

Appunti per un corso di comunicazione multimediale (6)

Problemi di immagine

Si dice che un'immagine vale più di mille parole, e spesso è vero. Quando poi è in gioco la comunicazione tra l'uomo e il computer, la maggior parte dei messaggi trasmessi dalla macchina compare sul monitor, ed è quindi affidata alla vista. Questo significa che nella comunicazione multimediale l'aspetto visivo svolge un ruolo di primo piano. È necessario quindi studiare con attenzione i meccanismi dell'acquisizione delle informazioni visive da un sistema digitale, attraverso un monitor o un videoproiettore, tenendo ben presenti le limitazioni che questi mezzi presentano allo stato attuale della tecnologia

di Manlio Cammarata

Dobbiamo fare un passo indietro e tornare un attimo al primo di questi appunti, pubblicato sul n. 136. In quelle pagine abbiamo visto che l'uomo dispone di *canali percettivi*, di solito definiti *sensi*, attraverso i quali riceve informazioni. In questa sede ci interessano il canale visivo e il canale uditivo. La comunicazione è affidata a *vettori*, genericamente definiti *media*, che sono la parola, l'immagine, il suono e così via.

Questa classificazione comporta la creazione di uno schema (fig. 1), nel quale vediamo che la parola può essere «ascoltata» oppure «vista», cioè può interessare di volta in volta il canale visivo o quello uditivo. Nella puntata del mese scorso abbiamo cercato di approfondire questo aspetto. Nel primo caso anche il

testo diventa immagine.

Ora consideriamo un fattore essenziale: la maggior parte delle informazioni che noi riceviamo da un sistema informatico appare su un video, cioè passa per il canale visivo. E questo è opportuno, perché attraverso la vista si può ricevere un'informazione più analitica e soprattutto meglio memorizzabile. Dunque le immagini costituiscono il vettore più efficace di cui disponiamo. Ma che cosa è l'immagine?

Dire che l'immagine è tutto ciò che si vede (nel nostro caso su un monitor) è banale, ma è vero e comporta due conseguenze non indifferenti. La prima, come si è già detto, è che anche la parola scritta è immagine; la sua efficacia dipende in buona parte dall'impaginazio-

ne e per questo il suo trattamento viene affidato a un grafico, cioè a un esperto di comunicazione visiva. La seconda conseguenza è che molto spesso sul video compaiono insieme informazioni di diverso tipo (scritte, fotografie, disegni, filmati, ecc.) e bisogna quindi capire sia come devono essere presentate le singole informazioni, sia da quali rapporti devono essere legate.

Fisse o in movimento

La prima cosa da fare è definire i diversi tipi di immagine che possono far parte di una comunicazione multimediale. La distinzione più immediata è tra immagini fisse e immagini in movimento.

Tra le immagini fisse possiamo avere fotografie, o disegni simili a fotografie, e disegni veri e propri. Tra i disegni bisogna distinguere quelli di tipo illustrativo da quelli di tipo tecnico, e fra questi ultimi comprendiamo anche le «chart», cioè schemi e diagrammi impiegati per l'illustrazione di statistiche e cose del genere.

Anche per le immagini in movimento si deve fare una distinzione, abbastanza simile alla precedente, tra *filmati* e *animazioni*. Nel filmato abbiamo una sequenza di fotogrammi, concettualmente simile al prodotto del cinema o della televisione, mentre con il termine «animazione» intendiamo una sequenza realizzata apportando una serie di modifiche da un'immagine iniziale a una finale. Per avere un esempio di quello che si intende per animazione, pensiamo a un oggetto disegnato con un programma CAD tridimensionale, che offra la possibilità di far ruotare l'oggetto stes-

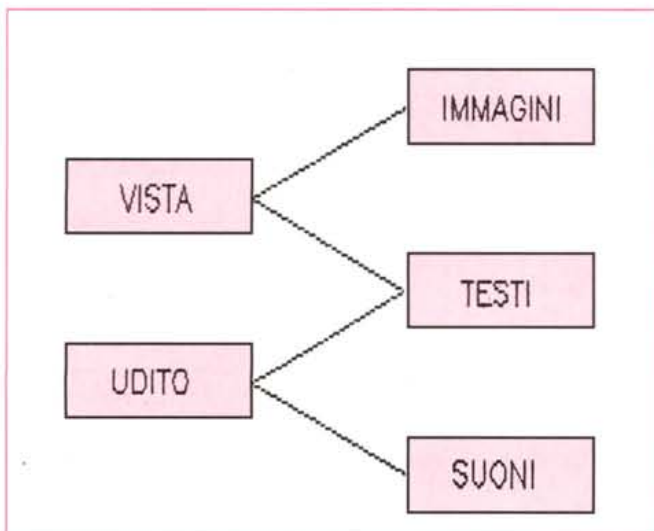


Figura 1.

