated by the acceleration of the controller itself. "For the industry's first mainstream game controller using MEMS acceleration sensors, we turned to Analog Devices, an industry leader whose acceleration sensors are used by Nintendo for popular games like Kirby's Tilt 'N Tumble for Game Boy Color," said Genyo Takeda, senior managing director/general manager, Nintendo's integrated R&D division, in a statement. "We selected the ADXL330 because its accuracy, small size, and extremely low power consumption were critical to the Wii console's design objectives and key for a wireless controller that will revolutionize the gaming industry."

The acceleration sensor's miniature dimensions (5x5x1.5mm³) significantly contribute to the light TV-control-shaped hand piece. Ultra-low power consumption of ST's MEMS devices prolongs the battery life of the controller and the high thermal stability avoids unwanted motion sensing in varying temperature conditions, the company said. The design of the sensor also provides high immunity to vibration and a shock resistance up to 10,000g. "With ST's 3-axis accelerometer, Nintendo is revolutionizing the gaming world by putting players into the action," said Benedetto Vigna, MEMS business unit director, STMicroelectronics, in a statement. "This valuable partnership confirms our conviction that MEMS sensors open up fascinating possibilities for intuitive man-machine interaction in consumer applications, which can also allow non-expert players to enjoy the excitement of the game."

Linking the controller movement to the main Wii console is Broadcom Corp.'s Bluetooth and Wi-Fi technologies. The Wii console is the first major gaming platform on the market to introduce Bluetooth technology as a standard feature, eliminating the cables between the console, hand-held game controllers and other accessories. Broadcom Bluetooth technology used in the controllers significantly increases the level of performance in the game controller, with special enhancements that provide low latency to accommodate the fast rate of interaction between the controller and the console. The Wii console also includes Broadcom's single-chip 54g Wi-Fi solution, which provides a wireless link for playing online or multi-player games. The console also includes Broadcom BroadRange technology, which allows users to maintain higher data rates further from the access point. Broadcom's family of 54g chipsets are featured in the leading brands of networking equipment, PCs, broadband modems and consumer electronics.

PARIS — In a move that may shift the focus in the videogame console battle from fast, powerful processing engines to more intuitive user interfaces, Nintendo unveiled a videogame controller Tuesday (May 9) that uses three-axis motion signal-processing technology.

The controller was shown at the E3 game show in Los Angeles. Analog Devices Inc. and STMicroelectronics are key suppliers for the microelectromechanical systems (MEMS) sensors used in the controller. The system comprises two controller units—main and freestyle—each powered by a motion sensor. ADI sensors are used in the main controller; ST won the socket for the freestyle unit. Nintendo developed the controller platform for its next-generation video console, called Wii (pronounced "whee"). Details about that console are not scheduled to be revealed until later this year. Not to be outdone, Sony Computer Entertainment Inc. disclosed this week that its Playstation 3 includes a videogame controller equipped with a six-axis sensing system. Sony was mum about how its new controller works and the identity of the motion sensor supplier. Whereas a conventional videogame controller is held with two hands, Wii's main and freestyle controllers are designed to be held in one hand. They can also be swung, aimed, twisted and turned as if they were the player's sword, motorcycle handle or tennis racket. Some games will use only the main controller; others will use both controllers, tethered by a cable.

Critical to the new generation of game applications is the MEMS device's ability to sense a player's movements in three dimensions: forward/backward, left/right and up/down. When the unit is grasped, the system senses motion, depth and positioning as dictated by the acceleration of the controller, ADI said. ADI is supplying its ADXL330, a standard three-axis motion signal processor, for Wii's main controller. "It's a single piece of silicon integrated with sensor and signal-conditioning elements," said Richard Mannherz, customer marketing manager for ADI's Micro Mechanical Product division. Benedetto Vigna, director of ST's MEMS Business Unit, said its device is optimized for Wii's freestyle controller. ST's system-in-package consists of two components: a three-axis MEMS accelerometer sensor and an interface chip that translates movement and inclination into electronic signals, amplifies them and provides analog output.

The motion sensors used in the Wii controllers are not necessarily "pin to pin" compatible, said ADI's Mannherz. "But it's a common practice among game console companies to hedge their bets and secure two suppliers for such key components," he said. The use of motion sensors in consumer devices is not new; indeed, their use mobile handsets and notebooks is on the rise. ST's three-axis linear accelerometer has been used in

applications such as mobile phones, PDAs and other battery-powered products to enable user interfaces based on movement. The same technology has also been used as the "airbag of a computer," said ST's Vigna. For example, an accelerometer embedded in a notebook computer could sense it falling off a desk and move the hard drive's head to a safe area so that it wouldn't crash onto the drive. ADI, which claims to be the leading supplier of motion sensors for auto airbag systems, began researching MEMS technology 15 years ago "as a platform to create highly reliable, extremely fast devices that could sense motion," Mannherz said.

ADI's two-axis accelerometer is in game cartridges designed five years ago for Nintendo's Game Boy Advanced. But the recent emergence of a highly integrated three-axis accelerometer is making it much easier and more cost-effective for game console makers to integrate motion sensors into mainstream videogame controllers, said ADI marketing manager Christophe Lemaire. ADI's motion signal-processing technology has been used in consumer devices ranging from a guitar effects controller to "intelligent" golf clubs. By transferring core MEMS sensor technology to the videogame world, "you can simplify [game play] and make it more natural," ST's Vigan argued. "You don't need to worry about which buttons to push anymore".

MEMS Sensors Breathe Life Into Wii Controller

May 25, 2006

Attendees lined up for hours at the Electronic Entertainment Expo in Los Angeles earlier this month for the chance to hold the tiny peripheral that is making a big splash in the gaming industry. A stark contrast to the traditional two-handed wired controller, the Wii Remote (or Wii-mote) can be held in one hand and is motion controlled (see Figure 1). Sans wires, the Wii-mote communicates with the gaming console via Bluetooth, and can be used up to 10 meters away.

Central to the Wii-mote's design are acceleration sensors by Analog Devices (ADI) and STMicroelectronics. Analog Devices and Nintendo have collaborated in the past on software for previous gaming systems. ADI's acceleration sensors were used in games like Kirby's Tilt'N Tumble for Game Boy Color. For the Wii-mote, Analog Devices's ADXL330 is used to sense motion of the game player in three dimensions of freedom: forward-backward, left-right, and up-down. When the Wii-mote is picked up and manipulated, it provides a quick element of interaction, sensing motion, depth, and positioning dictated by the acceleration of the controller itself.

"We selected the ADXL330 because its accuracy, small size, and extremely low power consumption were critical to the Wii Console's design objectives," said Genyo Takeda, senior managing director for the Integrated Research & Development Division at Nintendo, "... and key for a wireless controller that will revolutionize the gaming industry." The Wii-mote also includes STMicroelectronics' three-axis acceleration sensors. The sensors, which are based on the company's Micro Electro-Mechanical Systems (MEMS) technology, measure tilt to allow users to move characters, while three-axis sensing transforms the controller into a virtual sword, gearshift, or musical instrument. In an interview with The Wire, Nintendo industrial designer Lance Barr said that the Wii-mote operates on standard alkaline, lithium ion, or rechargeable AA batteries. Nintendo chose AA batteries, Barr said, because of their cost, but also because the adaptor port on the bottom of the controller presents limitations for an internally chargeable battery. Already, rival gaming console manufacturers are starting to pick up on the motion-controlled phenomenon. Industry insiders say that Sony has been working on a prototype for a motion-controlled controller for the Playstation 3.

Analog Devices has introduced a three-axis accelerometer that features power consumption as low as 200 μ A and targets applications in handsets and other handheld consumer products. The device will enable engineers to provide position awareness for applications such as games, interfaces, image-stabilizing cameras, and smart power management functions. The ADXL330, part of the company's iMEMS motion-signal-processing product line, consumes 200 μ A at 2.0V and 320 μ A at 3.0V—less than half as much as competitive devices, according to the company. The sensor integrates signal-conditioning circuitry and comes in a 4×4×1.45-mm LFCSP (lead-frame, chip-scale package). "This is the first three-axis motion sensor accelerometer...and the first single-chip three-axis MEMS accelerometer with signal conditioning and circuitry together on the same chip," said Mack Lund, the company's product marketing manager for micromachined products. The chip's mechanical sensing structures allow it to offer low power in a small package while also providing stable performance over its temperature range, he said.

In addition, the device's low supply voltage, from 2.0 to 3.6V, benefits battery-powered applications, Lund said. "If users only need to use this for short periods of time or they want to duty-cycle it, they can get even lower power than best-in-class," he said. The device can turn on in 1 msec,

take a measurement, and turn off again to save power, he said. The sensor also features a 10,000g shock rating, which is important in devices that are prone to dropping. The LFSCP is a hermetically-sealed package that protects the MEMS sensor structures. The ADXL330 is now sampling and costs less than \$2 in high volume, according to the company. It is slated for volume production during the first quarter of next year.

Analog Design's ADXL330

The ADXL330 is a small, low power (200µA at 2V) 3-axis iMEMS® accelerometer complete with signal conditioned voltage outputs on a single monolithic IC. The product measures acceleration with a minimum full-scale range of ±3g. For tilt sensing applications it measures the static acceleration of gravity. It also measures dynamic acceleration resulting from motion, shock, or vibration.

The ADXL330 is available in a small 4 mm × 4 mm × 1.45 mm, 16-lead, plastic LFCSP.

* Applications: Motion and Tilt sensing in Mobile Devices

* Handsets

* Motion-Enabled Gaming Devices

* Disk Drive Protection Systems

* Sports and Health Devices

"This is the first three-axis motion sensor accelerometer...and the first single-chip three-axis MEMS accelerometer with signal conditioning and circuitry together on the same chip," said Mack Lund, the company's product marketing manager for micromachined products. The chip's mechanical sensing structures allow it to offer low power in a small package while also providing stable performance over its temperature range, he said.

In addition, the device's low supply voltage, from 2.0 to 3.6V, benefits battery-powered applications, Lund said. "If users only need to use this for short periods of time or they want to duty-cycle it, they can get even lower power than best-in-class," he said. The device can turn on in 1 msec, take a measurement, and turn off again to save power, he said. The ADXL330 is now sampling and costs less than \$2 in high volume, according to the company. It is slated for volume production during the first quarter of next year.

Analog Devices and Nintendo collaboration drives video game innovation with imems motion signal processing technology

With ADI's ADXL330 3-axis iMEMS accelerometer, Nintendo's Wii Console redefines the gaming experience by bringing the gaming world into the physical world. Norwood, MA(5/9/2006) - Building on its relationship with Analog Devices, Nintendo is creating a truly interactive, lifelike, motion-based gaming experience for players of all ages while igniting the creative forces of game developers around the world. Nintendo breaks more than 20 years of video game history by abandoning the traditional controller held with two hands and introducing a new freehand-style unit held with one hand. Incorporating the Analog Devices, Inc. (NYSE: ADI) 3-axis ADXL330 iMEMS® acceleration sensor, the intuitive, innovative Wii™ controller allows players to run, jump, spin, slide, steer, accelerate, bank, dive, kick, throw and score in a way never experienced in the history of gaming.

Central to the Wii Console's design is accurate and reliable multi-axis linear acceleration sensing. When facing the design challenges of the new Wii controller, Nintendo collaborated with Analog Devices, a leader in signal processing technology and well known for innovative motion sensing. The ADXL330's robust 3-axis motion signal processing performance enables a new concept in console video game controllers and gaming interfaces by allowing the gamer's body motion to control his or her actions in the game in real time. The ADXL330 is used to sense motion of the game player in three dimensions of freedom: forward-backward, left-right, and up-down. When the new controller is picked up and manipulated, it provides a quick element of interaction, sensing motion, depth and positioning dictated by the acceleration of the controller itself. The Wii Console is more intuitive and realistic than existing game consoles that require buttons to be pushed and will expand the gaming market by appealing to new classes of users. To achieve a true-to-life gaming experience and the most realistic reproduction of game player motion, Nintendo relied on their experience with Analog Devices' iMEMS Motion Signal Processing™ technology. "For the industry's first mainstream game controller using MEMS acceleration sensors, we turned to Analog Devices, an industry leader whose acceleration sensors are used by Nintendo for popular games like Kirby's Tilt 'N Tumble for Game Boy Color," commented Genyo Takeda, senior managing director/general manager, Integrated Research & Development Division, Nintendo Co., Ltd. "We selected the ADXL330 because its accuracy, small size, and extremely low power consumption were

critical to the Wii Console's design objectives and key for a wireless controller that will revolutionize the gaming industry." "Motion is an integral part of our lives. We are delighted that Nintendo selected us again for another incredibly innovative motion-based product like the Wii Console and Controller," said Bill Giudice, vice president and general manager, Micromachined Products Division, Analog Devices. "Analog Devices offers unparalleled experience with integrating motion sensing to enhance the products we use every day, whether it's the automobiles we drive, the mobile phones we use, or the games we play. Our customers rely on our ability to deliver high performance, superior quality motion sensors in large volume for many exciting applications About Analog Devices' iMEMS technology. As a pioneer in the micromachine IC industry, ADI produced the first fully-integrated, single chip iMEMS® (integrated Micro Electrical Mechanical System) accelerometer in 1991. Since then, ADI has maintained a leadership position and remains the industry's only producer of high-volume, single-chip iMEMS accelerometers/motion sensors and continues to be the leading supplier of airbag sensors worldwide with shipments of more than 200 million units to date. ADI's iMEMS motion sensors are used in a wide range of consumer, automotive and industrial product applications.

STMicroelectronics Drives Gaming Revolution with Nintendo's Wii(TM)

STMicroelectronics (NYSE:STM), one of the world's leading semiconductor manufacturers, today announced that its three-axis acceleration sensors will be used to provide a motion-activated user interface for Nintendo's new home console, Wii. Expected to dramatically change the way people play games, the Wii controller includes ST's high-performance acceleration sensors that can detect the motion and tilt of a player's hand in all 3 dimensions and convert it into immediate game action. Driven by ST's Micro Electro-Mechanical Systems (MEMS) technology - micron-sized structures that interact with the physical world - the Wii controller can respond to changes in direction, speed, and acceleration, down to the most delicate movements. Specifically, the controllers use embedded acceleration sensors to enable players' wrist, arm, and hand movements to interact with the games. Tilt measurements allow users to move characters, while the accurate three-axis acceleration sensing easily transforms the controller into a virtual sword, gearshift, or musical instrument. Sounds like the first press release is speaking only of the three lateral DOF (sensor bar coordinated), while the second technology will be responsible for all tilt recognition (the other 3 DOF). "Commercially available inertial MEMS sensors include different types that can sense movement in one, two, or even three linear axis. Some types also directly sense rotational acceleration around one axis and others, gyroscopes, can implement the function of a Coriolis-effect, sensing the angular rate. These sensors share the same principle of operation and sense movements by measuring changes in capacitance between moving and fixed structures."

Pixart teams with Nintendo for sensor tracking technology for Wii™ controller

Pixart's Tracking Sensor Technology Provides Quick, Responsive Play Experience on Wii Remote - May 11, 2006.

PixArt Imaging Incorporation (PixArt), a market leader in CMOS sensor SoC (System-On-a-Chip) and related application semiconductors, today announces a strategic relationship with Nintendo Co., Ltd., to provide object tracking technology for Nintendo's new-generation gaming controller, the Wii Remote. PixArt's premium tracking technology will enable Nintendo to present innovative interaction gaming controllers for its new-generation gaming platform, Wii. PixArt's Multi-Object Tracking™ engine (MOT sensor™) technology can track multiple objects in an unbelievably quick and responsive way. As a result, Nintendo can enable its new gaming controller to interact with people by tracking the movement of the Wii Remote. The playing experience will be unprecedented, exciting and easy, even for young children or older people who cannot operate the traditional gaming controllers. "PixArt's technology enables a quick, dynamic play experience," commented Genyo Takeda, Senior Managing Director/General Manager, Integrated Research & Development Division, Nintendo Co., Ltd. "With PixArt's technology, Nintendo will usher in a new era of video games." "It is PixArt's pleasure to partner with Nintendo, a technology leader, to launch the most powerful and pleasing solution for a gaming platform," said Sen Huang, Chief Executive Officer of PixArt. "The Multiple-Object Tracking™ engine (MOT sensor™) has the highest performance ever in the market, and teaming with Nintendo means a gaming experience that is truly intuitive and inspiringly easy." "In the world of interaction, every realistic operation can be simulated in a computerized game. Through play of such simulation, any person can learn very quickly and efficiently in a virtual way. Integrated tracking technology from proven vendors like PixArt will enhance the gaming experience for the overwhelming majority of all-generation gamers." said Sen Huang.

[1] Inclinometer

Pitch is rotation around the lateral or transverse axis. This axis is parallel to the wings, thus the nose pitches up and the tail down, or vice-versa. An aircraft pitches up to climb and pitches down to dive. An aircraft increases or decreases the lift generated by the wings when it pitches up or down by increasing or decreasing the angle of attack (AOA) through the lateral axis.

Roll is rotation around the longitudinal axis—an axis drawn through the body of the vehicle from tail to nose. This is also known as **bank**.

Yaw is rotation about the normal axis—an axis perpendicular to the pitch and roll axes. If an airplane model placed on a flat surface is spun or pivoted around the center of mass (coordinate origin) it would be described as yawing.



An **inclinometer**, AKA clinometer is an instrument for measuring angles of elevation or inclination. A Clinometer (or Inclinometer) is an instrument for measuring angles of slope (or tilt), elevation or inclination of an object with respect to gravity. Also known as a tilt meter, tilt indicator, slope alert, slope gauge, gradient meter, gradiometer, level gauge, level meter, pitch & roll indicator. More particularly, it can refer to:

- An airplane instrument. See turn coordinator or slip indicator.
- Any of various surveying instruments for measuring angles of elevation, slope, or incline, as of an embankment. Also called a clinometer.
- An instrument for measuring slight differences in slopes, particularly for geophysics. Such inclinometers are, for instance, used for monitoring volcanoes.
- An inclinometer may also be used in civil engineering projects to measure movements in walls and/or the ground.
- An instrument for determining the dip of beds or strata, or the slope of an embankment or cutting; a kind of plumb level.
- An instrument used by surveyors in order to measure an angle of inclination or elevation
- An inclinometer may also be used in some automotive safety systems. An inclinometer is also used in certain rollover warning systems.
- use widely in geohazard monitoring programme

Tilt sensors and inclinometers generate an artificial horizon and measure angular tilt with respect to this horizon. They are used in cameras, aircraft flight controls, automobile security systems, and special switches. Also used for platform leveling, boom angle indication, anywhere tilt requires measuring. Important specifications to consider when searching for tilt sensors and inclinometers are the tilt angle range and number of axes. The tilt angle range is the range of desired linear output measured in degrees. The number of axes the inclinometer and tilt sensor measure on is another important specification. Common sensor technologies for [tilt] sensors and inclinometers are accelerometer, capacitive (liquid), electrolytic, gas bubble in liquid, and pendulum.

Technical speculation and fantasy retroengineering

Wii Controllers: Unlocking the Secrets Battery and power specifics, SYNCHRO functionality, rumble motor details, LED secondary purpose and... a camera?

July 14, 2006

The Wii-mote features 6KB of "non-volatile" memory, whose exact purpose remains a mystery. It's possible that this memory could enable players to store custom settings for the controller. Official documentation does not specify one way or the other. The Wii-mote is able to act as something of an eye, measuring coordinates between 0-1023 on the X axis and 0-767 on the Y axis, which means that it is more or less seeing a megapixel image. Whether or not this data can be interrupted into visual information remains unknown. One of the big questions about the Wii-mote has remained how it will be powered. Official documentation on the topic offers some answers. The pointer will accept two AA alkaline batteries, which will keep it going for a considerable amount of time. If the precision aim functionality of the pointer is being used in games, the device will run for approximately 30 hours. If only the accelerometer functionality is being utilized, the Wii-mote will operate for 60 hours on two AA batteries. The controller communicates with the Wii console via Bluetooth technology on a 2.4GHz band. Recent photos of the Wii console's front flap showcase a button called SYNCHRO, whose purpose seems fairly obvious. The SYNCHRO button is, according to documentation, used to identify Wii controllers that can be used with the console. It assigns each controller a wireless ID number. Evidently gamers first press the SYNCHRO button on the console itself and then find and press another SYNCHRO button located inside the battery compartment of the Wii controller. Documentation also suggests that gamers can hold down the 1 and 2 buttons on the Wii-mote to accomplish the same task. All of the buttons on the Wii controller are digital in nature. This includes (obviously) the D-Pad, as well as A, B, 1, 2, -, +, Power and SYNCHRO. The C and Z buttons on the nunchuk unit are also digital. The sensor bar that interacts with the Wii-mote must be placed "above or below" the television set. The bar itself is about 20 centimeters in length and features two sensors, one on each end. The Wii-mote's LEDs serve two roles. The first is to show which player is communicating with the console at a given moment. Player 1 will light up on the far left, Player 2 the next over, and so on. The second purpose is to illustrate battery life. Upon booting up, four LEDs blink when the controller has between 75% and full power. Three LEDs blink to show 50% and 75% power. Two LEDs blink to show 25% and 50% power. And one LED blinks to illustrate anything below that. The Wii-mote features a built-in rumble motor that can be turned on or off. It does not offer varying degrees of rumble sensitivity. However, according to documentation, more intense vibrations can be simulated if developers rapidly trigger the motor on and off. Light sources from fluorescent and halogen lamps, plastic, mirrors and more may occasionally interfere with the pointer, based on official documentation. To eliminate this interference, the pointer must identify the sensor bar and mark its two coordinates. When pointing with the Wii-mote, the unit is actually interacting with the sensor bar, which then translates data to the television, in effect simulating a direct aim to the television. Interestingly, according to documentation the Wii-mote is able to act as something of an eye, measuring coordinates between 0-1023 on the X axis and 0-767 on the Y axis, which means that it is more or less seeing a megapixel image. Whether or not this data can be interpreted into visual information remains unknown, but we're not ruling out the possibility that the pointer could sub as a camera. This is, of course, purely speculative on our part, but stranger things have certainly happened - like, for instance, an internal speaker. We contacted Nintendo of America for comments on our latest findings. "There are many details that



we've divulged about the console, but there is also more to learn," said the subsidiary's PR manager, Matt Atwood, in a telephone conversation. "However, Nintendo does not comment on rumors or speculation."

Nintendo's New Wii™ Video Game Console Uses NEC Electronics' Embedded DRAM

KAWASAKI, Japan, SANTA CLARA, Calif., June 19, 2006 - NEC Electronics today announced that Nintendo Co., Ltd. has selected NEC Electronics' 90-nanometer (nm) CMOS-compatible embedded DRAM (eDRAM) technology for Wii™, its innovative new video game console. Designed to provide advanced graphics functions for this new gaming platform, the new system LSI chips with eDRAM will be manufactured using advanced technologies on NEC Yamagata's 300-millimeter (mm) production lines. Embedded DRAM technology integrates DRAM on the same chip with logic circuits, and is viewed as an optimal solution for three-dimensional (3D) graphics acceleration systems and other applications that need to process high bandwidth data using low power. In the past, the integration of an eDRAM structure with a standard CMOS process proved challenging. NEC Electronics achieved its fully CMOS-compatible eDRAM technology by integrating a metalinsulator-metal 2 (MIM2) stacked DRAM capacitor on the company's standard CMOS process. NEC Electronics first introduced MIM2 technology on 90 nm eDRAM in 2005, and volume production started that same year. The technology requires a material with a high dielectric constant to be placed between two electrodes, and a large charge to be maintained on the capacitor with low leakage and a small cell size. NEC Electronics has achieved this by 1) using a MIM structure for the electrodes in the DRAM cell to achieve lower resistance values and higher data processing speeds, 2) using cobalt-silicide (CoSi) DRAM cell transistors to increase driving performance, and 3) using zirconium-oxide (ZrO₂) in the capacitance layer (ahead of other vendors) to increase capacitance of the unit area. These and other breakthroughs have allowed NEC Electronics to develop eDRAM chips using its most advanced 90 nm process and to secure its roadmap to 55 nm eDRAM and beyond. Available now with NEC Electronics' unique 90 nm CMOS-compatible eDRAM process, the ASICs deliver significant advantages that promise to continue along the technological roadmap toward 55 nm processes and beyond. NEC Electronics selected MoSys® as the DRAM macro design partner for the Wii devices because MoSys is experienced in implementing 1T-SRAM® macros on NEC Electronics' eDRAM process. MoSys designed the circuits and layout of high-speed 1T-SRAM macros on NEC Electronics' 90 nm CMOS-compatible eDRAM technology. The adoption of these eDRAM ASICs for Nintendo's new game console is a clear vote of confidence in NEC Electronics' eDRAM technology. The company is committed to delivering more value-added features to spur the adoption of its CMOS-compatible eDRAM solutions across a wide range of applications.



Concludendo... Il controller del Nintendo Wii, il wiimote, rivelato al Tokyo Game Show 2005, ha una forma che ricorda il telecomando della televisione ma funziona come un mouse da computer in uno spazio reale a tre dimensioni. Il controller ricorda il Nintendo 64 e il NES in termini di pulsanti, ma rappresenta il maggiore distacco dagli ultimi venti anni di design di console e di interfacce. Due sensori ai lati della televisione permettono al controller di percepire la propria posizione in uno spazio tridimensionale; altri sensori permettono di percepire l'inclinazione e la rotazione. I giocatori possono mimare delle azioni e "sentirle" attraverso delle vibrazioni invece che premere semplicemente pulsanti. Uno dei primi video promozionali del Wii mostrava attori mimare azioni come pescare, cucinare, suonare la batteria, dirigere un'orchestra, sparare, combattere con le spade e dilettarsi in un'operazione di chirurgia dentale. Per comunicare con i sensori, il controller del Wii usa una tecnologia Bluetooth. Un pad direzionale digitale è posizionato sulla parte superiore del controller, con un grosso tasto A in stile GameCube sotto e un grilletto nella parte inferiore che funziona da tasto B. Sotto il tasto A vi è una fila di tre piccoli tasti: Start, Home e Select. Vicino all'estremità inferiore del controller due bottoni aggiuntivi contrassegnati dalle lettere minuscole a e b (nei video promozionali erano X e Y con a e b leggibili con il controller ruotato in modo antiorario di 90 gradi), suggerendo che l'utente potrebbe girare il controller di 90 gradi e renderlo un classico controller del NES; si è in effetti scoperto che la stessa operazione trasforma il telecomando in un volante automobilistico. Sotto questi bottoni vi sono quattro luci che indicano quale "porta" sta venendo usata: attualmente possono giocare fino a quattro persone su un singolo Wii. Il controller ha una porta di espansione nel suo lato inferiore che permette di collegare diversi tipi periferiche. Nintendo ha rivelato che una di queste espansioni per il controller sarà una piccola unità di forma ovoidale composta da un joystick analogico e due pulsanti dorsali. La connessione all'unità principale del controller avviene tramite un breve cavo, da cui il nome "Nunchaku". Anche il "Nunchaku" ha al suo interno alcuni sensori di posizione. Nintendo ha annunciato inoltre un controller "shell", il "Classic-Style Expansion Controller", che assomiglia ad un gamepad tradizionale. Il "telecomando" si inserirà dentro questa espansione, che quindi potrà anche ritenere la sensibilità del "telecomando". Nonostante la somiglianza del controller alle light guns, che sono solo compatibili con le televisioni a scansione progressiva, Nintendo assicura che il Wii e il suo controller saranno compatibili con tutte le televisioni, incluse quelle a proiezioni digitali, come è ovvio che sia, considerata la totale differenza nella tecnologia di puntamento.

Last unsolved questions... The accelerometers can measure translations in space but they are prone to drift which aggregates over time, so they don't work as well when you want to keep track of position (like for a pointing game, or an FPS). Accelerometers are more for detecting quick jolts with the controller in a particular direction (stabbing down with the nunchuck in Zelda to finish enemies, or forward in Prime to fire grapple beam). Both the Nunchuck and the dualShock are using accelerometers, but not position tracking. The sensor bar is for locking the position of the remote in 3d space. It provides an external reference so the remote can calculate its position in relation to the sensor nodes. The accelerometer data can be used to help when the sensor bar fails. If you wave the controller around and point the remote at the sensor again, the game (hopefully) can go off of acceleration data until the tracker bar finds the eye again. The two work together to try and cover all necessary situations with good accuracy. The Accelerometers will determine orientation and motion of the controller and the CMOS sensors will detect location in 3d space and motion as well. The controller has an IR filter over the front of it so that it only sees IR light. Hopefully this wouldn't make it a very good camera. Plus, it looks like it has some image processing built in so that it's only sending information about the points (LEDs) it sees; it's not sending back a whole image to the console, that would require too much communication (be slow and drain battery power). With 2 LEDs in the sensor bar (really the bar has a 2x4 sensor array), the controller can tell what direction it is pointed with respect to the LEDs, and it can tell how far away it is from the bar (if it can see both LEDs). It cannot detect its own position in 3D space, though. You'd need a minimum of 3 non-collinear LEDs for that purpose (4 is better). The optical sensor can nail 4 degrees of freedom, limiting its possible position and orientation to being on a sphere (the centre of the sphere is not the centre of the sensor bar) with the orientation being different at each point on the sphere. Therefore by using the tilt sensors (2 degrees of freedom) it can calculate where on the sphere it lies, double checking with the accelerometers (another 4 degrees of freedom). Ingenious, I wouldn't have thought of this and I'm very impressed Nintendo has got this pointing device working as well as it seems to from the demos. It actually only tells its exact position and orientation within a reflection...although that'd be through the other side of the TV. This problem can also be solved by making each end of the sensor bar "look" different, e.g. different wavelength.

A

appendice



A.0. Appendice: alcuni concetti

Intelligenza artificiale forte

Nella filosofia dell'intelligenza artificiale l'intelligenza artificiale forte è l'idea che opportune forme di intelligenza artificiale possano veramente ragionare e risolvere problemi; l'intelligenza artificiale forte sostiene che è possibile per le macchine diventare sapienti o coscienti di sé, senza necessariamente mostrare processi di pensiero simili a quelli umani. Il termine intelligenza artificiale forte (AI forte, in inglese strong AI) fu originalmente coniato da John Searle, che scrisse:

«according to strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind [1]»

Il termine "intelligenza artificiale" dovrebbe equivalere allo stesso concetto di ciò che chiamiamo "IA forte", basandoci sul significato letterale di "artificiale" e "intelligenza". L'attività di ricerca iniziale sull'intelligenza artificiale, comunque, si concentrò su alcuni campi ristretti, quali il pattern recognition e la scheduling automatico, nella speranza di poter ricavare così una comprensione della vera intelligenza. Il termine "intelligenza artificiale" venne in questo modo ad indicare, oltre all'idea dell'IA forte, anche il lavoro svolto in questi settori limitati ("IA debole").

Intelligenza artificiale debole

In contrasto con l'intelligenza artificiale forte, l'intelligenza artificiale debole si riferisce all'uso di programmi per studiare o risolvere specifici problemi o ragionamenti che non possono essere compresi pienamente (o in alcuni casi, sono completamente al di fuori) nei limiti delle capacità cognitive umane. Un esempio di programma di intelligenza artificiale debole è un algoritmo per il gioco degli scacchi (si ricordi, ad esempio, Deep Blue). Diversamente dall'intelligenza artificiale forte, quella debole non realizza un'auto-consapevolezza e non dimostra il largo intervallo di livelli di abilità cognitive proprio dell'uomo, ma è esclusivamente un problem-solver (risolutore di problemi) specifico e, parzialmente, intelligente.

Alcuni sostengono che i programmi di intelligenza artificiale debole non possano essere chiamati propriamente "intelligenti", in quanto non possono realmente pensare. Rispondendo alla tesi secondo cui programmi come Deep Blue non siano realmente pensanti, Drew McDermott scrisse:

«Saying Deep Blue doesn't really think about chess is like saying an airplane doesn't really fly because it doesn't flap its wings.»

Egli sostenne che Deep Blue possieda un'intelligenza che difetta riguardo all'ampiezza stessa del suo intelletto. Altri notano invece che Deep Blue è meramente un potente albero di ricerca euristico, e che affermare che esso "pensi" agli scacchi è come affermare che gli organismi unicellulari "pensino" al processo di sequenza proteica; entrambi sono ignari di tutto, ed entrambi seguono un programma codificato al loro interno. Molti fra coloro che avanzano queste critiche riconoscono l'IA debole come l'unica possibile, affermando che le macchine non potranno mai divenire realmente intelligenti. Al contrario, i sostenitori della IA forte teorizzano la vera coscienza di sé ed il "pensiero" per come lo intendiamo ora possano richiedere uno speciale algoritmo progettato per osservare e prendere in considerazione i processi della propria stessa mente. Secondo alcuni psicologi dell'evoluzione gli umani potrebbero aver sviluppato un algoritmo di questo tipo, particolarmente avanzato, specificamente per l'interazione sociale o la mistificazione, due attività in cui il genere umano si dimostra nettamente superiore rispetto ad altre specie.

Intelligenza artificiale generale

La ricerca sull'intelligenza artificiale generale ha come obiettivo la creazione di una IA capace di replicare completamente l'intelligenza umana, solitamente chiamata Intelligenza Artificiale Generale ("Artificial General Intelligence", AGI) per distinguerla da progetti di IA meno ambiziosi. Ad oggi i ricercatori hanno mostrato poco interesse verso l'AGI, perlopiù seguendo la teoria per cui un'intelligenza umana è troppo complessa per essere replicata completamente. In ogni caso, alcuni gruppi indipendenti di informatici stanno portando avanti progetti di ricerca in questo campo. Tra le organizzazioni che perseguono ricerche sull'AGI sono presenti l'Adaptive AI, Artificial General Intelligence Research Institute (AGIRI), CCortex, Novamente LLC e il Singularity Institute for Artificial Intelligence. Una recentemente aggiuntasi è Numenta, il cui progetto è basato sulle teorie di Jeff Hawkins, creatore del Palm Pilot. Mentre Numenta si basa su un approccio computazionale verso l'intelligenza artificiale, Hawkins è inoltre il fondatore del RedWood Neuroscience Institute, che esplora il pensiero cosciente da una prospettiva biologica.

Filosofia dell'IA forte e della coscienza

John Searle e molti altri di coloro che sono coinvolti in questo dibattito discutono sul se una macchina che lavora tramite la sola trasformazione di dati codificati possa essere considerata una mente, mentre non considerano la più ampia questione del monismo opposto al dualismo, e cioè il problema del se una macchina di qualsiasi genere, includendo quelle biologiche, possa contenere una mente.

Searle sostiene nella sua argomentazione nota come la "Stanza Cinese" che gli elaboratori di informazione portino dati codificati che descrivono cose. I dati codificati in quanto tali risultano essere senza significato alcuno qualora venga a mancare un riferimento incrociato alle cose che essi descrivono. Questo porta Searle a dire che non c'è alcuna comprensione e significato nello stesso elaboratore di informazione. Come risultato egli dichiara che anche una macchina che superi il Test di Turing non debba essere necessariamente intelligente in senso umano.

Alcuni filosofi ritengono che, se l'IA debole è possibile, allora debba essere possibile anche l'IA forte. Daniel C. Dennett, nel suo "Consciousness Explained", sostiene che, se non c'è nessuna scintilla magica o anima, allora l'Uomo non sia altro che una macchina, e si chiede perché questo Uomo-macchina debba avere la posizione privilegiata su tutte le altre macchine possibili per quanto riguarda l'intelligenza o l'averne una "mente". Nello stesso lavoro propone il suo Modello a bozze multiple di coscienza. Simon Blackburn, nella sua introduzione alla filosofia "Think", fa notare che, sebbene una entità possa apparire intelligente, non vi è alcun modo per stabilire se quella intelligenza sia effettivamente reale (e cioè una "mente"). Comunque, se la discussione è limitata alla sola intelligenza artificiale forte, piuttosto che alla coscienza artificiale, è possibile identificare caratteristiche della mente umana che non occorrono nell'elaborazione delle informazioni in un computer.

Molti dei propositori dell'IA forte credono che la mente sia soggetta alla Tesi di Church-Turing. Questa teoria è tuttavia reputata da molti problematica e anti-intuitiva, in quanto un elaboratore di informazioni dovrebbe poter essere costruito anche basandosi solo su sfere e legnetti (si intenda per questo un meccanismo completamente analogico). Sebbene un tale dispositivo sarebbe alquanto lento e facile agli errori, dovrebbe essere in grado di svolgere qualsiasi cosa possa essere fatta da un moderno computer. Se la mente è Turing-compatibile, almeno in principio, un dispositivo fatto interamente di sfere rotanti e canali di legno può allora contenere una mente cosciente.

Roger Penrose ha attaccato l'applicabilità della tesi di Church-Turing direttamente, portando l'attenzione sull'halting problem, secondo cui certi tipi di computazione non possono essere eseguiti da sistemi informativi, sebbene possano essere risolti dalla mente umana.

La possibilità di creare una intelligenza artificiale forte dipende in definitiva dalla possibilità da parte di un elaboratore di informazioni artificiale di includere tutte le caratteristiche di una mente, fra cui la coscienza. Da questo punto di vista, l'intelligenza artificiale debole risulta essere slegata dal problema dell'IA forte. Basti pensare, ad esempio, che molti dei sistemi informativi utilizzati oggi sarebbero stati definiti "intelligenti" solo un secolo fa.

Metodi di realizzazione: simulazione computerizzata del modello cerebrale umano

Questa sembra essere la via più veloce per realizzare un'intelligenza artificiale forte, dal momento che non richiede una comprensione totale della mente umana. Necessita tre cose:

Hardware. Per realizzare questo modello sarebbe necessario un calcolatore estremamente potente: per il futurista Ray Kurzweil 1 milione MIPS. Seguendo la Legge di Moore, tale macchina sarà disponibile nel 2020 al costo di 1500€.

Software. Questa è considerata la parte difficile. Bisogna assumere che la mente umana sia data dal sistema nervoso centrale ed esso sia governato dalle leggi fisiche.

Comprensione. Infine, richiederebbe sufficiente conoscenza dei meccanismi mentali da essere in grado di riprodurli matematicamente. Ciò si potrebbe fare studiando il funzionamento del sistema nervoso centrale, o mappandolo e copiandolo. Le tecniche di neuroimaging migliorano rapidamente, Kurzweil prevede che una mappa di sufficiente qualità verrà creata all'incirca per lo stesso periodo in cui sarà disponibile la necessaria velocità di calcolo.

Una volta messo a punto, un tale modello potrà essere facilmente modificato e sperimentato, facendo progredire la ricerca sulla psiche umana, che a sua volta permetterà di migliorare il modello.

Future applicazioni: intelligenza artificiale a seme/Singularità tecnologica

Un'intelligenza artificiale forte potrebbe a questo punto migliorare ricorsivamente, ossia, partendo dal livello umano, migliorare autonomamente sé stessa, producendo tecnologie molto più velocemente degli scienziati umani. Sarebbe impossibile prevedere gli sviluppi di una tale intelligenza. Assumendo di prendere l'approccio del modello funzionale umano, sono necessarie alcune modifiche perché ciò avvenga. Le più significative sarebbero le modifiche alle motivazioni. La psicologia evolutiva sostiene che gli esseri umani sono completamente motivati da un intricato insieme di, 'desiderio per l'anticipazione del piacere' e 'desiderio per l'anticipazione dell'evitamento del dolore', sviluppati tramite la selezione naturale. Da ciò derivano tutti i desideri umani (compresi quelli che portano tutta l'intelligenza artificiale romanzata ad essere maligna, ovvero potere, auto conservazione, disgusto per l'inferiore, ecc).

Con la comprensione del modello, tutti i desideri del modello possono essere rimossi e possono esserne aggiunti di nuovi - essendo l'auto-miglioramento ricorsivo necessario per una singularità tecnologica. Si può sostenere che la cosa più importante sarebbe di equipaggiare l'intelligenza

artificiale a seme (che probabilmente raggiungerebbe poteri divini) solo con il desiderio di servire l'umanità - implicito in questo è l'auto-miglioramento. per questo motivo è stato istituito il Singularity Institute for Artificial Intelligence. Nota: se la psicologia evolutiva ha torto, sarebbe possibile scoprirlo con questo modello.

Le arti

Un'intelligenza artificiale più brillante di quella umana sarebbe (presumendo il funzionalismo) superiore in campo artistico oltre a quello scientifico. Quindi (in particolare se originariamente un'Intelligenza artificiale a seme) sarebbe in grado di produrre le migliori opere di musica, arte, letteratura e filosofia che il mondo abbia mai visto, potrebbe anche inventare nuove forme d'arte.

Robotica cognitiva

La robotica cognitiva prevede l'applicazione di diversi campi dell'intelligenza artificiale alla robotica. L'intelligenza artificiale forte sarebbe in particolare una risorsa preziosa per questo campo.

Confronto tra i calcolatori e il cervello umano: parallelismo contro velocità

La potenza del cervello umano sta nell'eseguire più operazioni contemporaneamente, mentre quella di una macchina nella velocità con cui sono eseguite queste operazioni. Il cervello umano può effettuare un gran numero di operazioni al secondo, in virtù del fatto che possiede all'incirca 100 bilioni di neuroni che operano simultaneamente, collegati da un centinaio di trilioni di sinapsi. Un moderno computer desktop ha al massimo due unità di calcolo.

Tuttavia, si stima che un neurone emetta 200 pulsazioni al secondo e questo limita il numero di operazioni. Tra loro i segnali sono trasmessi ad una velocità massima di 150 metri al secondo. Un moderno processore da 2 GHz esegue 2 bilioni di operazioni al secondo, e i segnali nelle componenti elettroniche viaggiano quasi alla velocità della luce (300.000 chilometri al secondo).

Se la nanotecnologia ha permesso di costruire dispositivi di dimensioni pari a quelle di un cervello umano e veloci quanto un moderno computer, un modello umano avvertirebbe lo scorrere del tempo più lentamente rispetto ad un uomo. Per questo un minuto potrebbe essere percepito molto più lungo da un cervello artificiale, probabilmente come se durasse qualche ora. Comunque, dal momento che la percezione della durata di un arco di tempo è differente dalla durata stessa, il modo in cui un'intelligenza artificiale tratta il tempo dipenderebbe dai calcoli e dal tipo specifico di cognizione durante quel periodo di tempo.

Bibliografia

Goertzel, Ben; Pennachin, Cassio (Eds.) Artificial General Intelligence. Springer: 2006. ISBN 354023733X.

John Searle: Minds, Brains and Programs Behavioral and Brain Sciences 3 (3): 417-457 1980.

Jose Maria de Espona: TUE: Tridimensional Understanding Engines, A method for artificial understanding and consciousness based on computer graphics tools and an euclidean reference system 2004.

Kurzweil, Ray; The Singularity is Near. ISBN 0670033847.

Expanding Frontiers of Humanoid Robots

How Intelligent is Deep Blue? di Drew McDermott. New York Times, 14 maggio 1997

Data mining

In informatica, il data mining (letteralmente: estrazione da una miniera di dati) è l'estrazione di informazione utile, eseguita in modo automatico o semiautomatico, da grandi quantità di dati. Questo tipo di attività è cruciale in molti ambiti della ricerca scientifica, ma anche in altri settori (per esempio in quello delle ricerche di mercato). Le tecniche e gli algoritmi di data mining hanno lo scopo di analizzare vasti campioni di dati, allo scopo di identificare interessanti regolarità dette pattern. I pattern così identificati possono essere, nella ricerca scientifica, il punto di partenza per ipotizzare e quindi verificare nuove relazioni di tipo causale fra fenomeni; in generale, possono servire in senso statistico per formulare previsioni su nuovi insiemi di dati.

Un concetto correlato al "data mining" è quello di machine learning (apprendimento automatico); infatti, l'identificazione di pattern può paragonarsi all'apprendimento, da parte del sistema di data mining, di una relazione causale precedentemente ignota, cosa che trova applicazione in ambiti come quello degli algoritmi euristici e della intelligenza artificiale. Tuttavia, occorre notare che il processo di data mining è sempre sottoposto al rischio di "rivelare" relazioni causali inesistenti. Una tecnica molto diffusa per il data mining è l'apprendimento mediante classificazione. Questo schema di apprendimento parte da un insieme ben definito di esempi di classificazione per casi noti, dai quali ci si aspetta di dedurre un modo per classificare esempi non noti. Tale approccio viene anche detto con supervisione (supervised), nel senso che lo schema di apprendimento opera sotto la supervisione fornita implicitamente dagli esempi di classificazione per i casi noti; tali esempi, per questo motivo, vengono anche detti training examples, ovvero esempi per l'addestramento. La conoscenza acquisita per apprendimento mediante classificazione può essere rappresentata con alberi di decisione.

Programmazione genetica

La programmazione genetica (GP) è una metodologia di programmazione automatizzata ispirata dall'evoluzione biologica per scoprire programmi informatici che svolgano in maniera (quasi) ottimale un determinato compito. È una particolare tecnica di apprendimento automatico che usa un algoritmo evolutivo per ottimizzare una popolazione di programmi di computer secondo un paesaggio adattativo determinato dall'abilità del programma di arrivare ad un risultato computazionalmente valido (ovvero di saper svolgere il compito dato). I primi esperimenti con la GP sono stati eseguiti da Stephen F. Smith (1980) e Michael L. Cramer (1985), come descritto nel famoso libro Genetic Programming: On the Programming of Computers by Means of Natural Selection di John Koza (1992).

I programmi creati con la GP possono essere scritti in molti linguaggi di programmazione. Nelle prime e tradizionali implementazioni della GP le istruzioni e i dati erano organizzati in strutture ad albero, quindi si preferiva l'uso di linguaggi che avessero queste strutture come tipo di dato

primitivo; un esempio importante di linguaggio utilizzato da Koza è il Lisp. Sono state suggerite e implementate con successo anche altre forme di GP, come la più semplice rappresentazione lineare che ben si adatta ai normali linguaggi imperativi (vedere a riguardo, Banzhaf e al. (1998)). Il software commerciale che implementa la GP Discipulus, ad esempio, usa la programmazione genetica lineare combinata coi linguaggi in codice macchina per ottenere migliori prestazioni. In maniera diversa il MicroGP usa una rappresentazione interna simile alla programmazione genetica lineare per generare programmi che utilizzino pienamente la sintassi di un dato linguaggio assembly.

Poiché la programmazione genetica richiede una grande capacità computazione, negli anni 90 fu usata principalmente per risolvere problemi relativamente semplici. Tuttavia, più recentemente, grazie ai recenti sviluppi nella tecnologia GP e alla ben conosciuta legge di Moore, la GP ha cominciato a fornire un certo numero di risultati. Al momento della scrittura di questo articolo, sono stati ottenuti almeno quaranta risultati competitivi con gli esseri umani, in aree come l'informatica quantistica, il design di componenti elettronici, i giochi, gli algoritmi di ordinamento, di ricerca e molte altre aree. Questi risultati includono la replicazione di molte invenzioni fatte dopo l'anno 2000 e la produzione di due invenzioni brevettabili.

Lo sviluppo di una teoria per la programmazione genetica fu compito arduo, cosa per cui, negli anni 90, fu considerata una sorta di paria tra le varie tecniche di ricerca. Tuttavia, dopo una serie di risultati positivi nei primi anni dopo il 2000, la teoria della GP ha avuto un formidabile e rapido sviluppo; così rapido che è stato possibile costruire modelli probabilistici esatti della GP (teorie, diagrammi e modelli tramite catene di Markov) e mostrare che la GP è più generale e come tale include gli algoritmi genetici. Le tecniche di programmazione genetica sono state applicate all'hardware evolutivo ed ai programmi per computer. La programmazione meta-genetica è la tecnica utilizzata per fare evolvere un sistema di programmazione genetica utilizzando la stessa programmazione genetica. I critici hanno sostenuto che ciò è in teoria impossibile, ma l'argomento necessita di approfondimento e di ulteriori ricerche.