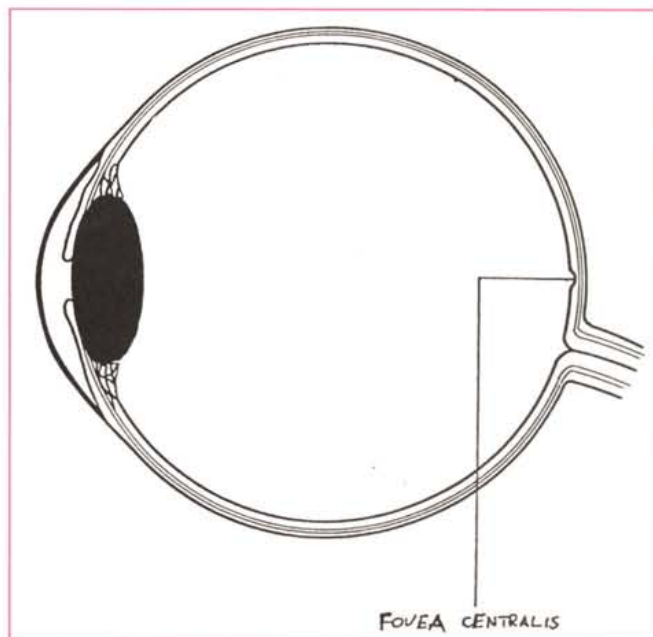
so per presentarlo sotto diverse prospettive. Oppure, più semplicemente, a un «diagramma a torta» con uno spicchio che si stacca per mettere in evidenza un dato.

A questo punto il problema è capire come deve essere impostata la parte visiva di una comunicazione multimediale. Per questo è necessario prima di tutto esaminare a grandi linee il meccanismo della visione, e quindi metterlo a confronto con le possibilità di restituzione delle immagini offerte dal monitor di un PC.

Come funziona l'occhio

Qualsiasi individuo dotato di una buona vista può abbracciare con lo sguardo, e senza muovere gli occhi, un campo di circa 180°. Tuttavia la percezione di questo spazio è piuttosto confusa, anche se di solito nessuno se ne rende conto. La visione nitida è limitata a un campo molto più ristretto, circa 2°, ed è dovuta al fatto che solo in una piccola zona della retina (*fovea centralis*, fig. 2) è presente una grande quantità di recettori. È questo il motivo per cui noi muoviamo continuamente gli occhi: dobbiamo fare in modo che l'oggetto del nostro interesse venga a cadere in corrispondenza della zona di visione nitida. È facile rendersene conto osservando gli occhi di una persona che legge: essi compiono una serie di piccoli scatti da sinistra verso destra, seguendo le singole parole, e poi compiono un salto più lungo verso sinistra per incominciare un'altra riga. Fate voi stessi un piccolo esperimento: fissate questa parola scritta in **neretto** e poi, senza muovere l'occhio, cercate di leggere qualcosa di quello che è scritto nella pagina accanto. Vi accorgete subito che, se non muovete l'occhio, il testo appare sfocato. Ora fate un'altra prova: allontanate lentamente la pagina, e a un certo punto il testo sarà così piccolo che non potrete più leggerlo. Al contrario, avvicinando la pagina all'occhio vedrete una serie di particolari (la grana della carta, al limite il «retino» di un'immagine a colori) che non sono visibili alla normale distanza di lettura. Questo avviene a causa del diverso ingrandimento dell'immagine determinato dalla distanza di visione. Quando un dettaglio è

Figura 2.



troppo piccolo in confronto alla dimensione dei recettori della parte centrale della retina, scompare. Quando invece l'ingrandimento è tale da interessare un sufficiente numero di recettori, ecco che il dettaglio più piccolo diventa visibile.