Bibliografia

Banzhaf, W., Nordin, P., Keller, R.E., Francone, F.D. (1998), Genetic Programming: An Introduction: On the Automatic Evolution of Computer Programs and Its Applications, Morgan Kaufmann

Cramer, Michael Lynn (1985), "A representation for the Adaptive Generation of Simple Sequential Programs" in Proceedings of an International Conference on Genetic Algorithms and the Applications, Grefenstette, John J. (ed.), Carnegie Mellon University

Koza, J.R. (1990), Genetic Programming: A Paradigm for Genetically Breeding Populations of Computer Programs to Solve Problems, Stanford University Computer Science Department technical report STAN-CS-90-1314. A thorough report, possibly used as a draft to his 1992 book.

Koza, J.R. (1992), Genetic Programming: On the Programming of Computers by Means of Natural Selection, MIT Press

Koza, J.R. (1994), Genetic Programming II: Automatic Discovery of Reusable Programs, MIT Press

Koza, J.R., Bennett, F.H., Andre, D., and Keane, M.A. (1999), Genetic Programming III: Darwinian Invention and Problem Solving, Morgan Kaufmann

Koza, J.R., Keane, M.A., Streeter, M.J., Mydlowec, W., Yu, J., Lanza, G. (2003), Genetic Programming IV: Routine Human-Competitive Machine Intelligence, Kluwer Academic Publishers

Langdon, W. B., Poli, R. (2002), Foundations of Genetic Programming, Springer-Verlag

Smith, S.F. (1980), A Learning System Based on Genetic Adaptive Algorithms, tesi di dottorato (University of Pittsburgh)

Rete neurale

Una rete neurale è un insieme di neuroni biologici tra loro interconnessi. Nell'uso moderno si intende però di solito con rete neurale una rete di neuroni artificiali, che cerca di simulare il funzionamento dei neuroni all'interno di un sistema informatico. Può essere composta sia da programmi che da hardware dedicato. Spesso viene utilizzata in congiunzione alla logica fuzzy.

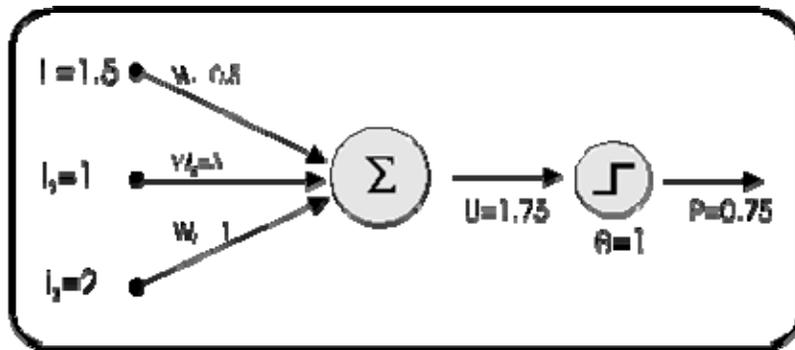
Fondamenti biologici

In quasi tutti gli organismi viventi sono presenti complesse organizzazioni di cellule nervose, con compiti di riconoscimento delle configurazioni assunte dall'ambiente esterno, memorizzazione e reazione agli stimoli provenienti dallo stesso. Il cervello umano rappresenta probabilmente il più mirabile frutto dell'evoluzione per le sue capacità di elaborare informazioni. Al fine di compiere tali operazioni, le reti biologiche si servono di un numero imponente di semplici elementi computazionali (neuroni) fittamente interconnessi in modo da variare la loro configurazione in risposta agli stimoli esterni: in questo senso può parlarsi di apprendimento ed i modelli artificiali cercano di catturare questo tratto distintivo della biologia.

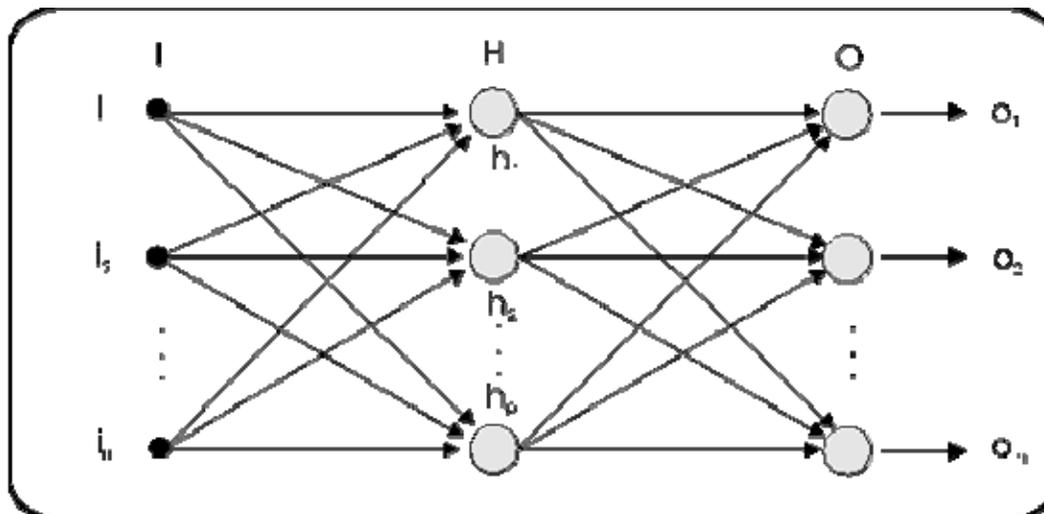
Tipicamente un neurone è costituito di 3 parti principali: il soma (corpo cellulare), l'assone (linea di uscita del neurone unica ma che si dirama in migliaia di rami) e il dendrite (linea di entrata del neurone che riceve segnali in ingresso da altri assoni tramite sinapsi. Il corpo cellulare fa una "somma pesata" (integrazione) dei segnali in ingresso. Se il risultato eccede un certo valore di soglia allora il neurone si attiva ed è prodotto un "potenziale di azione" che è mandato all'assone. Se il risultato non eccede il valore di soglia, il neurone rimane in uno stato di riposo. Una rete neurale artificiale riceve segnali esterni su uno strato di nodi (unità di elaborazione) d'ingresso, ciascuno dei quali è collegato con numerosi nodi interni, organizzati in più livelli. Ogni nodo elabora i segnali ricevuti e trasmette il risultato a nodi successivi.

Funzionamento di una rete neurale feedforward

Le reti neurali si basano principalmente sulla simulazione di neuroni artificiali opportunamente collegati. Il modello rappresentato in figura è quello proposto da McCulloch e Pitts.



I suddetti neuroni ricevono in ingresso degli stimoli e li elaborano. L'elaborazione può essere anche molto sofisticata ma in un caso semplice si può pensare che i singoli ingressi vengano moltiplicati per un opportuno valore detto peso, il risultato delle moltiplicazioni viene sommato e se la somma supera una certa soglia il neurone si attiva attivando la sua uscita. Il peso indica l'efficacia sinaptica della linea di ingresso e serve a quantificarne l'importanza, un ingresso molto importante avrà un peso elevato, mentre un ingresso poco utile all'elaborazione avrà un peso inferiore. Si può pensare che se due neuroni comunicano fra loro utilizzando maggiormente alcune connessioni allora tali connessioni avranno un peso maggiore.



I singoli neuroni vengono collegati alla schiera di neuroni successivi, in modo da formare una rete di neuroni. Normalmente una rete è formata da tre strati. Nel primo abbiamo gli ingressi (I), questo strato si preoccupa di trattare gli ingressi in modo da adeguarli alle richieste dei neuroni. Se i segnali in ingresso sono già trattati può anche non esserci. Il secondo strato è quello nascosto (H, hidden), si preoccupa dell'elaborazione vera e propria e può essere composto anche da più colonne di neuroni. Il terzo strato è quello di uscita (O) e si preoccupa di raccogliere i risultati ed

adattarli alle richieste del blocco successivo della rete neurale. Queste reti possono essere anche molto complesse e coinvolgere migliaia di neuroni e decine di migliaia di connessioni.

Storia

L'ampia varietà di modelli non può prescindere dal costituente di base, il neurone artificiale proposto da W.S. McCulloch e W. Pitts in un famoso lavoro del 1943, il quale schematizza un combinatore lineare a soglia, con dati binari multipli in entrata e un singolo dato binario in uscita: un numero opportuno di tali elementi, connessi in modo da formare una rete, è in grado di calcolare semplici funzioni booleane. Nel 1958, F. Rosenblatt introduce il primo schema di rete neurale, detto perceptron (perceptrone), antesignano delle attuali reti neurali, per il riconoscimento e la classificazione di forme, allo scopo di fornire un'interpretazione dell'organizzazione generale dei sistemi biologici. Il modello probabilistico di Rosenblatt è quindi mirato all'analisi, in forma matematica, di funzioni quali l'immagazzinamento delle informazioni, e della loro influenza sul riconoscimento dei patterns; esso costituisce un progresso decisivo rispetto al modello binario di McCulloch e Pitts, perché i suoi pesi sinaptici sono variabili e quindi il perceptrone è in grado di apprendere.

L'opera di Rosenblatt stimola una quantità di studi e ricerche, e suscita un vivo interesse e notevoli aspettative nella comunità scientifica, destinate tuttavia ad essere notevolmente ridimensionate allorché nel 1969 Marvin Minsky e Seymour A. Papert mostrano i limiti operativi delle semplici reti a due strati basate sui perceptroni, e dimostrano l'impossibilità di risolvere per questa via molte classi di problemi, ossia tutti quelli non caratterizzati da separabilità lineare delle soluzioni: questo tipo di rete neurale non è abbastanza potente, infatti non è in grado di calcolare neanche la funzione or esclusivo (XOR). Di conseguenza, a causa di queste limitazioni, ad un periodo di euforia per i primi risultati della cibernetica (come veniva chiamata negli anni '60), segue un periodo di diffidenza durante il quale tutte le ricerche in questo campo non ricevono più alcun finanziamento dal governo degli Stati Uniti d'America; le ricerche sulle reti tendono, di fatto, a ristagnare per oltre un decennio, e l'entusiasmo iniziale risulta fortemente ridimensionato.

Il contesto matematico per addestrare le reti MLP (Multi-Layers Perceptron, ossia perceptrone multistrato) fu stabilito dal matematico americano Paul Werbos nella sua tesi di dottorato (Ph.D.) del 1974. Uno dei metodi più noti ed efficaci per l'addestramento di tale classe di reti neurali è il cosiddetto algoritmo di retropropagazione dell'errore (error backpropagation), proposto nel 1986 da David E. Rumelhart, G. Hinton e R. J. Williams, il quale modifica sistematicamente i pesi delle connessioni tra i nodi, così che la risposta della rete si avvicini sempre di più a quella desiderata. L'algoritmo di backpropagation (BP) è una tecnica d'apprendimento tramite esempi, costituente una generalizzazione dell'algoritmo d'apprendimento per il perceptrone sviluppato da Rosenblatt nei primi anni '60. Mediante questa tecnica era possibile, come detto, trattare unicamente applicazioni caratterizzabili come funzioni booleane linearmente separabili.

L'algoritmo di apprendimento si basa sul metodo della discesa del gradiente che permette di trovare un minimo locale di una funzione in uno spazio a N dimensioni. I pesi associati ai collegamenti tra gli strati di neuroni si inizializzano a valori piccoli e casuali e poi si applica la regola di apprendimento presentando alla rete dei pattern di esempio. Queste reti neurali sono capaci di generalizzare in modo appropriato, cioè di dare risposte plausibili per input che non hanno mai visto. L'addestramento di una rete neurale di tipo BP avviene in due diversi stadi: forward-pass e backward-pass. Nella prima fase i vettori in input sono applicati ai nodi in ingresso con una propagazione in avanti dei segnali attraverso ciascun livello della rete (forward-pass). Durante questa fase i valori dei pesi sinaptici sono tutti fissati. Nella seconda fase la risposta della rete viene confrontata con l'uscita desiderata ottenendo il segnale d'errore. L'errore calcolato è propagato nella direzione inversa rispetto a quella delle connessioni sinaptiche. I pesi sinaptici infine sono modificati in modo da minimizzare la differenza tra l'uscita attuale e l'uscita desiderata (backward-pass).

Tale algoritmo consente di superare le limitazioni del perceptrone e di risolvere il problema della separabilità non lineare (e quindi di calcolare la funzione XOR), segnando il definitivo rilancio delle reti neurali, come testimoniato anche dall'ampia varietà d'applicazioni commerciali: attualmente la BP rappresenta un algoritmo di largo uso in molti campi applicativi.

Reti di Hopfield

Nel 1982, il fisico John J. Hopfield pubblica un articolo fondamentale in cui presenta un modello matematico comunemente noto appunto come rete di Hopfield: tale rete si distingue per "l'emergere spontaneo di nuove capacità computazionali dal comportamento collettivo di un gran numero di semplici elementi d'elaborazione". Le proprietà collettive del modello producono una memoria associativa per il riconoscimento di configurazioni corrotte e il recupero di informazioni mancanti. Inoltre, Hopfield ritiene che ogni sistema fisico può essere considerato come un potenziale dispositivo di memoria, qualora esso disponga di un certo numero di stati stabili, i quali fungano da attrattore per il sistema stesso. Sulla base di tale considerazione, egli si spinge a formulare la tesi secondo cui la stabilità e la collocazione di tali attrattori sono proprietà spontanee di sistemi costituiti, come accennato, da considerevoli quantità di neuroni reciprocamente interagenti.

Tra le nuove idee messe in luce da Hopfield, quella più degna di menzione riguarda il capovolgimento del rapporto, fino allora esistente, tra calcolo e numeri: mentre era universalmente noto che il calcolo producesse numeri, assai meno banale era l'osservazione di Hopfield che, viceversa, anche i numeri potessero spontaneamente generare calcolo, e che questo potesse emergere quale attributo collettivo di sistemi interattivi siffatti. Le applicazioni delle reti di Hopfield riguardano principalmente la realizzazione di memorie associative, resistenti all'alterazione delle condizioni operative, e la soluzione di problemi d'ottimizzazione combinatoriale. Da un punto di vista strutturale, la rete di Hopfield costituisce una rete neurale ricorrente simmetrica, di cui è garantita la convergenza.

Una rete ricorrente è un modello neurale in cui è presente un flusso bidirezionale d'informazioni; in altri termini, mentre nelle reti di tipo feedforward la propagazione dei segnali avviene unicamente, in maniera continua, nella direzione che conduce dagli ingressi alle uscite, nelle reti ricorrenti tale propagazione può anche manifestarsi da uno strato neurale successivo ad uno precedente, oppure tra neuroni appartenenti ad uno stesso strato, e persino tra un neurone e sé stesso.

Reti di Elman

Un significativo e noto esempio di semplice rete ricorrente è dovuto a Jeffrey L. Elman (1990). Essa costituisce una variazione sul tema del perceptrone multistrato, con esattamente tre strati e l'aggiunta di un insieme di neuroni "contestuali" nello strato d'ingresso. Le connessioni retroattive si propagano dallo strato intermedio (e nascosto) a tali unità contestuali, alle quali si assegna peso costante e pari all'unità. In ciascun istante, gli ingressi si propagano nel modo tradizionale e tipico delle reti feedforward, compresa l'applicazione dell'algoritmo d'apprendimento (solitamente la backpropagation). Le connessioni retroattive fisse hanno come effetto quello di mantenere una copia dei precedenti valori dei neuroni intermedi, dal momento che tale flusso avviene sempre prima della fase d'apprendimento. In questo modo la rete di Elman tiene conto del suo stato precedente, cosa che le consente di svolgere compiti di previsione di sequenze temporali che sono difficilmente alla portata dei perceptroni multistrato convenzionali.

Modalità d'apprendimento

Dal punto di vista della modalità d'apprendimento, è necessario operare una distinzione tra almeno tre diverse tipologie. In particolare, si può avere:

- un apprendimento supervisionato (supervised learning), qualora si disponga di un insieme di dati per l'addestramento (o training set) comprendente esempi tipici d'ingressi con le relative uscite loro corrispondenti: in tal modo la rete può imparare ad inferire la relazione che li lega. Successivamente, la rete è addestrata mediante un opportuno algoritmo (tipicamente, la backpropagation che è appunto un algoritmo d'apprendimento supervisionato), il quale usa tali dati allo scopo di modificare i pesi ed altri parametri della rete stessa in modo tale da minimizzare l'errore di previsione relativo all'insieme d'addestramento. Se l'addestramento ha successo, la rete impara a riconoscere la relazione incognita che lega le variabili d'ingresso a quelle d'uscita, ed è quindi in grado di fare previsioni anche laddove l'uscita non è nota a priori; in altri termini, l'obiettivo finale dell'apprendimento supervisionato è la previsione del valore dell'uscita per ogni valore valido dell'ingresso, basandosi soltanto su un numero limitato di esempi di corrispondenza (vale a dire, coppie di valori input-output). Per fare ciò, la rete deve essere infine dotata di un'adeguata capacità di generalizzazione, con riferimento a casi ad essa ignoti. Ciò consente di risolvere problemi di regressione o classificazione.

- un apprendimento non supervisionato (unsupervised learning), basato su algoritmi d'addestramento che modificano i pesi della rete facendo esclusivamente riferimento ad un insieme di dati che include le sole variabili d'ingresso. Tali algoritmi tentano di raggruppare i dati d'ingresso e di individuare pertanto degli opportuni cluster rappresentativi dei dati stessi, facendo uso tipicamente di metodi topologici o probabilistici. L'apprendimento non supervisionato è anche impiegato per sviluppare tecniche di compressione dei dati.

- un apprendimento per rinforzo (reinforcement learning), nel quale un opportuno algoritmo si prefigge lo scopo di individuare un certo modus operandi, a partire da un processo d'osservazione dell'ambiente esterno; ogni azione ha un impatto sull'ambiente, e l'ambiente produce una retroazione che guida l'algoritmo stesso nel processo d'apprendimento. Tale classe di problemi postula un agente, dotato di capacità di percezione, che esplora un ambiente nel quale intraprende una serie di azioni. L'ambiente stesso fornisce in risposta un incentivo o un disincentivo, secondo i casi. Gli algoritmi per il reinforcement learning tentano in definitiva di determinare una politica tesa a massimizzare gli incentivi cumulati ricevuti dall'agente nel corso della sua esplorazione del problema. L'apprendimento con rinforzo differisce da quello supervisionato poiché non sono mai presentate delle coppie input-output di esempi noti, né si procede alla correzione esplicita di azioni subottimali. Inoltre, l'algoritmo è focalizzato sulla prestazione in linea, la quale implica un bilanciamento tra esplorazione di situazioni ignote e sfruttamento della conoscenza corrente.

Algoritmo di Backpropagation

L'algoritmo di backpropagation è utilizzato nell'apprendimento con supervisione. Esso permette di modificare i pesi delle connessioni in modo tale che si minimizzi una certa funzione errore E . Tale funzione dipende dal vettore h-esimo di output \overline{out}^h restituito dalla rete, dato il vettore h-esimo di ingresso \overline{x}^h e dal vettore h-esimo di output \overline{y}^h che noi desideriamo (che fa parte del training set). Il training set è dunque un insieme di N coppie di vettori $(\overline{x}^h, \overline{y}^h)$, con $h = 1, \dots, N$. La funzione errore che si deve minimizzare si può scrivere come:

$$\frac{1}{2} \sum_h \sum_k (out_k^h - y_k^h)^2,$$

dove l'indice k rappresenta il valore corrispondente al k -esimo neurone di output. E è una funzione dipendente dai pesi (che in generale variano nel tempo). Per minimizzare $E(w)$ si può usare l'algoritmo della discesa del gradiente (gradient-descent). L'algoritmo parte da un punto generico $\overline{x}^{(0)}$ e calcola il gradiente $\nabla f(\overline{x}^{(0)})$. Il gradiente dà la direzione verso cui muoversi lungo la quale ho il massimo incremento (o decremento se considero $-\nabla$). Definita la direzione ci si muove di una distanza η predefinita a priori e si trova un nuovo punto $\overline{x}^{(1)}$ sul quale è calcolato nuovamente il gradiente. Si continua iterativamente finché il gradiente non è nullo. L'algoritmo di backpropagation può essere diviso in due passi:

Forward pass: l'input dato alla rete è propagato al livello successivo e così via ai livelli successivi (il flusso di informazioni si sposta in avanti, cioè forward). Si calcola dunque $E(w)$, l'errore commesso.

Backward pass: L'errore fatto dalla rete è propagato all'indietro (backward) e i pesi sono aggiornati in maniera appropriata.

Mappe auto-organizzanti o reti SOM (Self-Organizing Maps)

Infine, un'ultima interessante tipologia di rete è costituita dalla cosiddetta mappa auto-organizzante o rete SOM (Self-Organizing Map). Tale innovativa tipologia di rete neurale è stata elaborata da Teuvo Kohonen dell'Università Tecnologica di Helsinki; il suo algoritmo d'apprendimento è senza dubbio una brillante formulazione di apprendimento non supervisionato, e ha dato luogo a un gran numero di applicazioni nell'ambito dei problemi di classificazione. Una mappa o rete SOM è basata essenzialmente su un reticolo o griglia di neuroni artificiali i cui pesi sono continuamente adattati ai vettori presentati in ingresso nel relativo insieme di addestramento.

Tali vettori possono essere di dimensione generica, anche se nella maggior parte delle applicazioni essa è piuttosto alta. Per ciò che riguarda le uscite della rete, al contrario, ci si limita di solito ad una dimensione massima pari a tre, il che consente di dare luogo a mappe 2D o 3D. In termini più analitici, l'algoritmo può essere agevolmente descritto, come accennato, nei termini di un insieme di neuroni artificiali, ciascuno con una precisa collocazione sulla mappa rappresentativa degli output, che prendono parte ad un processo noto come winner takes all (asso pigliatutto), al

termine del quale il nodo avente un vettore di pesi più vicino ad un certo input è dichiarato vincitore, mentre i pesi stessi sono aggiornati in modo da avvicinarli al vettore in ingresso. Ciascun nodo ha un certo numero di nodi adiacenti. Quando un nodo vince una competizione, anche i pesi dei nodi adiacenti sono modificati, secondo la regola generale che più un nodo è lontano dal nodo vincitore, meno marcata deve essere la variazione dei suoi pesi. Il processo è quindi ripetuto per ogni vettore dell'insieme di training, per un certo numero, usualmente grande, di cicli. Va da sé che ingressi diversi producono vincitori diversi. Operando in tal modo, la mappa riesce infine ad associare i nodi d'uscita con i gruppi o schemi ricorrenti nell'insieme dei dati in ingresso. Se questi schemi sono riconoscibili, essi possono essere associati ai corrispondenti nodi della rete addestrata. In maniera analoga a quella della maggioranza delle reti neurali artificiali, anche la mappa o rete SOM può operare in due distinte modalità:

- durante la fase di addestramento si costruisce la mappa, pertanto la rete si configura ed organizza tramite un processo competitivo. Alla rete deve essere fornito il numero più grande possibile di vettori in ingresso, tali da rappresentare fedelmente la tipologia di vettore che le sarà eventualmente sottoposta nella seconda fase;

- nel corso della seconda fase ogni nuovo vettore d'ingresso può essere velocemente classificato o categorizzato, collocandolo in automatico sulla mappa ottenuta nella fase precedente. Vi sarà sempre un unico neurone vincente, quello il cui vettore dei pesi giace a minor distanza dal vettore appena sottoposto alla rete; tale neurone può essere determinato semplicemente calcolando la distanza euclidea tra i due vettori in questione.

Pregi

Le reti neurali per come sono costruite lavorano in parallelo e sono quindi in grado di trattare molti dati. Si tratta in sostanza di un sofisticato sistema di tipo statistico dotato di una buona immunità al rumore; se alcune unità del sistema dovessero funzionare male, la rete nel suo complesso avrebbe delle riduzioni di prestazioni ma difficilmente andrebbe incontro ad un blocco del sistema. I software di ultima generazione dedicati alle reti neurali richiedono comunque buone conoscenze statistiche; il grado di apparente utilizzabilità immediata non deve trarre in inganno, pur permettendo all'utente di effettuare da subito previsioni o classificazioni, seppure con i limiti del caso. Da un punto di vista industriale, risultano efficaci quando si dispone di dati storici che possono essere trattati con gli algoritmi neurali. Ciò è di interesse per la produzione perché permette di estrarre dati e modelli senza effettuare ulteriori prove e sperimentazioni.

Difetti

I modelli prodotti dalle reti neurali, anche se molto efficienti, non sono spiegabili in linguaggio simbolico umano: i risultati vanno accettati "così come sono", da cui anche la definizione inglese delle reti neurali come "black box". Come per qualsiasi algoritmo di modellazione, anche le reti neurali sono efficienti solo se le variabili predittive sono scelte con cura. Non sono in grado di trattare in modo efficiente variabili di tipo categorico (per esempio, il nome della città) con molti valori diversi. Necessitano di una fase di addestramento del sistema che fissi i pesi dei singoli Neuroni e questa fase può richiedere molto tempo, se il numero dei record e delle variabili analizzate è molto grande. Non esistono teoremi o modelli che permettano di definire la rete ottima, quindi la riuscita di una rete dipende molto dall'esperienza del creatore.

Utilizzi

Le reti neurali vengono solitamente usate in contesti dove i dati possono essere parzialmente errati oppure dove non esistano modelli analitici in grado di affrontare il problema. Un loro tipico utilizzo è nei software di OCR, nei sistemi di riconoscimento facciale e più in generale nei sistemi che

si occupano di trattare dati soggetti a errori o rumore. Esse sono anche uno degli strumenti maggiormente utilizzati nelle analisi di Data mining. Le reti neurali vengono anche utilizzate come mezzo previsionistico dell'analisi finanziaria o meteorologica. Negli ultimi anni è aumentata notevolmente la loro importanza anche nel campo della bioinformatica nel quale vengono utilizzate per la ricerca di pattern funzionali e/o strutturali in proteine e acidi nucleici. Mostrando opportunamente una lunga serie di input (fase di training o apprendimento), la rete è in grado di fornire l'output più probabile.

Algoritmo genetico

L'algoritmo genetico appartiene a una particolare classe di algoritmi utilizzati in diversi campi, tra cui l'intelligenza artificiale. È un metodo euristico di ricerca ed ottimizzazione, ispirato al principio della selezione naturale di Charles Darwin che regola l'evoluzione biologica. Questo tipo di algoritmi è detto genetico perché mutuata terminologia dalla genetica, branca della biologia. Gli algoritmi genetici sono applicabili alla risoluzione di un'ampia varietà di problemi d'ottimizzazione non indicati per gli algoritmi classici, compresi quelli in cui la funzione obiettivo è discontinua, non derivabile, stocastica, o fortemente non lineare.

Principi di funzionamento

Un tipico algoritmo genetico parte da un certo numero di possibili soluzioni (individui) chiamate popolazione e provvede a farle evolvere nel corso dell'esecuzione: a ciascuna iterazione, esso opera una selezione di individui della popolazione corrente, impiegandoli per generare nuovi elementi della popolazione stessa, che andranno a sostituire un pari numero d'individui già presenti, e a costituire in tal modo una nuova popolazione per l'iterazione (o generazione) seguente. Tale successione di generazioni evolve verso una soluzione ottima del problema assegnato.

La loro evoluzione viene ottenuta attraverso una parziale ricombinazione delle soluzioni, ogni individuo trasmette parte del suo patrimonio genetico ai propri discendenti, e l'introduzione di mutazioni casuali nella popolazione di partenza, sporadicamente quindi nascono individui con caratteristiche non comprese tra quelle presenti nel corredo genetico della specie originaria. Finita la fase di evoluzione la popolazione delle soluzioni viene analizzata e vengono tenute solo le soluzioni che meglio risolvono il problema: gli individui con le qualità più adatte all'ambiente in cui si trovano hanno quindi maggiori possibilità di sopravvivere e riprodursi. Queste soluzioni subiranno una nuova fase di evoluzione e così via.

Alla fine ci si aspetta di trovare una popolazione di soluzioni che riescano a risolvere adeguatamente il problema posto. Non vi è modo di decidere a priori se l'algoritmo sarà effettivamente in grado di trovare una soluzione accettabile. Di norma gli algoritmi genetici vengono utilizzati per problemi di ottimizzazione per i quali non si conoscono algoritmi di complessità lineare o polinomiale. Un caso particolare di applicazione di algoritmi genetici è Acovea, un software studiato per trovare il profilo migliore delle opzioni di ottimizzazione del compilatore gcc: un problema di elevata complessità.

Dettaglio del funzionamento

La soluzione del problema viene codificata in una struttura, di solito una stringa, detta gene. Inizialmente viene creato un certo numero di geni in maniera casuale e si definisce una funzione che restituisce la "bontà" di un gene come soluzione del problema, detta funzione di fitness. L'algoritmo consiste nell'applicazione di operazioni, che tendono a modificare la popolazione dei geni, nel tentativo di migliorarli in modo da ottenere una soluzione sempre migliore. L'evoluzione procede quindi in passi, per ognuno di questi viene per prima cosa eseguito un ordinamento dei geni sulla

base del risultato della funzione di fitness. Vengono poi eseguite le operazioni su un numero di geni stabilito dai parametri dell' algoritmo, che in generale determinano quanti geni devono subire crossover e mutazioni, e in quale misura.

L'algoritmo evolve quindi attraverso i seguenti punti:

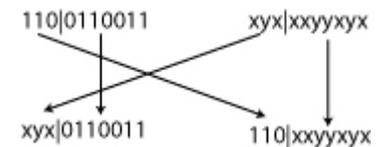
- generazione, in maniera casuale, una popolazione iniziale;
- creazione di una sequenza di nuove popolazioni, o generazioni. In ciascuna iterazione, gli individui della popolazione corrente sono usati per creare la generazione successiva, e a questo scopo si compiono degli ulteriori passi:
- ciascun membro della popolazione corrente è valutato calcolandone il rispettivo valore di fitness (idoneità);
- si determina un opportuno ordinamento di tali individui sulla base dei valori di fitness;
- gli individui più promettenti sono selezionati come genitori;
- a partire da tali individui si genera un pari numero di individui della generazione successiva, e ciò può avvenire secondo due modalità distinte, vale a dire effettuando cambiamenti casuali su un singolo genitore (mutazione) oppure combinando opportunamente le caratteristiche di una coppia di genitori (incrocio);
- gli individui così generati vanno a sostituire i genitori consentendo la formazione della generazione successiva;
- infine, l'algoritmo s'interrompe quando uno dei criteri d'arresto è soddisfatto.

Crossover

In base a un coefficiente stabilito inizialmente, alcune parti dei geni risultati migliori vengono scambiate, nell'ipotesi che questo possa migliorare il risultato della funzione di fitness nel successivo "passo evolutivo".

Single point crossover

Ci sono varie tecniche di crossover. Una delle più semplice è la "single point crossover" che consiste nel prendere due individui e tagliare le loro stringhe di codifica in un punto a caso. Si creano così due teste e due code. A questo punto si scambiano le teste e le code, ottenendo due nuovi geni. Il crossover non è applicato sempre, ma con una probabilità p_c . Nel caso in cui non viene applicato i figli sono semplicemente le copie dei genitori. Sperimentalmente si può vedere che il miglioramento diventa apprezzabile solo dopo un certo numero di passi. Questo a meno di casi fortunati, ovviamente.



Mutazione

La mutazione consiste nella modifica casuale di alcune parti dei geni con valore di fitness più basso, in base a coefficienti definiti inizialmente. Queste modifiche puntano a migliorare il valore della funzione per il gene in questione.

Basi teoriche e storia

La traduzione e l'estensione dei principi esposti, validi per i sistemi biologici e per i sistemi artificiali, si deve storicamente a John Holland. Dopo un non breve periodo di tempo, in cui il rilievo di tale lavoro non fu pienamente riconosciuto, l'impiego degli algoritmi genetici si è andato consolidando in ambito informatico, ingegneristico, finanziario ed ovviamente nel campo delle scienze sociali e naturali. Un teorema, dovuto a Holland, assicura che, sotto determinate ipotesi, gli individui con alti valori di fitness tendono a crescere esponenzialmente nella popolazione attraverso il meccanismo dell'incrocio, assicurando così la convergenza dell'algoritmo genetico verso una soluzione ottimale. Nel suo teorema sugli schemi (schema theorem), detto anche "teorema fondamentale degli algoritmi genetici", egli dimostra che uno schema (ossia una particolare combinazione di geni che occupano posizioni precise all'interno di un cromosoma) prolifera più rapidamente se, oltre ad avere un alto valore di fitness, contiene un piccolo numero di geni specifici non lontani l'uno dall'altro. Ciò, infatti, riduce la probabilità di distruggere lo schema durante la fase di riproduzione.

Brevi successioni di geni, che assumono particolari valori, definiscono i cosiddetti blocchi costitutivi (building blocks): favorendo l'incrocio dei cromosomi meglio adattati, in cui si riscontra statisticamente la presenza di peculiari blocchi costitutivi, l'algoritmo aumenta la probabilità che blocchi costituenti opportuni, provenienti da cromosomi diversi, si ritrovino in uno stesso cromosoma. Assumendo che l'associazione di siffatti blocchi sia dunque vantaggiosa, dovrà anche ritenersi probabile la comparsa di un cromosoma (soluzione) eccellente per il problema in esame, in un tempo ragionevole. La dimostrazione del teorema degli schemi è basata sull'ipotesi di codifica binaria, ma Wright (1991) l'ha estesa al caso di codifica con numeri reali; lo stesso Wright ha mostrato che una codifica reale è da preferirsi nel caso di problemi continui d'ottimizzazione. Herrera e Lozano (1998) hanno poi presentato un'ampia rassegna di operatori genetici applicabili a cromosomi codificati mediante numeri reali, compresi vari tipi di operatori di crossover (incrocio).

Pertanto, il campo dei numeri reali costituisce ormai un'appropriata e consolidata forma di rappresentazione per gli algoritmi genetici in domini continui. Tuttavia, a causa di complessi fenomeni di interazione non lineare (epistaticità) tra gruppi di valori di una stringa rappresentante un individuo, non si può affermare con certezza che la combinazione di blocchi costitutivi altamente performanti sia sempre destinata a produrre individui ancora migliori. In altri termini, non sempre l'operazione genetica di crossover produce risultati accettabili, e anzi a volte accade che, a partire da due genitori estremamente promettenti, si ottenga un discendente decisamente meno valido.

Calcolo evolutivo

In questi casi, il cosiddetto calcolo evolutivo, sviluppato principalmente da David B. Fogel, anch'esso ispirato all'evoluzione naturale ma non alla genetica, può essere impiegato con successo nelle applicazioni. Tale metodologia differisce dagli algoritmi genetici in quanto non utilizza l'operazione genetica di crossover, che invece per essi risulta imprescindibile.

Programmazione genetica

La programmazione genetica, elaborata fundamentalmente ad opera di John R. Koza, è invece un metodo per la generazione automatica di programmi, a partire da una descrizione ad alto livello del task da svolgere, e basato sul principio darwiniano della selezione naturale allo scopo di sviluppare una popolazione di programmi migliorativi nell'arco delle successive generazioni. Essa si avvale di operazioni capaci di alterare l'architettura di detti programmi e di prendere decisioni sull'uso delle subroutine, dei loop, della ricorsione e della memoria.

Da ciò si nota che la programmazione genetica costituisce in sostanza un'estensione degli algoritmi genetici al caso di popolazioni costituite da programmi di dimensione variabile; la programmazione genetica sostituisce in altri termini alla stringa di lunghezza costante, codificata in vario modo, un programma con struttura ad albero, il cui corpo (radice, nodi intermedi) è costituito da funzioni aritmetiche o logiche, mentre i nodi terminali rappresentano variabili o costanti numeriche. Pertanto la popolazione risulta ora composta da un numero opportuno di programmi, i quali mediante le operazioni genetiche di riproduzione (non è prevista alcuna mutazione) producono, in un certo numero di generazioni, il programma che risolve al meglio un problema assegnato, in forma topologica parametrizzata.

Neuroevoluzione

Infine, si denota col termine neuroevoluzione l'uso degli algoritmi genetici, o di altri metodi e tecniche evolutive, nella messa a punto delle reti neurali artificiali, per quanto riguarda sia l'architettura della rete (cioè la sua struttura intesa come numero di nodi e numero di connessioni tra i nodi stessi), sia i parametri relativi (ossia i pesi delle connessioni tra i nodi). Un metodo neuroevolutivo degno di nota è quello proposto recentemente da K. Stanley, denominato NEAT (NeuroEvolution of Augmenting Topologies), e basato su un processo di graduale incremento della complessità strutturale delle reti che si propongono di risolvere un problema assegnato (tipicamente un problema di reinforcement learning). A partire da reti estremamente semplici, in quanto completamente prive di neuroni intermedi, la procedura in questione sembra avere maggiori possibilità di determinare soluzioni efficaci e robuste rispetto a metodi analoghi, che però partono da topologie predeterminate o comunque casuali. I tre principi fondamentali su cui si basa il NEAT sono i seguenti:

- il primo principio è l'omologia: il NEAT codifica ciascun nodo e ciascuna connessione della rete attraverso un gene. Ogni volta che una mutazione strutturale sfocia nella creazione di un nuovo gene, quel gene riceve un contrassegno numerico che lo rende permanentemente rintracciabile. Tale marcatura storica è utilizzata in seguito per verificare la conciliabilità di geni omologhi durante l'operazione di crossover, e per definire un operatore di compatibilità;

- il secondo principio è la protezione dell'innovazione. L'operatore di compatibilità definito in precedenza è adoperato per dividere la popolazione, composta da reti neurali, in specie differenti, allo scopo di proteggere le soluzioni innovative da un'eliminazione prematura, e di prevenire l'incrocio di materiale genetico incompatibile. Tali innovazioni strutturali presentano una significativa possibilità di raggiungere il loro pieno potenziale, in quanto protette dal resto della popolazione attraverso la suddivisione in specie, cioè la creazione di nicchie o spazi riservati;

- da ultimo, il principio secondo cui la ricerca di una soluzione dovrebbe avvenire nel più piccolo spazio possibile (inteso come numero di dimensioni), da espandere poi in maniera graduale. Cominciando il processo evolutivo da una popolazione di elementi a struttura minima, le successive mutazioni topologiche comportano l'aggiunta di nuovi nodi e connessioni alle reti, conducendo pertanto ad una crescita incrementale della popolazione stessa. Dal momento che solo le modifiche strutturali vantaggiose tendono a sopravvivere nel lungo termine, le topologie che vengono raffinate tendono ad essere le minime necessarie alla soluzione del problema assegnato.

Bibliografia

D.B. Fogel, Evolutionary computation: toward a new philosophy of machine intelligence, IEEE Press, NewYork, 1995 ISBN 078035379X

J.R Koza, Genetic programming: on the programming of computers by means of natural selection, MIT Press, Cambridge, MA, 1992 ISBN 0262111705

Alden H. Wright, Genetic Algorithms for Real Parameter Optimization, 1991 [1]

Optical Character Recognition

I sistemi di Optical Character Recognition (riconoscimento ottico dei caratteri detti anche OCR) sono programmi dedicati alla conversione di un'immagine contenente testo in testo modificabile con un normale programma di videoscrittura. Solitamente le immagini sono acquisite da uno scanner d'immagini o da un sistema di digitalizzazione che si avvale di una telecamera o di una webcam. Il testo viene convertito in testo ASCII, Unicode o nel caso dei sistemi più avanzati in un formato in grado di contenere anche l'impaginazione del documento. I programmi di OCR si avvalgono dei progressi dell'intelligenza artificiale e dell'evoluzione degli algoritmi legati al riconoscimento delle immagini.

Letture ottica vs. riconoscimento digitale dei caratteri

Originalmente, le distinzioni fra lettura ottica dei caratteri (usando le tecniche ottiche quali gli specchi e gli obiettivi) e il riconoscimento digitale dei caratteri (usando gli algoritmi di separazione ed analisi del testo) erano notevoli ed infatti erano considerati campi separati. Poiché non è rimasta più quasi nessuna applicazione legata alle tecniche di lettura ottica si è esteso il termine OCR che ora indica il riconoscimento dei caratteri digitali indipendentemente dalla sorgente delle immagini.

Addestramento

I sistemi OCR per funzionare correttamente richiedono una fase di "addestramento". Durante questa fase al sistema vengono forniti degli esempi di immagini col corrispondente testo in formato ASCII o simile in modo che gli algoritmi si possano calibrare sul testo che usualmente andranno ad analizzare. Questo addestramento è fondamentale se si considera che gli elementi che analizzano il testo non sono altro che delle reti neurali e come tali richiedono un addestramento per funzionare. Gli ultimi software di OCR utilizzano algoritmi in grado di riconoscere i contorni e in grado di ricostruire oltre al testo anche la formattazione della pagina.

Breve storia dei programmi di OCR

Il sistema postale degli Stati Uniti d'America utilizza sistemi di OCR fin dal 1965. La necessità di riconoscere le destinazioni delle missive e di organizzarle in modo automatico ha spinto la ricerca nel settore dell'OCR. I sistemi OCR leggono il Codice Postale scritto sulle lettere e provvedono ad stampare sulle missive un codice a barre che rappresenta la destinazione della lettera. Per impedire che il codice a barre disturbi la lettura

dell'indirizzo e quindi complichino il lavoro dei postini il codice a barre viene stampato con un inchiostro visibile solo se illuminato da una luce con lunghezza d'onda nell'Ultravioletto. Il codice a barre viene utilizzato da macchine smistatrici per indirizzare la corrispondenza all'ufficio postale corrispondente che si preoccuperà di recapitarlo al destinatario. Un metodo analogo è in uso dalle poste italiane per la gestione della corrispondenza.

OCR di caratteri stampati

Mentre il riconoscimento esatto di un testo scritto con un alfabeto latino oramai è considerato un problema risolto quasi perfettamente, il riconoscimento della scrittura a mano libera e il riconoscimento degli alfabeti non latini è un problema che tuttora non ha trovato delle soluzioni realmente soddisfacenti e infatti è tuttora oggetto di studi e ricerche.

OCR a mano libera

Sistemi per riconoscere della scrittura a mano libera hanno avuto un discreto successo commerciale se integrati in prodotti tipo PDA o computer portatili. Il precursore di questi dispositivi è stato il dispositivo Newton prodotto dall'Apple. Gli algoritmi di questi dispositivi funzionano adeguatamente perché si impone all'utente di imparare a scrivere le lettere seguendo un certo schema predefinito in modo da minimizzare i possibili casi di ambiguità. Queste strategie non si possono applicare nei documenti scritti su carta infatti il riconoscimento a mano libera è un problema tutt'altro che risolto. I tassi di accuratezza dell'80%-90% sui caratteri scritti a mano in modo accurato e pulito possono essere raggiunti in modo relativamente semplice. Ma un tasso di accuratezza così basso produce diverse decine di errori per pagina rendendo le tecniche di scrittura a mano libera poco utili nella maggior parte dei casi.

OCR del corsivo

Il riconoscimento del testo scritto in corsivo è un campo di ricerca attivo, e attualmente l'accuratezza del riconoscimento è persino inferiore a quella di un testo scritto a mano. Più elevati livelli di accuratezza non saranno possibili fino a che non si useranno informazioni aggiuntive derivate da un'analisi contestuale o grammaticale del testo. Per esempio, riconoscere le intere parole da un dizionario è più facile che provando ad analizzare i diversi caratteri singolarmente: analizzare le parole intere consente di eliminare molte ambiguità legate al riconoscimento. Conoscere il contesto dello scritto consente di eliminare altre ambiguità, per esempio un documento che parla di storia conterrà probabilmente molte date e quindi una linea verticale seguita da un simbolo 9 consentirebbe di ipotizzare che probabilmente la linea è un 1 piuttosto che una l minuscola o una i maiuscola. La conoscenza della grammatica della lingua analizzata può contribuire a determinare se una parola è probabilmente un verbo o un nome, per esempio, consentendo un'accuratezza maggiore. Purtroppo i caratteri corsivi di molte lettere non contengono abbastanza informazioni per effettuare un'analisi corretta e infatti l'accuratezza difficilmente può superare il 98%.

Aree di Ricerca

Un problema particolarmente difficile per i calcolatori e gli esseri umani è quello del riconoscimento di documenti danneggiati contenenti molti nomi o comunque informazioni non deducibili dal contesto. Le pagine possono essere danneggiate dall'età, acqua o dal fuoco e dei nomi possono essere obsoleti o contenere errori d'ortografia. Le tecniche di elaborazione delle immagini dei calcolatori possono aiutare gli esseri umani nella lettura dei testi estremamente antichi come i documenti lasciati da Archimede o i rotoli del mar Morto. L'utilizzo del calcolatore come supporto

all'uomo e viceversa è un ambito di ricerca molto interessante e potenzialmente prolifico. Il riconoscimento dei caratteri è stato un settore soggetto ad un'intensa ricerca fin dai tardi anni cinquanta. Inizialmente è stato percepito come problema semplice, ma è risultato essere un problema molto più interessante. Serviranno ancora decenni di studi prima che il calcolatore sia in grado di riconoscere un testo con la stessa accuratezza di un essere umano, sempre che ciò sia possibile.

MICR

Un'applicazione dove l'esattezza e la velocità di riconoscimento dei sistemi OCR sui caratteri supera quella umana è quella dei MICR, dove l'accuratezza è molto elevata e gli errori variano intorno a un errore rilevato su 20.000 - 30.000 controlli. Questa precisione si ottiene grazie all'utilizzo di inchiostri speciali contenenti materiale magnetico (ossido di ferro).

Statistical pattern recognition, neural networks and learning selected papers

- S. Mitra, S.K. Pal and P. Mitra, "Data mining in soft computing framework: a survey", IEEE Transactions on Neural Networks, vol. 13, no. 1, 2002, 3-14
- J.F. Roddick and M. Spiliopoulou, "A survey of temporal knowledge discovery paradigms and methods", IEEE Transactions on Knowledge and Data Engineering, vol. 14, no. 4, 2002, 750.
- M. Egmont-Petersen, D. de Ridder and H. Handels, "[Image processing using neural networks - a review](#)", Pattern Recognition, vol. 35, no. 10, 2002, 2279-301.
- A.K. Jain, R.P.W. Duin and J. Mao, "[Statistical pattern recognition: a review](#)", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, 2000, 4-37.
- G.P. Zhang, "Neural networks for classification: a survey", IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews, vol. 30, no. 4, 2000, 451-462.
- S. Mitra and Y. Hayashi, "Neuro-fuzzy rule generation: survey in soft computing framework", IEEE Transactions on Neural Networks, vol. 11, no. 3, 2000, 748-768.
- M. Sarkar, "Modular pattern classifiers: a brief survey", IEEE International Conference on Systems, Man, and Cybernetics, vol. 4, 2000, 2878-2883.
- A. Hyvarinen, "[Survey on Independent Component Analysis](#)", Neural Computing Surveys, vol. 1, no. 2, 1999, 94-128.
- Y. Bengio, "Markovian models for sequential data", Neural Computing Surveys, vol. 1, no. 2, 1999, 129-162.
- A. Baraldi and P. Blonda, "A survey of fuzzy clustering algorithms for pattern recognition", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 29, no. 6, 1999, 778-801.
- S.R. Kulkarni, G. Lugosi and Venkatesh, S.S., "Learning pattern classification-a survey", IEEE Transactions on Information Theory, vol. 44, no. 6, 1998, 2178-2206.
- C.J.C. Burges, "A tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, vol. 2, no. 2, 1998, 121-167.
- T.Y. Kwok and D.Y. Yeung, "Constructive algorithms for structure learning in feedforward neural networks for regression problems", IEEE Transactions on Neural Networks, vol. 8, no. 3, 1997, 630-645.
- A.K. Jain, J. Mao, K.M. Mohiuddin, "Artificial Neural Networks: a Tutorial," Computer, vol. 29, no. 3, 1996, 31-44.
- J. Wood, "Invariant pattern recognition: a review," Pattern Recognition, vol. 29, no. 1, 1996, 1-17.
- M. Bianchini and M. Gori, "Optimal learning in artificial neural networks - a review of theoretical results", Neurocomputing, vol. 13, no. 2-4, 1996, 313-346.
- H.H. Thodberg, "Review of Bayesian neural networks with an application to near infrared spectroscopy", IEEE Transactions on Neural Networks, vol. 7, no. 1, 1996, 56-72.
- P.F. Baldi and K. Hornik, "Learning in linear neural networks: A survey," IEEE Transactions on Neural Networks, vol. 6, no. 4, 1995, 837-858.
- Yu.G. Smetanin, "Neural Networks as Systems for Pattern Recognition: a Review," Pattern Recognition and Image Analysis, vol. 5, no. 2, 1995, 254-293.
- A.C. Tsoi and A.D. Back, "Locally recurrent globally feedforward networks - a critical review of architectures," IEEE Transactions on Neural Networks, vol. 5, no. 2, pp. 229-239, 1994.
- R. Reed, "Pruning algorithms - A survey," IEEE Transactions on Neural Networks, vol. 4, no. 5, pp. 740-747, 1993.
- S.J. Raudys and A.K. Jain, "Small sample size effects in statistical pattern recognition: recommendations for practitioners," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 3, pp. 252-264, 1991.
- S.B. Gelfand and E.J. Delp, "On Tree Structured Classifiers," in: Artificial neural networks and statistical pattern recognition, ed. I.K. Sethi and A.K. Jain, pp. 51 - 70, North-Holland, Amsterdam, 1991.
- S.B. Gelfand, C.S. Ravishankar, and E.J. Delp, "An iterative growing and pruning algorithm for classification tree design," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 2, pp. 163- 174, 1991.
- A.K. Jain and B. Chandrasekaran, "Dimensionality and Sample Size Considerations in Pattern Recognition Practice," in: Handbook of Statistics, vol. 2, ed. P.R. Krishnaiah and L.N. Kanal, pp. 835 - 855, North-Holland, Amsterdam, 1987.