Forse si può capire meglio il meccanismo osservando il monitor del computer in un punto dove l'immagine è bianca con l'aiuto di una lente di ingrandimento. Con un po' di attenzione si possono vedere i puntini verdi, rossi e blu che formano l'immagine (ogni triade forma un *pixel*, acronimo di *picture element*). Nulla che sia più piccolo di un pixel può essere visualizzato, mentre le immagini composte da molti pixel sono più «leggibili». La nostra retina funziona più o meno nello stesso modo.

Si vede col cervello

Questi meccanismi fisiologici sono alla base delle operazioni mentali che portano alla visione «intelligente», oltre la semplice percezione di forme e di colori. In primo luogo il nostro cervello ricava la sensazione della dimensione di un oggetto confrontando la sua distanza

(la cui sensazione dipende in primo luogo dall'effetto stereoscopico della visione binoculare) con la sua dimensione sulla retina, o con il movimento panoramico dello sguardo necessario per abbracciarlo tutto. La dimensione di un oggetto determina una serie di effetti psicologici molto importanti: pensate al senso di timore che possono suscitare le grandi ruote di un autotreno viste da vicino stando in una piccola automobile o all'effetto di un bel panorama visto da un luogo elevato o su un grande schermo cinematografico; oppure alle diverse sensazioni che possono dare la vista dal vero di una lucertola e quella di un coccodrillo, che sullo schermo di un computer, in mancanza di qualche elemento di paragone, possono sembrare molto simili.

Un'altra fondamentale conseguenza dei meccanismi fisiologici della visione è il modo in cui «leggiamo» le immagini, influenzato anche da fattori culturali. Il nostro occhio tende a esplorare un'immagine «piatta», cioè senza elementi di particolare evidenza per forma, colore o dimensioni, partendo dall'alto e da sinistra, come se leggesse un testo. Se invece c'è un particolare che, come

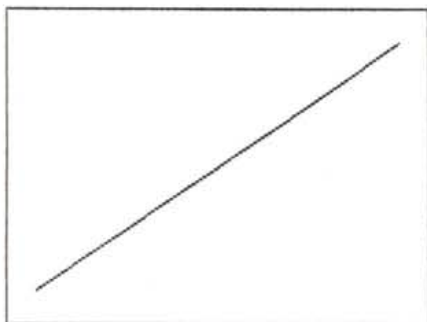


Figura 3a.

si dice «salta all'occhio» per la forma, la dimensione o il colore, noi lo vediamo prima del resto, e poi tendiamo a considerare gli altri elementi dell'immagine in relazione al più importante. Se poi nell'inquadratura compaiono delle linee significative, esse diventano come frecce e guidano la direzione del movimento dell'occhio, o comunque dell'attenzione. Ma questo ha anche un effetto psicologico: una linea obliqua che parte dall'angolo destro in basso per finire in alto a sinistra (fig. 3a) dà un senso di progresso, di salita, di ottimismo (perché la leggiamo da sinistra), mentre una linea che va dall'alto a sinistra verso l'angolo inferiore destro (fig. 3b) rende l'effetto contrario, di discesa. Ma se trasformiamo queste linee in frecce, possiamo cambiare l'effetto (figg. 3c e 3d): la salita diventa discesa, e viceversa.

Gli effetti delle dimensioni e della composizione dell'immagine si combinano: maggiore è la dimensione, più ampi sono i movimenti dell'occhio, e quindi l'effetto della composizione è più evidente. Quando si parla di dimensione dell'immagine, però, si parla di dimensione relativa al campo visivo, ovvero di angolo di visione. Infatti un'immagine grande vista da lontano può fare lo stesso effetto di una più piccola vista a una distanza inferiore.

Anche il colore ha un'importanza fondamentale. Tutti conoscono i diversi effetti dei colori «caldi» e di quelli «freddi», sanno che tinte come il verde e il marrone infondono calma, che il rosso stimola l'aggressività e il blu la malinconia, e via discorrendo.