J.R. Quinlan, "Simplifying decision trees," *Int. J. Man - Machine Studies*, vol. 27, pp. 221-234, 1987.

K.S. Fu, "Learning control systems: review and outlook," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 3, pp. 327-343, 1986.

K.S. Fu and T.L. Booth, "Grammatical inference: Introduction and survey: Part I," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 3, pp. 343-360, 1986.

K.S. Fu and T.L. Booth, "Grammatical inference: Introduction and survey: Part II," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 3, pp. 360-376, 1986.

B.M.E. Moret, "Decision trees and diagrams," *Computing Surveys*, vol. 14, no. 4, pp. 593-623, 1982.

Image, speech and signal processing

C.C. Chibelushi, F. Deravi, and J.S.D. Mason, "A review of speech-based bimodal recognition", *IEEE Transactions on Multimedia*, vol. 4, no. 1, 2002, 23-37.

J. Zhang and T. Tan, "Brief review of invariant texture analysis methods", *Pattern Recognition*, vol. 35, no. 3, 2002, 735-747.

M.-H. Yang, D.J. Kriegman and N. Ahuja, "Detecting faces in images: a survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, 2002, 34-58.

P.M. Djuric, Yufei Huang and T. Ghirmai, "Perfect sampling: a review and applications to signal processing", *IEEE Transactions on Signal Processing*, vol. 50, no. 2, 2002, 345-356.

A. Nikolaidis, S. Tsekeridou, A. Tefas and V. Solachidis, "A survey on watermarking application scenarios and related attacks", *Proceedings International Conference on Image Processing*, 2001, vol. 3, 991-994.

R.J. Campbell and P.J. Flynn, "A survey of free-form object representation and recognition techniques", *Computer Vision and Image Understanding*, vol. 81, no. 2, 2001, 166-210.

S.D. Olabarriaga and A.W.M. Smeulders, "Interaction in the segmentation of medical images: a survey", *Medical Image Analysis*, vol. 5, no. 2, 2001, 127-142.

A.F. Frangi, W.J. Niessen and M.A. Viergever, "Three-dimensional modeling for functional analysis of cardiac images: a review", *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, 2001, 2-25.

S.F. Chen and R. Rosenfeld, "A survey of smoothing techniques for ME models", *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 1, 2000, 37-50.

J. Battle, A. Casals, J. Freixenet and J. Marti, "A review on strategies for recognizing natural objects in colour images of outdoor scenes", *Image and Vision Computing*, vol. 18, no. 6-7, 2000, 515-530.

T. Randen and J.H. Husoy, "Filtering for texture classification: a comparative study", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, 1999, 291-310.

B. Wohlberg, and G. De Jager, "A review of the fractal image coding literature", *IEEE Transactions on Image Processing*, vol. 8, no. 12, 1999, 1716-1729.

H. Lester and S.R. Arridge, "A survey of hierarchical non-linear medical image registration", *Pattern Recognition*, vol. 32, no. 1, 1999, 129-149.

O. Egger, P. Fleury, T. Ebrahimi and M. Kunt, "High-performance compression of visual information - a tutorial review, I. Still pictures", *Proceedings of the IEEE*, vol. 87, no. 6, 1999, 976-1013.

T.M. Lehmann, C. G'nnner, and K. Spitzer, "Survey: interpolation methods in medical image processing", *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, 1999, 1049-1075.

J. Jiang, "Image compression with neural networks - a survey", *Signal Processing: Image Communication*, vol. 14, no. 9, 1999, 737-760.

J.B.A. Maintz and M.A. Viergever, "A survey of medical image registration", *Medical Image Analysis*, vol. 2, no. 1, 1998, 1-36.

A. Hammoude, "Endocardial border identification in two-dimensional echocardiographic images: review of methods", *Computerized Medical Imaging and Graphics*, vol. 22, no. 3, 1998, 181-193.

A.N. Akansu, P. Duhamel, X. Lin, and M. de Courville, "Orthogonal transmultiplexers in communication: a review", *IEEE Transactions on Signal Processing*, vol. 46, no. 4, 1998, 979-995.

S. Loncaric, "A survey of shape analysis techniques", *Pattern Recognition*, vol. 31, no. 8, 1998, 983-1001.

F.A. Westall, "Review of speech technologies for telecommunications", *Electronics & Communication Engineering Journal*, vol. 9, no. 5, 1997, 197-207.

G. Sharma and H.J. Trussell, "Digital color imaging", *IEEE Transactions on Image Processing*, vol. 6, no. 7, 1997, 901-932.

J.C. Pinoli, "A general comparative study of the multiplicative homomorphic, log-ratio and logarithmic image processing approaches", *Signal Processing*, vol. 58, no. 1, 1997, 11-45.

A.J. den Dekker and A. van den Bos, "Resolution - a survey", *Journal of the Optical Society of America A - Optics and Image Science*, vol. 14, no. 3, 1997, 547-557.

T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey", *Medical Image Analysis*, vol. 1, no. 2, 1996/1997, 91-108.

S. Young, "A review of large-vocabulary continuous-speech", *IEEE Signal Processing Magazine*, vol. 13, no. 5, 1996, 45-37.

Y.J. Zhang, "A survey on evaluation methods for image segmentation", *Pattern Recognition*, vol. 29, no. 8, 1996, 1335-1346.

T. Lindeberg, "[Scale-space theory: A framework for handling image structures at multiple scales](#)," in: *Proc. CERN School of Computing*, Egmond aan Zee, The Netherlands, Sep. 8--21, 1996.

P.C. Cosman, R.M. Gray, and M. Vetterli, "Vector quantization of image subbands: A survey," *IEEE Transactions on Image Processing*, vol. 5, no. 2, 1996, 202-225.

C.F. Barnes, S.A. Rizvi, and N.M. Nasrabadi, "Advances in residual vector quantization: A review," *IEEE Transactions on Image Processing*, vol. 5, no. 2, 1996, 226-262.

J.C. Brailean, R.P. Kleihorst, S. Efstratiadis, A.K. Katsaggelos and R.L. Lagendijk, "Noise reduction filters for dynamic image sequences: a review", *Proceedings of the IEEE*, vol. 83, no. 9, 1995, 1272-1292.

G.M. Blair, "A review of the discrete Fourier transform. 2. Non-radix algorithms, real transforms and noise", *Electronics & Communication Engineering Journal*, vol. 7, no. 5, 1995, 187-194.

[T. Lindeberg Scale-space theory: A basic tool for analysing structures at different scales](#)", *J. of Applied Statistics*, 21(2), pp. 224--270, 1994.

H. Li, A. Lundmark and R. Forchheimer, "Image sequence coding at very low bitrates: a review", *IEEE Transactions on Image Processing*, vol. 3, no. 5, 1994, 589-609.

E.V. Zabrodskii, L.S. Korolyuk, and B.P. Rusyn, "Methods for the detection of homogeneous regions in image recognition: a review", *Pattern Recognition and Image Analysis*, vol. 4, no. 3, 1994, 272-277.

N. Akrouf, R. Prost and R. Goutte, "Image compression by vector quantization: a review focused on codebook generation", *Image and Vision Computing*, vol. 12, no. 10, 1994, 627-637.

T.R. Reed and J.M.H. Dubuf, "A review of recent texture segmentation and feature extraction techniques," *CVGIP - Image Understanding*, vol. 57, no. 3, pp. 359-372, 1993.

N.R. Pal and S.K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277-1294, 1993.

L. Lam, S.-W. Lee, and C.Y. Suen, "Thinning methodologies: A comprehensive survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, pp. 869-885, 1992.

P. Duhamel and M. Vetterli, "Fast Fourier transforms: a tutorial review and a state of the art", *Signal Processing*, vol. 19, no. 4, 1990, 259-299.

S. Marshall, "Review of shape coding techniques", *Image and Vision Computing*, vol. 7, 1989, 281-294.

T.Y. Kong and A. Rosenfeld, "Digital topology: introduction and survey," *Computer Vision, Graphics, and Image Processing*, vol. 48, no. 3, pp. 357-393, 1989.

J. Kittler and J. Illingworth, "Relaxation labelling algorithms-a review", *Image and Vision Computing*, vol. 3, 1985, 206-216.

Vision

J.S. Suri, J.S., Kecheng Liu, S. Singh, S.N. Laxminarayan, Xiaolan Zeng and L. Reden, "Shape recovery algorithms using level sets in 2-D/3-D medical imagery: a state-of-the-art review", *IEEE Transactions on Information Technology in Biomedicine*, vol. 6, no. 1, 2002, 8-28.

G.N. Desouza and A.C. Kak, "Vision for mobile robot navigation: a survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, 2002, 237-267.

T.B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture", *Computer Vision and Image Understanding*, vol. 81, no. 3, 2001, 231-268.

J. Montagnat, H. Delingette and N. Ayache, "A review of deformable surfaces: topology, geometry and deformation", *Image and Vision Computing*, vol. 19, no. 14, 2001, 1023-1040.

Z.-L. Lu and G. Sperling, "Three-systems theory of human visual motion perception: review and update", *Journal of the Optical Society of America A - Optics and Image Science*, vol. 18, no. 9, 2001, 2331-2370.

Xuming Xu, Shi Li, Zhen Ye and Zeng Qi Sun, "A survey: RoboCup and the research", *Proceedings of the 3rd World Congress on Intelligent Control and Automation*, vol. 1, 2000, 207-211.

D.M. Gavrilu, "The visual analysis of human movement: a survey", *Computer Vision and Image Understanding*, vol. 73, no. 1, 1999, 82-98.

R. Zhang, P.-S. Tsai, J.E. Cryer and M. Shah, "Shape from shading: a survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, 1999, 690-706.

J.K. Aggarwal and Q. Cai, "Human motion analysis: a review", *Computer Vision and Image Understanding*, vol. 73, no. 3, 1999, 428-440.

J.K. Aggarwal, Q. Cai, W. Liao and B. Sabata, "Nonrigid motion analysis - articulated and elastic motion", *Computer Vision and Image Understanding*, vol. 70, no. 2, 1998, 142-156.

M. Bolduc and M.D. Levine, "A review of biologically motivated space-variant data reduction models for robotic vision", *Computer Vision and Image Understanding*, vol. 69, no. 2, 1998, 170-184.

J. Batlle, E. Mouaddib and J. Salvi, "Recent progress in coded structured light as a technique to solve the correspondence problem: a survey", *Pattern Recognition*, vol. 31, no. 7, 1998, 963-982.

V.I. Pavlovic, R. Sharma, and T.S. Huang, "Visual interpretation of hand gestures for human-computer interaction: a review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, 1997, 677-695.

D. Crevier and R. Lepage, "Knowledge-based image understanding systems - a survey", *Computer Vision and Image Understanding*, vol. 67, no. 2, 1997, 161-185.

K.J. Parker, L. Gao, R.M. Lerner and S.F. Levinson, "Techniques for elastic imaging: a review", *IEEE Engineering in Medicine and Biology Magazine*, vol. 15, no. 6, 1996, 52-59.

A. Fenster and D.B. Downey, "3-D ultrasound imaging: a review", *IEEE Engineering in Medicine and Biology Magazine*, vol. 15, no. 6, 1996, 41-51.

C. Cedras and M. Shah, "Motion based recognition: A survey", *Image and Vision Computing*, vol. 13, no. 2, 1995, 129-155.

S. Sarkar and K.L. Boyer, "Perceptual organization in computer vision - a review and a proposal for a classifactory structure", *IEEE Transactions on Systems Man and Cybernetics*, vol. 23, no. 2, 1993, 382-399.

I.J. Cox, "A review of statistical data association techniques for motion correspondence", *International Journal of Computer Vision*, vol. 10, 1993, 53-66.

R.J. Prokop and A.P. Reeves, "A Survey of Moment-Based Techniques for Unoccluded Object Representation and Recognition," *CVGIP - Graphical Models and Image Processing*, vol. 54, no. 5, pp. 438 - 460, 1992.

B. Sabata and J.K. Aggarwal, "Estimation of Motion from a Pair of Range Images: A Review," *CVGIP - Image Understanding*, vol. 54, no. 3, pp. 309 - 324, 1991.

P. Meer, D. Mintz, A. Rosenfeld, and D.Y. Kim, "Robust regression methods for computer vision: A review", *International Journal of Computer Vision*, vol. 6, 1991, 59-70.

R.M. Haralick and L.G. Shapiro, "Glossary of computer vision terms," *Pattern Recognition*, vol. 24, no. 1, pp. 69-93, 1991.

K.W. Bowyer and C.R. Dyer, "Aspect graphs: an introduction and survey of recent results", *International Journal of Imaging Systems and Technologies*, vol. 2, pp. 315-328, 1990.

U.R. Dhond and J.K. Aggarwal, "Structure from stereo-a review", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, 1989, 1489-1510.

M.J. Magee and J.K. Aggarwal, "Using multisensory images to derive the structure of three-dimensional objects -- a review", *Computer Vision, Graphics, and Image Processing*, vol. 32, 1985, 145-157.

Optical character recognition and text processing

R. Plamondon and S.N. Srihari, "Online and off-line handwriting recognition: a comprehensive survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, 2000, 63-84.

A. Amin, "Off-line Arabic character recognition: the state of the art", *Pattern Recognition*, vol. 31, no. 5, 1998, 517-530.

D. Doermann, "The retrieval of document images: a brief survey", *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, vol. 2, 1997, 945-949.

D. Trier, A.K. Jain, and T. Taxt, "Feature extraction methods for character recognition - a survey," *Pattern Recognition*, vol. 29, no. 4, pp. 641-662, Apr. 1996.

Y.Y. Tang, S.W. Lee, and C.Y. Suen, "Automatic document processing - a survey", *Pattern Recognition*, vol. 29, no. 12, 1996, 1931-1952.

B. Albadr and S.A. Mahmoud, "Survey and bibliography of Arabic optical text recognition," *Signal Processing*, vol. 41, no. 1, 1995, 49-77.

T.H. Hildebrandt and W. Liu, "Optical recognition of handwritten Chinese characters: advances since 1980," *Pattern Recognition*, vol. 26, no. 2, pp. 205-225, 1993.

S. Mori, C.Y. Suen, and K. Yamamoto, "Historical review of OCR research and development", *Proceedings of the IEEE*, vol. 80, 1992, 1029-1058.

C.C. Tappert, C.Y. Suen and T. Wakahara, "The state of the art in on-line handwriting recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 8, 1990, 787-808.

D.G. Elliman and I.T. Lancaster, "A review of segmentation and contextual analysis techniques for text recognition," *Pattern Recognition*, vol. 23, no. 3/4, pp. 337-346, 1990.

V.K. Govindan and A.P. Shivaprasad, "Character recognition -- A review," *Pattern Recognition*, vol. 23, no. 7, pp. 671-683, 1990.

F. Nouboud and R. Plamondon, "On-line recognition of handprinted characters: survey and beta tests," *Pattern Recognition*, vol. 23, no. 9, pp. 1031-1044, 1990.

R.H. Davis and J. Lyall, "Recognition of handwritten characters-a review", *Image and Vision Computing*, vol. 4, 1986, 208-218.

Other application areas

W.R. Anis Ibrahim and M.M. Morcos, "Artificial intelligence and advanced mathematical tools for power quality applications: a survey", *IEEE Transactions on Power Delivery*, vol. 17, no. 2, 2002, 668-673.

R. Gutierrez-Osuna, "Pattern analysis for machine olfaction: a review", *IEEE Sensors Journal*, vol. 2, no. 3, 2002, 189-202.

R. Jain, S. Antani and R. Kasturi, "A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video", *Pattern Recognition*, vol. 35, no. 4, 2002, 945-965.

W.G. Fenton, T.M. McGinnity and L.P. Maguire, "Fault diagnosis of electronic systems using intelligent techniques: a review", *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, vol. 31, no. 3, 2001, 269-281.

Y. Dote, and S.J. Ovaska, "Industrial applications of soft computing: a review", *Proceedings of the IEEE*, vol. 89, no. 9, 2001, 1243-1265.

H.S. Hippert, C.E. Pedreira and R.C. Souza, "Neural networks for short-term load forecasting: a review and evaluation", *IEEE Transactions on Power Systems*, vol. 16, no. 1, 2001, 44-55.

E. Hjelm, A.W.M. Smees and B. Kee Low, "Face Detection: a survey", *Computer Vision and Image Understanding*, vol. 83, no. 3, 2001, 236-274.

A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-based image retrieval at the end of the early years", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, 2000, 1349-1380.

M. Ibnkahla, "Applications of neural networks to digital communications - a survey, *Signal Processing*, vol. 80, no. 7, 2000, 1185-1215.

E.L. Hines, E. Llobet and J.W. Gardner, "Electronic noses: a review of signal processing techniques", *IEEE Proceedings on Circuits, Devices and Systems*, vol. 146, no. 6, 1999, 297-310.

H. Mayer, "Automatic Object Extraction from Aerial Imagery - A Survey Focusing on Buildings", *Computer Vision and Image Understanding*, vol. 74, no. 2, 1999, 138-149.

A.R. Jimenez, A.K. Jain, R. Ceres and J.L. Pons, "Automatic fruit recognition: a survey and new results using range/attenuation images", *Pattern Recognition*, vol. 32, no. 10, 1999, 1719-1736.

M. de Marsicoi, L. Cinque, S. Levialdi, "Indexing pictorial documents by their content: a survey of current techniques", *Image and Vision Computing*, vol. 15, no. 2, 1997, 119-141.

M. Moganti, F. Ercal, C.H. Dagli, and S. Tsunekawa, "Automatic PCB inspection algorithms: a survey", *Computer Vision and Image Understanding*, vol. 63, no. 2, 1996, 287-313.

R. Romanroldan, P. Bernalgalvan, and J.L. Oliver, "Application of information theory to DNA sequence analysis: a review", *Pattern Recognition*, vol. 29, no. 7, 1996, 1187-1194.

G. Kerr, "A review of fully interactive video on demand", *Signal Processing: Image Communication*, vol. 8, no. 3, 1996, 173-190.

T.S. Newman and A.K. Jain, "A survey of automated visual inspection", *Computer Vision and Image Understanding*, vol. 61, no. 2, 1995, 231-262.

D. Valentin, H. Abdi, A.J. Otoole, and G.W. Cottrell, "Connectionist Models of Face Processing - A Survey," *Pattern Recognition*, vol. 27, no. 9, pp. 1209-1230, 1994.

G. Liu, "Applied stereology in materials science and engineering - Invited review," *Journal of Microscopy - Oxford*, vol. 171, no. 1, pp. 57-68, 1993.

A. Samal and P.A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: a survey," *Pattern Recognition*, vol. 25, no. 1, pp. 65-77, 1992.

M.W. Roth, "Survey of Neural Network Technology for Automatic Target Recognition," *IEEE Transactions on Neural Networks*, vol. 1, no. 1, pp. 28-43, 1990.

Richard P. Lippmann, "Review of neural networks for speech recognition," *Neural Computation*, vol. 1, no. 1, pp. 1-38, 1989.

A.1. L'Intelligenza Artificiale: la storia e le idee

È difficile dare una definizione dell'IA e dei suoi obiettivi che sia unanimemente condivisa dai ricercatori. L'origine della difficoltà sta anche nel fatto che da sempre l'IA si è presentata sotto un duplice profilo: quello di disciplina ingegneristica, il cui obiettivo è di costruire macchine in grado di assistere l'uomo, e magari di competere con esso, in compiti soprattutto intellettuali, e quello di disciplina psicologica, il cui obiettivo è di costruire macchine le quali, riproducendo da vicino le caratteristiche essenziali dell'attività cognitiva umana, gettino una nuova luce su alcuni tradizionali enigmi della mente, ad esempio sul cosiddetto problema mente-corpo. Forse la base programmatica dell'IA più comunemente accettata è ancora quella utilizzata nella presentazione del seminario organizzato da John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon nel giugno del 1956 negli Stati Uniti, a Dartmouth (New Hampshire), nella quale si legge: *In linea di principio si può descrivere ogni aspetto dell'apprendimento e dell'intelligenza con una precisione tale da permetterne la simulazione con macchine appositamente costruite. Si cercherà di costruire macchine in grado di usare il linguaggio, di formare astrazioni e concetti, di migliorare se stesse e risolvere problemi che sono ancora di esclusiva pertinenza degli esseri umani.*

Nel corso di quello storico seminario si gettarono le basi dell'IA, individuando alcune aree di ricerca rimaste classiche e presentando i primi programmi per calcolatore cosiddetti "intelligenti". Le macchine alle quali pensavano i pionieri dell'IA sono i calcolatori digitali. È possibile circoscrivere con una certa precisione alcune proprietà fondamentali di queste macchine, proprietà che ci aiutano anche a comprendere perché di esse si è parlato come di "macchine intelligenti". Ricorderemo varie caratteristiche dei calcolatori fisicamente realizzati e di loro modelli matematici, facendo riferimento a risultati cruciali della teoria della calcolabilità, le cui basi vennero gettate da alcuni logici nell'arco del quinquennio che va dal 1931 al 1936 - e cioè prima che i moderni calcolatori digitali venissero effettivamente costruiti. I pionieri della nuova disciplina videro nel calcolatore digitale uno strumento con capacità di elaborazione simbolica ineguagliate da qualsiasi macchina sviluppata nella storia dell'umanità e per questo motivo adatta al confronto con alcuni aspetti tra i più peculiari dell'intelligenza umana.

Nei giorni di Dartmouth, pronto per "girare" su calcolatore era un programma ideato da Newell, Shaw e Simon, l'ormai mitico Logic Theorist, che dimostrava alcuni teoremi del calcolo proposizionale. Ma bisogna ricordare almeno un altro programma la cui sperimentazione era cominciata qualche tempo prima, quello di Arthur Samuel per il gioco della dama, in grado di migliorare le proprie prestazioni al punto di arrivare a battere ottimi giocatori. Altri programmi sarebbero presto seguiti: dalla Geometry Machine di Herbert Gelertner e Rochester ai primi programmi per gli scacchi, stimolati da alcune precedenti intuizioni di Shannon (Somenzi, Cordeschi, 1994). I protagonisti della prima IA si concentrarono su problemi relativi ad ambiti ben delimitati, per risolvere i quali bastano regole esplicite per l'elaborazione simbolica e poca conoscenza specializzata. In questa scelta iniziale sono confluiti sia motivi pratici, come la modesta potenza di calcolo dei calcolatori dell'epoca, sia motivi più teorici, come la diffusa concezione che l'elaborazione simbolica sia la vera marca dell'intelligenza. Ma la descrizione degli intenti dell'IA sopra ricordata si rivela in questo caso opportunamente generica, poiché non assegna esplicitamente all'elaborazione simbolica un ruolo privilegiato. Programmi di ricerca ispirati ad altre concezioni dell'intelligenza sono stati portati avanti nell'ambito dell'IA, spostando l'accento dall'elaborazione simbolica alla conoscenza tacita, alle abilità senso-motorie, alle capacità di adattamento all'ambiente naturale o alle interazioni sociali con altri agenti naturali o artificiali. Programmi di ricerca così diversi tra loro sono emersi anche in reazione agli ostacoli che questa disciplina ha incontrato lungo un percorso non lineare, per quanto costellato da notevoli successi. La diffusione capillare di tecniche e sistemi dell'IA in molti prodotti della tecnologia avanzata testimonia il successo dell'IA intesa come disciplina ingegneristica. Ma quale bilancio possiamo trarre a proposito della costruzione di macchine che riproducano caratteristiche essenziali dell'attività cognitiva umana? Cosa aggiunge alla nostra conoscenza del mentale un sistema dell'IA che simula con successo un essere umano nello svolgimento di un certo compito cognitivo? E che posto occupa l'IA, così intesa, nell'ambito di ricerche interdisciplinari sul sistema cervello-mente che coinvolgono anche le neuroscienze e la psicologia? Vedremo come le tesi del funzionalismo, della realizzabilità di sistemi intelligenti con sostrati materiali differenti, dell'indipendenza di vari livelli di analisi e spiegazione del mentale emergono nelle riflessioni dei ricercatori dell'IA prima di ricevere una formulazione canonica da parte dei filosofi della mente. Infine, toccheremo brevemente i problemi delle emozioni e dei cosiddetti qualia e il problema "difficile" della coscienza intesa come esperienza soggettiva, che resta un enigma per l'IA così come per ogni altra impostazione scientifica allo studio della mente.

Macchine astratte e calcolatori potenzialmente universali

I calcolatori digitali sono dunque le macchine prescelte per realizzare i sistemi dell'IA. Nella forma di calcolatori "generali" (general purpose), essi hanno la fondamentale caratteristica di essere macchine simboliche, che non eseguono solo un repertorio limitato di calcoli numerici, come le ordinarie macchine calcolatrici. Con i calcolatori general purpose si comincia a disporre per la prima volta di una macchina capace di manipolare strutture di simboli che il programmatore fa corrispondere in modo naturale alle entità più diverse: parole di una lingua naturale, espressioni matematiche, posizioni del gioco degli scacchi, oggetti da riconoscere e classificare. Questa capacità di elaborazione simbolica, insieme alla presenza dell'istruzione di "salto condizionato" nei programmi, sollecitò da subito a parlare di "macchine intelligenti". L'istruzione di salto condizionato consente di cambiare l'ordine di esecuzione delle istruzioni: se una data condizione è soddisfatta, allora vengono effettuate operazioni specificate da una certa parte del programma, altrimenti ne vengono eseguite altre, specificate da una diversa parte del programma. Il salto condizionato conferisce a un programma una certa capacità discriminativa, che sfocia nella selezione della sequenza di istruzioni da eseguire. Nell'EDSAC le due caratteristiche che abbiamo ricordato erano pienamente realizzate. L'EDSAC, realizzato nel 1949, era il primo grande calcolatore con programma memorizzato: nella sua memoria interna erano depositati non solo i dati, ma anche le istruzioni per manipolarli, ovvero il programma, che poteva essere modificato non meno dei dati mediante operazioni algoritmiche. Era questa l'architettura dei calcolatori che John von Neumann aveva sintetizzato nel celebre First Draft del 1945 e che sarebbe rimasta sostanzialmente immutata negli anni a venire (un comune calcolatore da tavolo è una macchina di questo tipo).

I primi calcolatori (1941-51)

- 1941 : Z3 di K. Zuse. Il primo calcolatore programmabile digitale e general purpose effettivamente costruito. Era elettromeccanico, con una memoria composta di 2600 relè telefonici. Poteva convertire i decimali in numerazione binaria e viceversa. Utilizzava un nastro magnetico perforabile per il programma in ingresso.
- 1943 : Primo dei calcolatori COLOSSUS di T. H. Flowers, W. W. Chandler e altri. Era un grande calcolatore elettronico, con una memoria composta di 1500 valvole. Usava cinque processori paralleli e poteva leggere un nastro perforato alla velocità di 5000 caratteri al secondo.
- 1944 : Harvard Mark I di H. Aiken. Era elettromeccanico e usava la numerazione binaria. Era specializzato in problemi di tipo matematico, ma aveva già il programma memorizzato.
- 1946 : ENIAC di J. P. Eckert e J. W. Mauchly. Il primo grande e veloce calcolatore elettronico, general purpose e programmabile. Era composto di più di 18000 valvole e utilizzava un sistema di calcolo parallelo.
- 1949 : EDSAC (Electronic Delay Storage Automatic Computer) di M. Wilkes, W. Renwick, D. J. Wheeler e collaboratori. Il primo calcolatore elettronico general purpose con programma memorizzato. MADM (Manchester Automatic Digital Machine) di F. Williams, T. Kilburn e M. H. A. Newman (con la collaborazione di A. M. Turing per gli aspetti relativi alla programmazione). Da questa macchina prototipo fu costruito nel 1951, in collaborazione con la Ferranti Ltd., il Ferranti Mark I. BINAC (Binary Automatic Computer) di J. P. Eckert e J. W. Mauchly. Era di tipo general purpose, il primo calcolatore costruito da privati fuori da centri di ricerca.
- 1950 : ACE (Automatic Computing Engine) di G. G. Alway, D. Davies, J. H. Wilkinson e M. Woodger (progetto di A. M. Turing del 1945). Costruito presso il National Physical Laboratory di Teddington, era general purpose e usava la numerazione binaria.
- 1951 : EDVAC (Electronic Discrete Variable Automatic Computer) di R. L. Snyder, S. E. Gluck e W. H. Boghosian. Il progetto iniziale, tra gli altri di J. von Neumann, lo descriveva come un calcolatore a programma memorizzato. Era composto di 3600 valvole e usava schede perforate. UNIVAC (Universal Automatic Computer) di J. P. Eckert e J. W. Mauchly. Il primo grande calcolatore elettronico prodotto per scopi commerciali.

I programmi che determinano i processi di elaborazione dei calcolatori digitali sono procedimenti algoritmici specificati in un qualche linguaggio di programmazione dato. La nozione generale di programma dipende dunque da quella di procedimento algoritmico, che è del tutto intuitiva e dai contorni sfumati - per quanto sia sufficientemente chiara da consentire a tutti di riconoscere e di eseguire vari procedimenti algoritmici, per calcolare almeno funzioni molto semplici, come la somma o il prodotto di numeri naturali.

Alcune proprietà fondamentali dei procedimenti algoritmici sono state tuttavia individuate con precisione, soprattutto grazie alle ricerche di vari logici, come Alonzo Church, Kurt Gödel, Stephen Kleene, Emil Post e Alan Turing. In particolare, attraverso la cosiddetta tesi di Church-Turing, enunciata indipendentemente da Church e Turing tra il 1935 e il 1936, è stata fornita una caratterizzazione precisa e generalmente ritenuta soddisfacente delle funzioni calcolabili mediante procedimenti algoritmici. Nella versione direttamente collegata al lavoro di Turing (1937), la tesi asserisce che ogni funzione calcolabile mediante un procedimento algoritmico è calcolabile da una macchina di Turing. Dunque, per la tesi di Church-Turing la classe delle funzioni calcolabili mediante procedimenti algoritmici è inclusa nella classe delle funzioni calcolabili dalle macchine di Turing. Quest'ultima classe, al contrario della prima, ha una definizione precisa, che è basata a sua volta sulla definizione di macchina di Turing.

Una macchina di Turing è un calcolatore idealizzato, in quanto si suppone che essa non sia soggetta ad alcune limitazioni proprie dei calcolatori fisicamente realizzati: non si impone un limite alle sue capacità di memoria (la memoria di una macchina di Turing è costituita da un nastro indefinitamente espandibile, uniformemente diviso in caselle, ciascuna delle quali può contenere al più un simbolo da un alfabeto finito e prefissato); non si impone nemmeno un limite alla durata dei processi di calcolo e si assume che la macchina funzioni sempre perfettamente. Le operazioni eseguibili da una macchina di Turing sono estremamente elementari: dotata di una testina di lettura per osservare una casella per volta, una macchina di Turing che si trova in una determinata configurazione interna può stampare un simbolo sulla casella osservata o cancellare il simbolo già stampato su quest'ultima, spostare la testina di lettura di una casella a sinistra o a destra della casella osservata e cambiare la propria configurazione interna. Quali di queste operazioni vengono eseguite a ogni dato istante dipende esclusivamente dal contenuto della casella osservata e dalla configurazione interna della macchina.

Gli aspetti fondamentali di questa descrizione informale e intuitivamente visualizzabile di una macchina di Turing possono essere espressi in un linguaggio più astratto e rigoroso. Si consideri la lista di simboli s_1, s_2, s_3, \dots ; q_1, q_2, q_3, \dots ; D, S . Un'espressione è una successione finita di simboli presi da questa lista. Un'istruzione è una quintupla, cioè un'espressione della forma "q_is_jsk₁M" dove M sta per la lettera D o per la lettera S. Intuitivamente, un'istruzione si interpreta in questo modo: se la configurazione interna è q_i e il simbolo osservato è s_j , allora la macchina scriverà al suo posto il simbolo s_k entrando nella configurazione interna q_1 e spostandosi di una casella, a destra o a sinistra, a seconda che M sia D oppure S. Si conviene generalmente che il simbolo s_1 indichi la casella vuota; pertanto rimpiazzare s_j con s_1 equivale all'azione di cancellare s_j .

Definizione: Una macchina di Turing è un insieme finito (ma non vuoto) di quintuple che non contiene nessuna coppia di quintuple che coincidano nei loro primi due simboli. Questa definizione impone cruciali condizioni di finitezza (relative, per esempio, al numero di istruzioni, all'alfabeto di simboli sul quale opera la macchina e alle sue possibili configurazioni interne) e di determinatezza (ad ogni istante dato una macchina di Turing potrà applicare al più un'istruzione). Essa identifica una macchina di Turing con un programma formato da istruzioni scritte in una forma standard. È anche possibile dare una descrizione rigorosa di quali siano, ad ogni passo di esecuzione di un tale programma, la casella osservata, i contenuti del nastro e la configurazione interna della macchina (per una trattazione esauriente cfr. Davis, 1982, pp. 5-8).

Alcune macchine di Turing sono dette universali, perché in grado di simulare fedelmente il comportamento di una qualsiasi macchina di Turing dato un qualsiasi ingresso. Poiché una macchina universale di Turing può calcolare tutte le funzioni calcolabili da ogni macchina di Turing, dalla tesi di Church-Turing segue che una tale macchina universale può calcolare ogni funzione calcolabile da un calcolatore digitale. Tra i calcolatori general purpose fisicamente realizzati, anche i calcolatori da tavolo e i portatili più comuni hanno una stretta relazione con le macchine universali di Turing. Facendo astrazione da possibili guasti e dalle limitazioni relative alla memoria e alla durata dei calcoli, i nostri calcolatori da tavolo hanno la stessa potenza di calcolo di una macchina universale di Turing. Per questo motivo siamo autorizzati a chiamarli calcolatori potenzialmente universali. Pertanto si preserva la massima generalità rispetto alla classe di funzioni calcolabili se si sceglie, per simulare dei comportamenti intelligenti, una macchina potenzialmente universale. Questo punto è stato fortemente sottolineato agli albori dell'IA dallo stesso

Turing (1950) e rielaborato con originalità da Newell e Simon attraverso la nozione di sistema fisico di simboli. Il lavoro che portò alla tesi di Church-Turing fu principalmente motivato dal problema di generalizzare i teoremi di incompletezza, scoperti da Gödel nel 1931. Nella forma generalizzata ottenuta facendo ricorso alla tesi di Church-Turing, il primo teorema di Gödel implica che non vi è nessun procedimento algoritmico che consenta di generare tutti gli enunciati veri dell'aritmetica elementare senza generare al contempo nessun enunciato falso. In altre parole, se un procedimento algoritmico consente di operare in modo corretto nell'ambito dell'aritmetica elementare (cioè non permette di generare enunciati falsi), allora esso risulta essere incompleto (cioè non permette di generare tutti gli enunciati aritmetici veri). Il problema di generare tutte e sole le proposizioni vere dell'aritmetica elementare non è risolvibile, nemmeno in linea di principio, mediante un procedimento algoritmico. Si è ampiamente discusso se gli esseri umani siano soggetti ad analoghe limitazioni e, in caso di risposta negativa, se ciò ponga dei limiti alle ambizioni dell'IA.

Elaborazione simbolica e intelligenza

La partizione in problemi risolvibili o non risolvibili mediante un procedimento algoritmico è ancora troppo grossolana per gli scopi dell'IA. Nel caso di molti problemi risolvibili "in linea di principio", accade che le risorse di calcolo da impiegare per arrivare alla soluzione desiderata eccedano la potenza di un qualsiasi calcolatore già realizzato o perfino fisicamente realizzabile. In questi casi, la locuzione "in linea di principio" indica una mera possibilità matematica, che non implica possibilità fisica e tanto meno la concreta realizzabilità di sistemi in grado di risolvere tali problemi. La teoria della complessità algoritmica o computazionale studia le risorse di calcolo necessarie per la soluzione dei problemi (per un'introduzione pertinente ai temi ai quali qui accenniamo cfr. Burattini, 1993).

Si sono rivelate proibitive, in base alle stime della teoria della complessità, le risorse di calcolo necessarie per risolvere problemi che hanno attratto da subito l'attenzione dei pionieri dell'IA, come quello di determinare una strategia ottimale in una partita a scacchi. È stato sostenuto che questi risultati pongono un ostacolo teorico alle ambizioni dell'IA. Ma i ricercatori di questa disciplina hanno adottato un punto di vista molto diverso da quello della teoria della complessità, che ha permesso loro di aggirare il presunto ostacolo in molti casi significativi. I primi programmi dell'IA, come abbiamo ricordato, elaborano strutture di simboli e, soprattutto attraverso sequenze di applicazioni di regole di salto condizionato, simulano almeno i prodromi di una capacità ritenuta peculiare dell'intelligenza umana: quella di scegliere, davanti a un problema che dà luogo all'esplosione combinatoria delle mosse lecite, solo alcune sequenze di mosse che potrebbero portare alla soluzione.

Classico è il caso del giocatore di scacchi, che all'inizio della partita ha di fronte a sé un numero ultra-astronomico di possibili mosse alternative, che Shannon calcolò nell'ordine di 10^{120} . L'intelligenza fu primariamente identificata con questa capacità selettiva della mente umana, del resto ben documentata dalla psicologia dei processi del pensiero. Fu proprio con quest'ultima che l'IA delle origini venne a confrontarsi, sia nelle vesti della Information Processing Psychology, o "psicologia dei processi dell'informazione", che aspirava a simulare i processi cognitivi umani in modo psicologicamente realistico, sia nella costruzione di programmi per calcolatore che, senza essere psicologicamente realistici, avessero comunque prestazioni efficienti. Il successo dei primi programmi alimentò la convinzione che compito principale dell'IA fosse lo studio delle strategie di soluzione di problemi efficacemente selettive, o "euristiche".

La programmazione euristica, il settore di ricerca nel quale l'IA raccolse i suoi primi successi, ha risposto alle sfide della complessità algoritmica tenendo conto delle strategie effettivamente adottate dagli esseri umani. Questi ultimi, dovendo fornire risposte adeguate ai problemi in tempi ragionevoli, per lo più evitano di condurre una ricerca "cieca" o basata sulla forza bruta di soluzioni ottimali a un problema dato, limitandosi a esplorare solo una parte dei percorsi che, in base alle informazioni in loro possesso, potrebbero portare a una soluzione accettabile del problema dato. Ci si comporta in questo modo in tante situazioni della vita di tutti i giorni (quando, poniamo, si sceglie una scuola per i propri figli o una località per le vacanze), accontentandosi di individuare una soluzione che soddisfi alcuni requisiti irrinunciabili. Ma così facendo si corre il rischio che tutte le soluzioni soddisfacenti al problema dato non si trovino lungo i percorsi prescelti; per questo motivo, in IA si pone grande cura nello sviluppo di metodi per individuare i percorsi risolutivi più promettenti. Gli algoritmi euristici dell'IA non sono ispirati solo al comportamento cognitivo degli esseri umani (come la strategia "mezzi-fine" introdotta all'origine dell'IA da Allen Newell e Herbert Simon).

L'evoluzionismo darwiniano, con le nozioni di mutazione e selezione, ha ispirato gli algoritmi evolutivi, applicati sia in IA sia nella cosiddetta Vita Artificiale (Langton, 1995). Vi sono inoltre euristiche che prendono lo spunto da comportamenti individuali o collettivi osservati nel mondo animale, come le strategie cooperative di ricerca e segnalazione di fonti di cibo messe in opera da colonie di formiche (Dorigo, Gambardella, 1997). La programmazione euristica ha continuato comunque a costituire un settore di ricerca importante dell'IA. Tipicamente fondati sulla programmazione euristica sono i sistemi basati sulla conoscenza, e tra questi i cosiddetti sistemi esperti, che generalmente richiedono un supporto hardware più potente di quello dei primi calcolatori. Un sistema esperto risolve problemi in un ambito specialistico (ad esempio in un qualche settore diagnostico) con prestazioni, almeno nei casi riusciti, comparabili a quelle di un esperto umano del settore. Una metodologia comunemente impiegata nella costruzione di un sistema esperto prevede anzitutto la registrazione di conoscenze e di tecniche di ragionamento euristico dell'esperto umano. Tali conoscenze sono registrate, sia pure in forme drasticamente semplificate, in strutture dati chiamate basi di conoscenza, che generalmente assumono vaste dimensioni.

Sulla base di conoscenza opera un motore inferenziale, che risolve i problemi posti anche simulando tecniche di ragionamento euristico dell'esperto umano. Nei grandi sistemi basati sulla conoscenza, al di là di differenze, anche marcate, negli stili di rappresentazione ed elaborazione della conoscenza, si riconosce uno dei prodotti principali di quella che viene comunemente chiamata "IA simbolica". Un bilancio equilibrato del lavoro svolto in questo ambito impone di ricordare, ma senza nessuna pretesa di esaustività, alcuni problemi che le metodologie soggiacenti ai sistemi basati sulla conoscenza non permettono di affrontare adeguatamente.

1. Conoscenza tacita. Varie forme di comportamento intelligente sono guidate da ciò che chiamiamo comunemente il "saper fare", cioè da un tipo di abilità che appare difficile e spesso innaturale descrivere attraverso conoscenze dichiarative e regole esplicite di manipolazione simbolica.

2. Azione in tempo reale. Gli algoritmi euristici dell'IA simbolica, anche se soddisfacenti rispetto a ricerche esaustive o a quelle con un grado controllato di approssimazione all'ottimalità, non sempre riescono a generare risposte adeguate in tempi utili. Ostacoli del genere sono stati incontrati, per esempio, nella progettazione di sistemi esperti per la diagnosi di guasti in situazioni critiche o di sistemi per la pianificazione delle azioni di un robot in ambienti poco prevedibili.

3. Robustezza. I sistemi che operano in base a rappresentazioni esplicite della conoscenza si rivelano fragili, spesso incapaci di operare in situazioni diverse da quelle previste dai progettisti. Piccole discrepanze tra lo stato percepito del mondo e le condizioni previste per l'esecuzione di un procedimento algoritmico talvolta portano a un rapido degrado delle prestazioni. Ciò non si osserva negli esseri umani e in altre specie animali, le cui prestazioni sono caratterizzate da una buona tolleranza al rumore e alle variazioni ambientali.

Negli ultimi venti anni, attraverso il connessionismo, la nuova robotica e la cognizione situata, la tematizzazione di queste e altre difficoltà ha permesso di delineare programmi di ricerca che si collocano all'interno di quella che viene talvolta chiamata "nuova IA". Ci limitiamo a indicarne qui alcune caratteristiche generali. Il connessionismo attuale trae ispirazione remota da schematici modelli funzionali delle cellule nervose, introdotti nella scia del lavoro pionieristico di Warren McCulloch e Walter Pitts del 1943. Successivamente, attraverso la mediazione delle celebri ricerche sui meccanismi dell'apprendimento di Donald Hebb del 1949, lo studio dei processi di apprendimento e classificazione in reti con neuroni a la McCulloch e Pitts si è incontrato, nel Percettrone di Rosenblatt, con la tradizione del connessionismo psicologico e neurologico risalente a Ivan Pavlov, Edward Thorndike e Clark Hull (alcuni testi di McCulloch e Pitts, di Rosenblatt e di Hebb sono raccolti in Anderson, Rosenfeld, 1988). Dopo un periodo di relativa stasi - attribuita da molti sia all'impossibilità di sintetizzare alcune semplici funzioni booleane con Percettroni formati da un solo "strato" di neuroni sia alla mancanza di validi algoritmi di apprendimento per reti a più strati - lo studio delle reti neurali è ripreso con grande vigore intorno al 1985, grazie all'introduzione di nuovi modelli di neuroni e di nuovi algoritmi di apprendimento per reti multistrato (come quello di retropropagazione), ma anche grazie alla disponibilità di calcolatori più potenti per simulare il comportamento di reti formate da un elevato numero di unità. Il connessionismo ha affrontato, non di rado con maggior successo della prima IA, quelli che abbiamo chiamato problemi di conoscenza tacita e di robustezza, ad esempio in relazione all'elaborazione di segnali sensoriali. La nuova robotica e la cognizione situata hanno invece affrontato i problemi dell'azione in tempo reale. Contrastando una tendenza, manifestatasi all'interno della prima IA, a scindere nettamente tra sistemi percettivi, sistemi motori e sistemi cognitivi in senso stretto, la cognizione situata ha variamente tematizzato l'interazione tra mente, corpo e ambiente circostante (cfr. Clark, 1997). Il modello ideale di un sistema di successo dell'IA è diventato così un robot immerso in

un ambiente non predeterminato dal progettista e in grado di mostrare comportamenti adattivi in tempo reale, piuttosto che un calcolatore in grado di sostenere, poniamo, un'educata conversazione salottiera, come nel gioco dell'imitazione di Turing (1950), o di battere alla maniera di Deep Blue un grande maestro di scacchi. Inizialmente, la proposizione di questo nuovo modello ideale è sfociata in posizioni estreme, che hanno proposto l'eliminazione di ogni forma di rappresentazione ed elaborazione simbolica dal repertorio degli strumenti dell'IA. In alcuni scritti che si collocano alle origini della nuova robotica, Rodney Brooks ha assunto una posizione non lontana da un atteggiamento eliminativista "forte" nei confronti delle rappresentazioni. Le concezioni di Brooks hanno guidato lo sviluppo di architetture innovative per il controllo dell'azione in robot che si sono mostrati capaci di agire in tempo reale e in ambienti non completamente strutturati, senza fare ricorso a sistemi di controllo centralizzato.

Brooks ha anche riproposto importanti questioni epistemologiche e ontologiche relative all'analisi e alla simulazione dei sistemi intelligenti. Alle carenze nel controllo dell'azione in tempo reale si aggiunge, tra i motivi di insoddisfazione verso le rappresentazioni simboliche della prima IA, il fatto di non riuscire a definire in modo soddisfacente che cosa veramente siano le rappresentazioni (ad esempio quali siano le loro condizioni di identità) o quale sia il loro ruolo nei sistemi intelligenti, naturali e artificiali. Sebbene queste difficoltà siano innegabili, è veramente arduo rinunciare alle rappresentazioni nella costruzione e nella spiegazione del comportamento di sistemi intelligenti, come indicano gli stessi ostacoli incontrati dal programma di ricerca di Brooks. Per questa ragione, il richiamo a non trascurare l'interazione tra corpo, cognizione e mondo in altri casi è sfociato in proposte anche molto diverse di integrazione tra gli strumenti della prima IA e quelli dell'IA "nuova".

Come risultato si è avuta la progettazione di sistemi con connessioni retroattive più profonde fra i processi percettivo-motori e i processi di rappresentazione ed elaborazione simbolica (cfr. ad esempio Carlucci Aiello, Nardi, Pirri, 2001). L'assenza di modelli interni del mondo e di elaborazioni simboliche rende infatti il comportamento dei sistemi à la Brooks troppo dipendente da stimoli sensoriali locali. Il punto di vista originario di Brooks si espone in effetti a obiezioni simili a quelle già sollevate sia contro impostazioni comportamentiste allo studio dell'intelligenza, basate esclusivamente sulle correlazioni fra stimoli e risposte, sia contro i primi modelli cibernetici dell'azione rivolta a uno scopo, nei quali lo scopo veniva identificato con un oggetto fisico presente nel raggio di azione dei sensori del sistema modellato (Somenzi, Cordeschi, 1994, pp. 23-7). È inoltre difficile immaginare come gli austeri strumenti consentiti da Brooks possano fornire una base adeguata per simulare capacità di ragionamento che troviamo anche nei sistemi della prima IA, oppure per affrontare le problematiche dell'IA distribuita e dei sistemi multiagente. Un sistema intelligente immerso in un ambiente nel quale operano altri agenti deve essere infatti capace di comunicare con altri agenti, formarsi aspettative sui risultati di possibili interazioni, ragionare a proposito degli obiettivi e delle capacità degli altri agenti.

Proprio la dimensione sociale dell'intelligenza è diventata in anni recenti uno dei settori principali di indagine dell'IA, anche perché, oltre a Internet, le interfacce uomo-macchina e sistemi diversi di controllo negli elettrodomestici pullulano di piccoli agenti intelligenti specializzati, i quali devono interagire con gli utenti umani o almeno, se "invisibili" a questi ultimi, con altri agenti artificiali. I vantaggi largamente complementari mostrati dai sistemi basati sulla conoscenza e da quelli connessionisti o da quelli ispirati alle tesi della cognizione situata vengono spesso sfruttati nella progettazione di agenti "ibridi" in IA. Questa spregiudicatezza metodologica è legittima e auspicabile in una visione ingegneristica dell'IA. Ma per l'IA più strettamente collegata alla scienza cognitiva si pone il vincolo ulteriore e più severo di una maggiore aderenza ai processi cognitivi degli esseri umani o di altre specie animali. È opportuno notare a questo riguardo come la nozione di agente, attorno alla quale si è formato un vocabolario condiviso da ricercatori operanti nei vari programmi di ricerca che abbiamo menzionato, abbia un forte impatto metaforico, mentre richiederebbe un confronto più puntuale con quelli che consideriamo esempi genuini di agenti cognitivi. Nel diffuso impiego metaforico di un vocabolario cognitivista bisogna indicare uno dei vizi ricorrenti dell'IA, che ha suscitato fin dalle origini giuste critiche di vuoto ideologismo e una serie di polemiche verbali, a detrimento di una più equilibrata valutazione dei risultati sia ingegneristici sia teorici ottenuti da questa disciplina nel corso della sua breve storia.

Quali sono, dunque, le metodologie che potrebbero essere adottate per condurre un confronto sistematico soddisfacente tra i sistemi algoritmici dell'IA e altri agenti cognitivi?

Modelli algoritmici nello studio della mente

Turing (1950) ha discusso la possibilità di controllare, attraverso ciò che egli chiamava «gioco dell'imitazione», se un calcolatore opportunamente programmato potesse sostenere una conversazione nei modi tipici di un essere umano. La simulazione di comportamenti intelligenti, seppure in forme diverse da quelle contemplate nel gioco dell'imitazione (oggi noto anche come test di Turing), è rimasta un obiettivo centrale dell'IA. Ma per quanto riguarda l'IA come "nuova scienza della mente", la tesi che la simulazione, purché funzioni, possa essere conseguita sulla base di qualunque mezzo algoritmico ha dato luogo a diversi e gravi equivoci. La storia dell'IA ha mostrato che si possono simulare con successo, ad esempio, varie forme di comportamento verbale senza gettare alcuna luce sui meccanismi che presiedono alla generazione o alla comprensione del linguaggio naturale. Esempi noti e spettacolari sono i programmi vincitori del premio Loebner, o ELIZA di Joseph Weizenbaum, un programma che dialogando con un utente umano «scimmiettava», per dirla con Schank (1984), la comprensione della lingua inglese. Viceversa, la costruzione di programmi simulativi che siano rilevanti sul piano della spiegazione di certe capacità cognitive è per certi aspetti assimilabile al normale processo di costruzione e revisione dei modelli scientifici.

Per comprendere anche sviluppi più recenti in questa direzione, è bene partire da un breve esame della metodologia messa a punto da Newell, Shaw e Simon (cfr. Cordeschi, 1984, per maggiori dettagli). Essi furono infatti i primi a porsi esplicitamente il problema della simulazione come strategia di costruzione di modelli del mentale. La psicologia dell'elaborazione dell'informazione (PPI d'ora in avanti) aveva appunto questo obiettivo: costruire modelli algoritmici, cioè programmi per calcolatore che simulassero i processi cognitivi umani in modo psicologicamente realistico. Non meno dell'IA delle origini, la PPI era interessata in primo luogo alla questione del controllo. Una routine di controllo che selezionava le regole in funzione dell'obiettivo venne sperimentata da Newell, Shaw e Simon nella soluzione di problemi di logica proposizionale da parte del Logic Theorist.

Si trattava di un procedimento euristico "guidato" dall'obiettivo, diventato presto noto come analisi mezzi-fine: invece di applicare sistematicamente tutte le regole della logica proposizionale, la routine, attraverso il confronto tra la formula che rappresentava lo stato iniziale del problema e quella che rappresentava lo stato finale, selezionava solo quelle regole che permettevano di eliminare le differenze tra le due formule. La soluzione del problema consisteva nel generare una successione di formule progressivamente più simili a quella che rappresentava lo stato finale, coincidente con l'ultima formula della successione. Questa forma di controllo è stata ampiamente sperimentata in numerosi programmi di IA dedicati a compiti diversi. Nella PPI essa divenne inizialmente il fulcro di una metodologia simulativa che, sviluppata in particolare da Newell e Simon, è entrata nella storia delle scienze della mente del Novecento.

Essi pensarono di registrare il protocollo verbale di un soggetto che risolveva un problema riferendo ad "alta voce" i procedimenti selettivi o euristici che usava e di implementare questi ultimi in un programma. Il confronto tra il protocollo verbale e la traccia del programma avrebbe dovuto mostrare se e fino a che punto la simulazione aveva avuto successo, cioè se e fino a che punto processi umani e processi algoritmici di soluzione di problemi erano gli stessi, almeno sotto il profilo dell'elaborazione dell'informazione. Un aspetto importante di questa metodologia simulativa era che il programma doveva tener conto dei limiti effettivi del solutore di problemi umano, limiti di memoria, di velocità di elaborazione dei dati e così via.

Il programma doveva cioè soddisfare alcune restrizioni se voleva essere una simulazione psicologicamente realistica del comportamento del solutore di problemi umano. Non bastava che la prestazione della macchina fosse la stessa dell'essere umano, che la macchina, per così dire, si sostituisse a quest'ultimo nel dare la risposta giusta (la soluzione del problema). Il confronto protocollo-traccia doveva garantire che la simulazione si estendesse ai processi di soluzione del problema. Era questa condizione che distinse subito l'impostazione di Newell, Shaw e Simon, e anche l'uso stesso del termine "euristica" fatto da questi ultimi, rispetto a varianti più ingegneristiche dell'IA. Dunque non ogni programma dell'IA poteva essere considerato una simulazione interessante sotto il profilo della spiegazione psicologica, in breve un "modello". Il modello prendeva la forma di un programma di un tipo ben definito, che fu anche chiamato «microteoria» da Newell e Simon.

Una microteoria riguardava il comportamento di un singolo solutore umano di problemi, mentre la teoria generale dell'elaborazione, umana dell'informazione era un corpus di generalizzazioni qualitative (relative alle caratteristiche della memoria a breve e a lungo termine di tutti i

solutori umani di problemi, ai loro tempi di reazione ecc.), non più esprimibili sotto forma di programmi, ma ricavabili dallo studio dei singoli programmi simulativi o microteorie (Newell, Simon, ...).

Quando il programma "girava" sul calcolatore, diventava possibile individuarne le lacune, suggerendo le modifiche per migliorarlo e trovando nella sua prestazione anche corroborazioni o smentite della teoria generale. Per esempio, del Logic Theorist, che sotto questo profilo risultò essere una simulazione molto grossolana, venne data ben presto una nuova versione, basata su un confronto più puntuale delle euristiche implementate nel programma con i protocolli verbali di singoli soggetti alle prese con problemi di logica. Parliamo del ben noto General Problem Solver che, a giudizio dei suoi ideatori, dimostrava in molti casi di essere una buona simulazione del protocollo verbale di un singolo soggetto umano alle prese con la dimostrazione di un semplice teorema di logica proposizionale.

Sulla base di queste esperienze, Newell e Simon hanno ritenuto che il test di Turing fosse un test «debole», come essi dicevano, proprio perché non riguardava i processi del pensiero, ma solo la prestazione finale. E per la stessa ragione la sconfitta di un essere umano da parte di un calcolatore che giochi a scacchi, come nel caso di Deep Blue, benché interessante, non è un risultato particolarmente significativo dal punto di vista di quanti sono rimasti interessati al realismo psicologico della simulazione. «Processi diversi danno luogo a risultati simili»: questo è del resto il commento dei costruttori di Deep Blue, i quali, con parole che Newell e Simon avrebbero probabilmente sottoscritto, sottolineano il «contrasto di stili» tra uomo e macchina durante il gioco (come si legge nella pagina web dell'IBM dedicata al confronto tra Kasparov e Deep Blue).

In definitiva, i programmi sviluppati da Simon ai tempi della prima IA possono essere visti come altrettanti tentativi di costruire modelli cognitivi, impiegati per attribuire una struttura funzionale interna a parti del sistema cognitivo e per dare conto su questa base di varie proprietà esibite dal sistema stesso. Questi modelli vennero considerati significativi anche sotto il profilo predittivo. Attraverso il confronto con il comportamento dei soggetti umani, una simulazione algoritmica diventava uno strumento per controllare e raffinare ipotesi empiriche sui processi di pensiero. Può un approccio del genere portare alla costruzione di buoni modelli del mentale?

Tra le varie critiche sollevate in merito a questo punto, ci sono state quelle che hanno sottolineato l'incompletezza dei protocolli verbali: poiché non tutti i processi mentali coinvolti nella soluzione di un problema sono accessibili all'introspezione, la costruzione di un programma simulativo deve attingere a ipotesi arbitrarie su processi di pensiero non documentati nel protocollo. Per quali vie, allora, si possono introdurre ulteriori vincoli empirici sul modello, riducendone gli aspetti arbitrari?

Mentre Simon ha continuato a percorrere la strada della costruzione di microteorie o di teorie di «media ampiezza», come egli disse, affinando le tecniche dell'analisi dei protocolli verbali, Newell finì ben presto per ritenere che la costruzione di singole microteorie simulative portasse in un vicolo cieco, perché non permetteva di generalizzare plausibilmente i risultati raggiunti in direzione di una teoria "generale" o "unificata" della cognizione (l'antico obiettivo che la PPI aveva affrontato in termini di generalizzazioni qualitative). Egli propose di costruire i programmi simulativi sulla base di nuove restrizioni, indipendenti dai singoli compiti cognitivi e dai singoli soggetti, e che includessero proprietà invarianti del sistema cervello-mente (Newell, 1990). L'individuazione di proprietà invarianti dell'architettura funzionale del sistema cognitivo è stata al centro anche di un approccio molto diverso alla scienza cognitiva, quello di Pylyshyn (1984).

Altri, come Fodor (1983), hanno sostenuto la tesi di un'architettura modulare della mente, dunque opposta a quella dell'architettura unificata che Newell e diversi collaboratori hanno sperimentato nel sistema SOAR. Si tratta di una questione che è tuttora oggetto di diverse valutazioni all'interno della scienza cognitiva, ma va sottolineato che, rispetto alla PPI, fondamentalmente legata alla psicologia cognitivista, il contesto interdisciplinare nel quale si collocano le ricerche su SOAR si apre alle neuroscienze, ponendo nuove e difficili sfide epistemologiche per i modelli cognitivi algoritmici. La convergenza in un singolo modello di restrizioni provenienti da ambiti disciplinari diversi richiede che ci si interroghi sulla compatibilità di assunzioni teoriche e metodologiche soggiacenti le varie discipline coinvolte. Un'impostazione diversa, molto influente nell'IA e nella scienza cognitiva, al problema della costruzione di modelli algoritmici è stata suggerita da David Marr. Secondo Marr (1982), per costruire un buon modello algoritmico di una qualche capacità cognitiva bisogna anzitutto specificare che cosa dovrebbe essere in grado di fare un soggetto umano che possiede quella capacità, e cioè quali corrispondenze algoritmiche tra ingressi e uscite (ovvero quali funzioni Turing-calcolabili) caratterizzano l'esercizio di quella particolare capacità cognitiva. In assenza di tale analisi non si ha un'idea sufficientemente chiara di che cosa si vuole simulare o modellare.

Newell e Simon, a giudizio di Marr, avrebbero mostrato poca attenzione verso questo problema, concentrandosi prematuramente sui modi in cui i soggetti umani svolgono un compito cognitivo - un problema, questo, subordinato epistemologicamente e temporalmente all'analisi astratta del compito. In un contesto ancora preteorico, Marr ha proposto di identificare il compito del sistema visivo umano con la trasformazione dei valori di intensità luminosa registrati dai fotorecettori della retina (gli ingressi del sistema visivo) in informazioni relative alla forma e alla posizione degli oggetti nel campo visivo (le uscite finali del sistema visivo). All'interno di questo sistema complessivo Marr ha distinto vari sottosistemi deputati a compiti più specifici e ha proposto di analizzarli con il medesimo impianto metodologico: bisogna anzitutto mostrare che le corrispondenze fra gli ingressi e le uscite del sottosistema oggetto di indagine possono essere ottenute, in linea di principio, mediante un procedimento algoritmico. Una volta fatto ciò, si può tentare di sviluppare il modello algoritmico analizzando i processi mediante i quali i soggetti umani determinano queste corrispondenze.

Marr ha proposto di astrarre, in una prima fase di costruzione del modello, dai concreti processi di esecuzione del compito e dalle limitazioni del sistema cognitivo degli esseri umani che influenzano l'esecuzione del compito. È evidente a questo punto il conflitto con Newell e Simon, secondo i quali, per arrivare a modelli algoritmici con un qualche potere predittivo ed esplicativo, bisogna da subito prendere in considerazione i limiti del sistema cognitivo e le sue strategie euristiche. Nei modelli connessionisti, l'analisi astratta richiesta da Marr si riduce generalmente alla selezione di esempi significativi di ingressi o di uscite associati all'esecuzione del compito preso in considerazione. Elaborando tali esempi mediante un appropriato algoritmo di apprendimento, la rete neurale dovrebbe imparare a svolgere correttamente il compito anche in casi non considerati nella fase di apprendimento. Il confronto tra processi cognitivi e processi di elaborazione neurale, al di là di mere relazioni ingresso/uscita, è reso problematico dalla stessa natura dell'elaborazione neurale, che avviene attraverso la propagazione di una miriade di segnali spesso privi di una plausibile interpretazione cognitiva.

Nei modelli connessionisti più rappresentativi l'unità significativa sul piano cognitivo o simbolico è un gruppo più o meno ampio di neuroni in uno stato complessivo determinato. Ma va riconosciuto che i limiti dei modelli cognitivi connessionisti, nonostante l'ardita rinascita del connessionismo, restano per ora quelli di sempre: non è ancora chiaro se e come essi possano riuscire a simulare le capacità cognitive "superiori" (ad esempio inferenziali) e non solo quelle "inferiori" (associazioni, semplici forme di apprendimento).

Come hanno concluso gli stessi Anderson e Rosenfeld: "Al momento le nostre reti, dopo trent'anni di progressi, funzionano ancora come "cervelli lesionati" [incapaci di comportamento simbolico, secondo l'espressione di Rosenblatt]. Resta aperta la questione di quali severe modifiche apportare alla teoria delle reti neurali affinché esse riescano a raggiungere le funzioni cognitive superiori." (Anderson, Rosenfeld, 1998, p. 91).

Nei modelli situati, è la dinamica del sistema formato da un agente algoritmico dotato di un corpo e dall'ambiente circostante a diventare oggetto principale di indagine. Anche in questo caso, i vantaggi dei modelli situati (robustezza, azione in tempo reale e, particolarmente nel caso dei robot evolutivi, anche miglioramento delle prestazioni) si pagano con alcuni limiti (incapacità di ragionamento e di pianificazione complessa). E anche in questo caso si sperimentano approcci variamente "ibridi", che cercano di combinare al meglio le caratteristiche positive dei diversi modelli. Ma nonostante tutto, i modelli situati non sembrano ancora accontentare i teorici radicali della cognizione situata. Clancey (1997) ritiene ad esempio che i robot à la Brooks o behavior based sono ancora troppo "pre-programmati" dal progettista, e non sono in grado di sviluppare in modo sufficientemente autonomo l'interazione con l'ambiente.

IA e filosofia della mente

Una delle prime discussioni filosofiche intorno all'IA ha preso le mosse dai teoremi di incompletezza e di indecidibilità. Già Turing (1950) aveva discusso il problema da una prospettiva particolarmente adatta agli scopi di una modellistica cognitiva, utilizzando il teorema della fermata da lui dimostrato quindici anni prima. In base alla tesi di Church-Turing, questo risultato ci permette di concludere che nessun algoritmo consente di rispondere correttamente a tutte le infinite domande della seguente forma: «Sia X un programma e Y un particolare ingresso per X. Si dà il caso che X si fermerà in un tempo finito fornendo un'uscita determinata in corrispondenza dell'ingresso Y?».

E inoltre, per ogni dato programma a, è possibile esibire una domanda di questo tipo tale che, se a risponde in un modo qualsiasi, allora la sua risposta sarà dimostrabilmente errata. Turing commentava: "Ogni volta che viene posta a una di queste macchine un'opportuna domanda critica,

ed essa dà una risposta definita, noi sappiamo che questa risposta deve essere errata, e questo ci dà un senso di superiorità. Questo senso di superiorità è illusorio? Certamente esso è genuino, ma non credo vi si possa attribuire troppa importanza. Diamo troppo spesso risposte errate anche noi, per sentirci giustificati nel provar piacere davanti a tali prove della possibilità di errore da parte della macchina (Turing, 1950, p. 179)." Turing sosteneva che bisogna concedere la possibilità di sbagliare alle macchine che intendiamo usare per imitare o modellare i comportamenti cognitivi degli esseri umani.

Dopotutto, ammettiamo senza battere ciglio questa possibilità nel caso dei matematici più dotati. Ma se ammettiamo che le macchine sbagliano, allora non sarà più possibile concludere, in base al teorema della fermata, che le macchine alla quali siamo interessati nella modellistica cognitiva non sono in grado di fornire risposte a certe domande critiche. In altre parole, ammettere macchine fallibili come modelli ragionevoli delle capacità cognitive umane blocca alla radice la possibilità di sfruttare il teorema della fermata per argomentare che gli esseri umani sanno dare risposte corrette a certi problemi matematici mentre i loro presunti modelli algoritmici non sono in grado di farlo. Analoghe considerazioni valgono per tentativi simili che prendono lo spunto dai teoremi di incompletezza di Gödel (cfr. Tamburrini, 1997).

Turing sottolineava come forti idealizzazioni a proposito delle capacità matematiche degli esseri umani sono necessarie (ma non sufficienti) per ipotizzare una qualche rilevanza dei teoremi limitativi in questo contesto, ma suggeriva che l'utilità di tali idealizzazioni è perlomeno dubbia. Per quanto la storia della scienza mostri come in molti casi le giuste idealizzazioni siano state cruciali per arrivare a leggi generali nell'ambito della fisica, non disponiamo di una base teorica e sperimentale analoga per decidere se e quali idealizzazioni siano le più opportune nello studio algoritmico di determinate capacità cognitive, e comunque non c'è alcun accordo tra i ricercatori su questo punto.

Abbiamo ricordato, tra i tanti possibili esempi, come Marr sostenesse l'opportunità di introdurre forti idealizzazioni nei modelli algoritmici, astraendo da vari fattori che limitano di fatto le prestazioni cognitive, mentre, a partire dai primi lavori di PPI, Newell e Simon abbiano percorso altre strade, a loro volta molto diverse tra loro. Fin dalle origini l'IA ha esercitato una notevole influenza sui filosofi interessati al problema mente-corpo e, più in generale, allo statuto epistemologico delle scienze tradizionalmente coinvolte nella spiegazione della vita mentale, come la psicologia e le neuroscienze.

L'esempio più chiaro di tale influsso, evidente già a partire dal celebrato articolo di Putnam (1960), è la tesi del funzionalismo, collegata strettamente alla tesi della realizzabilità multipla dell'intelligenza e, conseguentemente, all'ipotizzata autonomia dei livelli di spiegazione del mentale. Queste varie tesi sono state indipendentemente dibattute dagli stessi ricercatori della prima IA. Tanto la PPI quanto l'IA successivamente coinvolta nella scienza cognitiva hanno considerato la mente umana come un sistema di elaborazione dell'informazione, proponendo di indagarne i processi a uno specifico livello di analisi attraverso la costruzione di modelli algoritmici. Inizialmente in questo progetto non erano direttamente coinvolte le neuroscienze.

Per la PPI, ad esempio, il futuro si sarebbe incaricato di mostrare se e fino a che punto i modelli algoritmici possono essere messi in una qualche relazione (eventualmente di riduzione) con teorie biologiche o fisiche del cervello. In assenza di conoscenze approfondite sulla biologia e la fisica dei processi cognitivi, la psicologia, come PPI, avrebbe potuto svilupparsi autonomamente al livello dell'elaborazione dell'informazione (proprio come la chimica si è consolidata inizialmente al livello molecolare senza aspettare che fosse nota la struttura atomica della materia). La relativa autonomia dei livelli di spiegazione trovava ispirazione e insieme conferma nell'architettura del calcolatore. È proprio questa che suggerì ai padri dell'IA di respingere la posizione del riduzionismo estremo, o "laplaciano", che svela la propria paradossalità allorché si immagini di applicarlo ai possibili livelli di descrizione di un calcolatore. È come se noi, volendo spiegarne il funzionamento, rifiutassimo di descriverlo al livello del linguaggio di programmazione (il LISP o il C, poniamo), scegliendo quello, sottostante, del linguaggio macchina, o addirittura quello dello hardware, cioè del funzionamento del calcolatore come macchina fisica - un livello solo illusoriamente ("laplacianamente") assumibile come "più fondamentale", ma in realtà «acme della riduzione e dell'incomprensibilità», per usare le parole di Simon (1973, p. 26).

Questa concezione dei livelli del calcolatore ha fornito il nucleo teorico a differenti programmi di ricerca prima in PPI e poi in scienza cognitiva, e anche all'impostazione dei rapporti tra le varie discipline coinvolte, in particolare la psicologia e le neuroscienze. Una tesi "forte" sul ruolo del calcolatore in scienza cognitiva è quella relativa ai Sistemi fisici di simboli (SFS) proposta da Newell e Simon (1976). Un SFS è un calcolatore potenzialmente universale, relativamente al quale certe espressioni simboliche hanno una funzione denotante e possono essere interpretate. In particolare, si dice che un SFS interpreta un'espressione simbolica se questa denota un processo algoritmico e l'SFS può eseguire tale processo

quando riceve in ingresso quell'espressione simbolica. La tesi "forte" di Newell e Simon afferma che gli SFS, con le nozioni collegate di denotazione e interpretazione, forniscono una base esplicativa per tutto il comportamento intelligente degli esseri umani. Al giusto livello di analisi, le menti umane sarebbero identificabili con degli SFS, i quali sono dotati delle capacità di elaborazione simbolica necessarie e sufficienti per esibire intelligenza.

Newell (1982) ha successivamente interpretato l'impresa della scienza cognitiva attraverso l'individuazione di tre livelli fondamentali di descrizione di un SFS: il livello della conoscenza, quello sintattico o dei simboli e quello fisico. Questa idea è stata ripresa da Pylyshyn, che l'ha coniugata con l'ipotesi del "linguaggio del pensiero" di Fodor (sul quale cfr. Di Francesco, 2001). Per quest'ultimo, gli stati mentali sono rappresentazioni. Le rappresentazioni condividono con gli enunciati di un linguaggio naturale o formale una struttura combinatoria o, come più spesso si dice, compositiva: a partire dagli elementi base del linguaggio, iterando l'applicazione di regole per formare enunciati sintatticamente corretti, si può generare un numero illimitato di enunciati.

Le regole generative degli enunciati dei più comuni sistemi logico-formali esemplificano in modo elementare questa proprietà. Un'altra analogia con la logica è fondamentale per il linguaggio del pensiero di Fodor. Le regole sintattiche della logica hanno una giustificazione semantica. In alcuni sistemi di logica proposizionale, ad esempio, da una congiunzione della forma A&B possiamo immediatamente ottenere l'enunciato A.

Questa transizione è semanticamente corretta (necessariamente se A&B è un enunciato vero allora anche A sarà vero). Anche le regole di manipolazione delle rappresentazioni devono permettere transizioni tra rappresentazioni che abbiano una qualche giustificazione semantica. In altre parole, il linguaggio del pensiero, come ogni altro linguaggio, avrebbe una semantica delle rappresentazioni in grado di giustificare regole di manipolazione delle rappresentazioni, le quali costituiscono una sorta di "motore inferenziale" cognitivo. In forme diverse, questa concezione ha permeato molti settori della scienza cognitiva.

Le rappresentazioni e le loro regole di manipolazione possono avere realizzazioni multiple, in sistemi o agenti fisici diversi: nel calcolatore sotto specie di hardware a loro volta diversi, negli esseri umani sotto specie di strutture biologiche. Ma i tre livelli sopra ricordati (della conoscenza, dei simboli, fisico) sono stati anche considerati (ad esempio da Pylyshyn) autonomi e irriducibili l'uno all'altro in quanto livelli di spiegazione. Secondo questa concezione, non è possibile descrivere, poniamo, una scelta compiuta da un agente razionale senza rifarci alle sue intenzioni, convinzioni e conoscenze, in breve al vocabolario della psicologia del senso comune. Se rinunciassimo ad analizzare il comportamento al livello della conoscenza, dal momento che quella stessa scelta può essere effettuata in una molteplicità di modi, che coinvolgono diversissimi processi neurologici o fisici dell'agente, è difficile concepire la possibilità di formare leggi e spiegazioni generali del comportamento intelligente. Le ipotesi del funzionalismo e della realizzabilità multipla, unitamente al ruolo centrale attribuito alle rappresentazioni, hanno caratterizzato molti programmi di ricerca in scienza cognitiva e possono essere sintetizzate nel modo seguente: il calcolatore fornisce un adeguato strumento per un'indagine materialistica sui processi cognitivi, dal momento che esso dimostra come certe strutture fisiche, le quali codificano strutture di simboli, possono svolgere il ruolo di causa del comportamento intenzionale di un sistema.

Più specificamente, questo ruolo causale è assolto dalle rappresentazioni in virtù delle loro proprietà fisiche. Dunque, «i codici sono "psicologicamente reali", [...] il cervello è un tipo di sistema che elabora tali codici, e [...] i codici hanno un effettivo contenuto semantico» (Pylyshyn, 1984, p. 40). Critici diversi, come Paul Churchland, Patricia Churchland, Gerald Edelman e i teorici della cognizione situata, nel respingere l'ipotesi degli SFS hanno respinto anche il funzionalismo e quella che considerano un suo corollario, la tesi della realizzabilità multipla. Quest'ultima, mentre stabilisce l'autonomia della spiegazione psicologica, avrebbe contemporaneamente indotto a bandire lo studio del cervello, del corpo e spesso dell'ambiente reale dall'impresa della scienza cognitiva. Ma il funzionalismo e la tesi della realizzabilità multipla non sono un tratto esclusivo dell'IA cosiddetta simbolica: in realtà, è ben difficile che essi possano essere rifiutati da qualunque teorico dei modelli, sia esso sostenitore di modelli a reti neurali o di robot situati o di quello che Edelman chiama il «metodo neurale sintetico». In tutti i casi, infatti, i modelli incorporano in quanto artefatti (siano essi simulati o "situati"), le ipotesi di una teoria del comportamento.

Questi modelli sono costruiti per condividere l'organizzazione funzionale (non, ovviamente, la struttura specifica, biologica) degli organismi dei quali si vuole spiegare il comportamento, introducendo le restrizioni al livello (o ai livelli) scelti. È la scelta di tali restrizioni, la scelta del livello di idealizzazione o semplificazione, che caratterizza i diversi approcci della scienza cognitiva, piuttosto che l'opzione, che resta inevitabilmente comune, per il funzionalismo e per la tesi della realizzabilità multipla. Per dirla in breve, il metodo dei modelli è dato dal funzionalismo più le

restrizioni, e la tesi della realizzabilità multipla, al di là delle speculazioni di cui essa è spesso oggetto in filosofia della mente, resta il cuore dell'impresa di qualunque teorico dei modelli. Una prospettiva che appare effettivamente superata è l'idea che sia sempre possibile trascurare, nella costruzione dei modelli di IA, le restrizioni suggerite dalle neuroscienze, limitandosi a quelle che riguardano, per esempio, diversi aspetti dell'architettura cognitiva. Questo punto è stato enfaticamente sottolineato dallo stesso Newell: "L'indipendenza dalla struttura del cervello è stato un elemento retorico importante nella storia della PPI. Ritengo che sia stato all'epoca un atteggiamento importante e giusto; ma non ci può essere più questa voragine tra elaborazione della conoscenza e struttura cerebrale [...]. Il grande vantaggio è che molte altre restrizioni possono ora entrare in gioco [...]. Ma questo è un modo troppo astratto di porre la questione. Il grande vantaggio è che finalmente possiamo raccordare la biologia e la psicologia tramite una superautostrada [...]" (Newell, 1990, p. 483).

Altre carenze dei programmi di ricerca legati all'ipotesi degli SFS sono state evidenziate criticamente al suo stesso interno per quanto riguarda, per esempio, la mancata considerazione da parte della prima IA dell'interazione degli agenti con il mondo reale (Vera, Simon, ...) o della cooperazione tra agenti (Okada, Simon, ...). Ma è anche opportuno rilevare come i modelli della cognizione che si distaccano maggiormente da questa tradizione, introducendo restrizioni che riguardano il corpo, l'ambiente, l'evoluzione, lo sviluppo, realizzino efficacemente certe prestazioni (risposte in tempo reale, robustezza, abilità percettivomotorie) ma non riescono a realizzare con altrettanta efficacia funzioni cognitive "superiori" (processi decisionali, dimostrazione automatica, elaborazione del linguaggio naturale), ambiti nei quali riescono meglio modelli algoritmici caratterizzati da altre restrizioni.

Sembra perciò poco fondata l'idea che vi sia una classe di restrizioni privilegiate per i modelli, visto che quelle relative al corpo, o in generale quelle biologiche, non consentono di ottenere sempre i modelli algoritmici più soddisfacenti. Il metodo dei modelli, in tutte le versioni che abbiamo discusso, adotta una strategia inevitabilmente anti-olistica nei confronti di un oggetto così complesso come il sistema cervello-mente, cercando di costruire modelli approssimati di "parti" del sistema. Troppo spesso, tuttavia, non ci si è posto il problema delle relazioni di queste "parti" con tutto il resto del sistema. E troppo spesso l'IA ha interpretato questa strategia come l'autorizzazione a trascurare i fattori emotivi che influenzano le capacità e le prestazioni cognitive, nonostante sia da tempo noto l'impatto degli stati emotivi sui processi di apprendimento, oppure il ruolo di selezione svolto dalle emozioni nei processi decisionali. Vi sono, a partire dagli albori dell'IA, importanti eccezioni a questa tendenza a separare nettamente emozione e cognizione (Turing 2950, p. 150; Simon, 1967; Dyer, 1987). Ma questi spunti sono stati spesso approfonditi entro una prospettiva puramente cognitivista sulle emozioni: i processi emotivi producono informazioni, sotto forma di un bilancio dei vantaggi o dei danni potenziali risultanti da una determinata situazione (cfr. LeDoux, 1998, cap. 2; Castelfranchi, 1991).

I limiti di questa prospettiva sono emersi con prepotenza attraverso i lavori di vari neuroscienziati, tra i quali Damasio (1995), che ha studiato il rapporto tra emozioni e processo deliberativo razionale, partendo da osservazioni cliniche su pazienti affetti da lesioni alle cortecce prefrontali. Posti di fronte a una situazione che normalmente evoca forti reazioni emotive, questi pazienti sono incapaci di provare alcunché, pur conservando la capacità di descrivere vantaggi o danni potenziali di quella situazione. All'insensibilità emotiva, in questi pazienti si associa una marcata incapacità deliberativa: essi si perdono in lunghi ragionamenti, senza operare una scelta o compiendo scelte disastrose. Damasio ipotizza che solo le emozioni provate nei confronti di situazioni reali o immaginate consentono agli esseri umani di scartare immediatamente un gran numero di scenari possibili, potando drasticamente in questo modo lo spazio di un problema deliberativo e passando all'esame dei meccanismi meno veloci della valutazione razionale solo le poche alternative sopravvissute. Provare emozioni costituirebbe un potente metodo euristico per esplorare lo spazio di un problema decisionale.

La conclusione di Damasio, se corretta, pone l'IA davanti a importanti sfide. Anzitutto, è possibile progettare macchine che abbiano prestazioni comparabili agli esseri umani nel ragionamento pratico, senza che queste provino emozioni? E segnatamente per l'IA che ambisce a fornire modelli algoritmici dei processi deliberativi degli esseri umani: in che senso un calcolatore può provare piacere o dolore e come bisogna modellare il ruolo causale attribuito da Damasio all'esperire la qualità emotiva di una situazione? Le emozioni che si provano di fronte a situazioni reali o immaginate possono essere viste come qualità soggettivamente esperite di particolari stati mentali, cioè come particolari tipi di qualia. In filosofia della mente (Di Francesco, 2001), un quale è, più genericamente di un'emozione, la qualità soggettivamente esperita di uno stato mentale (per esempio, ciò che si prova nel vedere un oggetto di un determinato colore).

Le considerazioni di Damasio indicano un possibile ruolo causale, poco evidenziato nella discussione filosofica, di alcuni tipi di qualia nel processo cognitivo. Ma le discussioni filosofiche sui qualia riguardano soprattutto il tema dell'esperienza cosciente, di quella che viene anche detta coscienza fenomenica, spesso distinta dall'autocoscienza, che si fonda sul possesso di una idea del sé, dalla coscienza introspettiva, per la quale è determinante la capacità di auto-osservazione o da altre accezioni ancora del termine "coscienza". Se sia possibile costruire una macchina cosciente, in una o più accezioni di questo termine, è un problema al quale già accennava Turing (1950, p. 179), ed è ora ampiamente discusso nella filosofia della mente contemporanea (cfr. ad esempio Chalmers, 1996); ma nell'ambito dell'IA sono stati considerati solo aspetti della coscienza legati all'introspezione e alla riflessione, attraverso sistemi (come lo stesso SOAR) capaci di fornire resoconti articolati sui propri stati e processi interni e di modificare i propri programmi sulla base di tali resoconti (Cordeschi, Tamburrini, Trautteur, 1999).

Conclusioni

La breve storia dell'IA è caratterizzata da uno sviluppo tortuoso e anche tumultuoso, segnato da proposte e ripensamenti, entusiasmi e delusioni. Tutto questo è stato spesso descritto, all'interno e all'esterno della comunità dell'IA, come una sorta di lotta tra "paradigmi" contrapposti. A insistere sulla contrapposizione del loro paradigma "subsimbolico" con quello "simbolico" dell'IA sono stati soprattutto i connessionisti degli anni ottanta. In un primo momento, è sembrato di assistere al riaffiorare di una sorta di trauma: Rosenblatt aveva ragione, il libro di Minsky e Papert (1969) sui Percettroni, dando un peso eccessivo ad alcune limitazioni funzionali di vari tipi di Percettroni, aveva ingiustamente cancellato le reti neurali dal mondo della ricerca.

È vero che dopo il libro di Minsky e Papert ci fu una drastica riduzione nei finanziamenti della ricerca sulle reti neurali; ma questa fu proseguita da diversi ricercatori, come James Anderson, Eduardo Caianiello, Stephen Grossberg, Teuvo Kohonen. Un autorevole connessionista, James McClelland, ha dichiarato di non credere che l'evento decisivo per l'arresto della ricerca sulle reti neurali sia stato il libro di Minsky e Papert. Tenendo conto del fatto che la ricerca sulle reti si fa simulandole su calcolatore, semplicemente «non si era pronti per la ricerca sulle reti neurali. [...] La potenza di calcolo dei [calcolatori dei] primi anni sessanta era del tutto insufficiente» (citato da Crevier, 1993, p. 309).

È bene ricordare che questi limiti delle prestazioni dei calcolatori sono gli stessi che hanno condizionato lo sviluppo dell'IA simbolica: a puro titolo d'esempio, si pensi alla scelta del paradigma dei sistemi basati sulla conoscenza rispetto al precedente paradigma della ricerca euristica sui toy-problems o problemi giocattolo - una scelta che semplicemente non si sarebbe posta senza disporre di calcolatori con grandi memorie e una grande potenza di calcolo.

Le limitazioni dei primi calcolatori incoraggiavano la sperimentazione di euristiche "deboli" su problemi giocattolo, in quel momento considerati la vera Drosophila dell'IA. Visti retrospettivamente, i successi della prima IA possono apparire scontati, ma per l'epoca erano tali da incoraggiare certe scelte piuttosto che altre: la scelta per la "manipolazione euristica di simboli", invece che per l'"imitazione del cervello", invece che per la "rappresentazione della conoscenza". Nessuna linea di ricerca viene spazzata via da un libro se non è già debole per conto suo. Prima e dopo la contrapposizione simbolico-subsimbolico, sono state proposte altre contrapposizioni tra paradigmi, anche all'interno dell'IA simbolica.

Di volta in volta è stato detto che erano contrapposti il paradigma della ricerca euristica e quello della conoscenza; quello logicista e quello antilogicista; quello dichiarativista e quello proceduralista, quello individuale e quello sociale o multi-agente. Ma è evidente che nessuno di questi può essere riconosciuto come un paradigma nel senso tecnico introdotto da Thomas Kuhn nel suo celebre volume La struttura delle rivoluzioni scientifiche. Piuttosto, ciascuno di essi descrive la parola d'ordine di indirizzi di ricerca diversi e anche rivali, in una storia come quella dell'IA in cui si intraprendevano e poi si abbandonavano le strade più diverse, salvo poi riprenderne qualcuna a distanza di tempo.

In questo tumultuoso panorama, l'evento dell'emarginazione e poi della ripresa delle reti neurali non appare in fondo così sconcertante come è stato qualche volta descritto. E il fatto che sia possibile accostare e mescolare IA simbolica, connessionismo, cognizione situata in tante realizzazioni "ibride" sembra confermarlo, rendendo un esercizio retorico, se non propagandistico, la loro separazione in paradigmi contrapposti.

Riferimenti bibliografici

- ANDERSON J. A., ROSENFELD E. (eds.) (1988), *Neurocomputing*, MIT Press, Cambridge (MA).
- BURATTINI E. (1993), Reti neuroniche e sistemi esperti, in "Lettera Matematica PRISTEM", 9, pp. 31-7.
- CARLUCCI AIELLO L., NARDI D., PIRRI F. (2001), Case Studies in Cognitive Robotics, in V. Cantoni, V. Di Gesù, A. Setti, D. Tegolo (eds.), *Human and Machine Perception .: Thinking, Deciding, and Acting*, Kluwer, Dordrecht.
- CASTELFRANCI C. (1991), Emozione e regolazione del comportamento, in T. Magri, F. Mancini (a cura di), *Emozione e conoscenza*, Editori Riuniti, Roma.
- CHALMERS D. J. (1996), *The Conscious Mind*, Oxford University Press, Oxford (trad. it. *La mente cosciente*, McGraw Hill-Italia, Milano 1999).
- CLANCEY W. J. (1993), *Situated Action*, in "Cognitive Science", 17, pp. 87-116.
- CLARK A. (1997), *Being there*, MIT Press, Cambridge (MA) (trad. it. *Dare corpo alla mente*, McGraw Hill-Italia, Milano 1999).
- CORDESCHI R. (1984), La teoria dell'elaborazione umana dell'informazione. Aspetti critici e problemi metodologici, in V. Somenzi (a cura di), *Evoluzione e modelli*, Editori Riuniti, Roma, pp. 321-422.
- CORDESCHI R., TAMBURRINI G., TRAUTTEUR G. (1999), The Notion of Loop in the Study of Consciousness, in C. Musio, C. Taddei Ferretti (eds.), *Neuronal Bases and Psychological Aspects of Consciousness*, World Scientific, Singapore.
- CREVIER D. (1993), *AI. The Tumultuous History of the Search for Artificial Intelligence*, Basic Books, New York.
- DAMASIO A. (1995), *L'errore di Cartesio*, Adelphi, Milano.
- DAVIS M. (1982), *Computability and Unsolvability*, Dover, New York.
- DI FRANCESCO M. (2001), *Introduzione alla filosofia della mente*, Carocci, Roma.
- DORIGO M., GAMBARDELLA L. M. (1997), Ant Colonies for the Traveling Salesman Problem, in "Biosystems", 43, pp. 73-81.
- DYER M. G. (1987), Emotions and their Computations. Three Computer Models, in "Cognition and Emotion", 3, pp. 323-47.
- FODOR J. (1983), *The Modularity of Mind*, MIT Press, Cambridge (MA) (trad. it. *La mente modulare*, Il Mulino, Bologna .1988).
- LANGTON C. G. (ed.) (1995), *Artificial Life. An Overview*, MIT Press, Cambridge (MA).
- LEDOUX J. (1998), *Il cervello emotivo*, Baldini e Castoldi, Milano.
- MARR D. (1982), *Vision*, Freeman, New York.
- MINSKY M. L., PAPER S. (1969), *Perceptrons*, MIT Press, Cambridge (MA). Ristampato nel 1988 con una prefazione e una postfazione degli autori.
- NEWELL A. (1982), The Knowledge Level, in "Artificial Intelligence", 18, pp. 87- 127.
- ID. (1990), *Unified Theories of Cognition*, Harvard University Press, Cambridge(MA).
- NEWELL A., SIMON H. A. (1982), *Human Problem Solving*, Prentice-Hall, Englewood Cliffs (NJ).
- IDD. (1976), Computer Science as Empirical Inquiry: Symbols and Search, in "Communications of the ACM", 19, pp. .113-26 (trad. it. in B. G. Bara, a cura di, *Intelligenza artificiale*, Angeli, Milano 1978).
- OKADA T., SIMON H. A. (1997), Collaborative Discovery in a Scientific Domain, in "Cognitive Science", 21, pp. 109-46.
- PUTNAM H. (1960), *Minds and Machines*, in S. Hook (ed.), *Dimensions of Mind*, New York University Press, New York. Ristampato in H. Putnam, *Minds, Language and Reality*, Cambridge University Press, Cambridge (MA) 1975 (trad. it. *Menti e macchine*, in *Mente, linguaggio e realtà*, Adelphi, Milano 1987)
- PYLYSHYN Z. W. (1984), *Computation and Cognition. Toward a Foundation for Cognitive Science*, MIT Press, Cambridge (MA).
- SCHANK R. C. (1984), *The Cognitive Computer*, Addison-Wesley, Reading (MA) (trad. it. *Il computer cognitivo*, Giunti, Firenze 1989).
- SIMON H. A. (1967), Motivational and Emotional Control of Cognition, in "Psychological Review", 74, pp. 29-39.
- ID. (1973), The Organization of Complex Systems, in H. H. Pattee (ed.), *Hierarchy Theory*, Braziller, New York, pp. 1-28.
- SOMENZI V., CORDESCHI R. (1994), *La filosofia degli automi. Origini dell'Intelligenza Artificiale*, Bollati Boringhieri, Torino.
- TAMBURRINI G. (1997), Mechanistic Theories in Cognitive Science: The Import of Turing's Thesis, in M. L. Dalla Chiara et al. (eds.), *Logic and Scientific Methods*, Kluwer, Dordrecht.

- TURING A. M. (1937), On Computable Numbers, with an Application to the Entscheidungsproblem, in "Proceedings of the London Mathematical Society", 42, pp. 230-65.
- ID. (1950), Computing Machinery and Intelligence, in "Mind", 59, pp. 433-60 (trad. it. Macchine calcolatrici e intelligenza, in Somenzi, Cordeschi, 1994).
- VERA A. H., SIMON H. A. (1993), Situated Action: a Symbolic Interpretation, in "Cognitive Science", 17, pp. 7-48.

Per saperne di più I seguenti due volumi sono ottime guide nello studio degli argomenti della scienza cognitiva:

- BECHTEL W., GRAHAM G. (eds.) (...), A Companion to Cognitive Science, Blackwell, Oxford.
- WILSON R. A., KEIL F. C. (eds.) (...), The MIT Encyclopedia of Cognitive Sciences, MIT Press, Cambridge (MA).

Anche da consultare:

- FLORIDI L. (ed.) (...), The Blackwell Guide to the Philosophy of Computing and Information, Blackwell, Oxford.

Inoltre:

- BECHTEL W., ABRAHAMSEN A. (...), Connectionism and the Mind, Blackwell, Oxford. Un'eccellente introduzione ai problemi della filosofia della mente in relazione al connessionismo e alle reti neurali.
- CORDESCI R. (...), The Discovery of the Artificial, Kluwer, Dordrecht. Una ricerca sui modelli della vita mentale e sulle loro implicazioni filosofiche a partire dal connessionismo di Thorndike e Hull fino alla cibernetica e agli attuali sviluppi della "nuova" IA.
- RANDELL B. (ed.) (...), The Origins of Digital Computers, Springer, Berlin. Una raccolta di testi fondamentali nella storia dei calcolatori. Comprende anche parte del celebre First Draft di von Neumann.
- RUSSELL S. J., NORVIG P. (...), Artificial Intelligence. A Modern Approach, Simon and Schuster, Englewood Cliffs (NJ) (trad. it. a cura di L. Carlucci Aiello, Intelligenza artificiale. Un approccio moderno, UTET, Torino ...). Uno dei manuali più completi e aggiornati di IA, con numerose schede sulla sua storia. Il suo studio costituisce l'approfondimento ideale degli argomenti trattati in tutti i capitoli del presente volume. È attualmente in preparazione la seconda edizione inglese.
- TRAUTTEUR G. (a cura di) (...), Consciousness: Distinction and Reflection, Bibliopolis, Napoli. Una raccolta di testi sulla coscienza rilevanti dal punto di vista dell'IA.

Siti Internet

- <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>. Lo storico documento di Dartmouth nella sua versione completa.
- <http://www.chess.ibm.com/>. Il sito della IBM dedicato ai confronti scacchistici tra uomo e calcolatore.
- <http://vimp.museophile.com/computing.html>. Un sito completo sulla storia dei calcolatori.
- <http://www-history.mcs.st-and.ac.uk/history/Mathematicians/Turing.html>. Sito dedicato ad Alan Turing.
- <http://heinz.library.cmu.edu/Newell/>. Archivio degli scritti di Allen Newell presso la biblioteca della Carnegie Mellon University di Pittsburgh.
- <http://www.psy.cmu.edu/psy/faculty/hsimon/hsimon.html>. La home page di Herbert Simon.
- <http://www-formal.stanford.edu/>. La home page del Formal Reasoning Group della Stanford University, dalla quale si può accedere alla home page di John McCarthy.
- <http://www.ai.mit.edu/people/minsky/minsky.html>. La home page di Marvin Minsky.