Se combiniamo insieme questi tre aspetti fondamentali (dimensioni, elementi significativi e colore), vediamo quanto complesso possa essere il trattamento delle immagini in qualsiasi tipo di comunicazione, e particolarmente in quella multimediale. Per questa dobbiamo prendere in considerazione anche

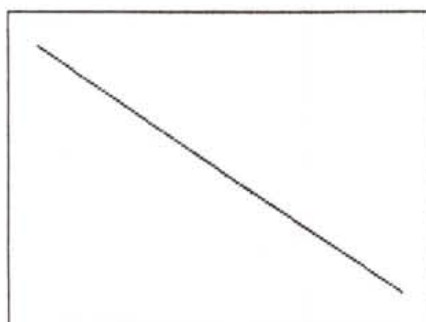


Figura 3b.

un altro elemento di grande importanza: la qualità dell'immagine che può essere raggiunta, allo stato attuale della tecnologia, dal video di un computer o da un sistema di videoproiezione, a seconda dell'impiego a cui è destinato un prodotto.

In sintesi

Nell'inserimento di qualsiasi tipo di immagine in un prodotto multimediale è necessario soddisfare una serie di requisiti che derivano dai meccanismi fisiologici e psicologici della visione.

Di un'immagine bisogna considerare non solo il contenuto esplicito, ma anche tutti gli elementi che possono influenzarne la comprensione.

In particolare va considerata la presenza di elementi che possono attirare l'attenzione o, al contrario, creare distrazione; bisogna anche valutare la composizione geometrica e l'effetto dei colori.

Tutto questo deve essere correlato alle dimensioni finali dell'immagine.

I limiti del video

Eccoci al dunque: il video non è una grande lastra fotografica, sulla quale possono essere registrati praticamente tutti i dettagli e tutte le sfumature di colore visibili dall'occhio umano, ma è un supporto di capacità piuttosto limitate. Il punto più critico è la *risoluzione*, cioè la riproducibilità di dettagli molto piccoli. Una pellicola fotografica del formato di 24 x 36 mm può facilmente raggiungere la risoluzione di 100 linee per millimetro, cioè 2.400 punti in verticale e 3.600 in orizzontale. Questo dato può ri-

manere teoricamente inalterato ingrandendo il fotogramma anche di dieci o cento volte (pari a un'immagine di 3,6 m di base). Invece il monitor di un PC ha una risoluzione, nel migliore dei casi, di 1024 x 1280 pixel, e questo significa che il dettaglio minimo che può riprodurre è grande più del doppio, in misura lineare, di quello del fotogramma ottico, cioè la risoluzione è la metà. C'è poi il problema della restituzione dei colori. Sulla pellicola è virtualmente possibile una gamma illimitata di colori, mentre in un'immagine digitalizzata il numero dei colori riproducibili è in funzione del numero di bit scelti per rappresentare ogni punto. Facciamo un esempio: un'immagine monocromatica digitalizzata a 8 bit per ogni pixel, può rendere (se il monitor lo consente) 256 sfumature di grigio. Un'immagine a colori, digitalizzata a 8 bit per ognuno dei tre colori primari (rosso, verde e blu, cioè 24 bit per punto) può rendere la bellezza di 16.000.000 di colori. Ma se moltiplichiamo 24 per il numero di punto dello schermo (1024 x 1280), scopriamo che un'immagine digitalizzata a questo livello supera i 30.000.000 di bit, cioè occupa quasi quattro megabyte di memoria. Si pone quindi il problema di immagazzinare queste immagini e di trattarle in tempi ragionevolmente brevi, operazioni che al giorno d'oggi sono alla portata di grandi sistemi informatici, non di personal computer. Se poi passiamo alla riproduzione di immagini in movimento, troviamo che a 24 fotogrammi al secondo (la cadenza del cinema), un minuto di film richiede (o, meglio, richiederebbe) circa 4 x 24 x 60 MB, cioè circa 5 gigabyte!

Il problema viene risolto con diversi accorgimenti. Il punto di partenza, nella maggior parte dei casi, è quello della risoluzione televisiva (525 linee verticali nel sistema nordamericano), molto più bassa di quella fotografica o cinematografica, quindi ci si accontenta di un minor numero di colori, e in questo modo la dimensione di un'immagine in bit si riduce di parecchi ordini di grandezza. E poi, soprattutto per le immagini in movimento, si rinuncia alla rappresentazione a pieno schermo, utilizzando solo una finestra che nella maggior parte dei casi è pari a circa un quarto della superficie disponibile. Ma in questo modo ci troviamo di fronte a una risoluzione di circa 260 x 190 linee di pixel: circa 50.000 pixel, contro gli oltre 2.000.000 di un fotogramma cinematografico 35mm. I colori, nella maggior parte dei casi, si riducono a 256. La quantità di informazioni, e quindi la qualità della comunicazione

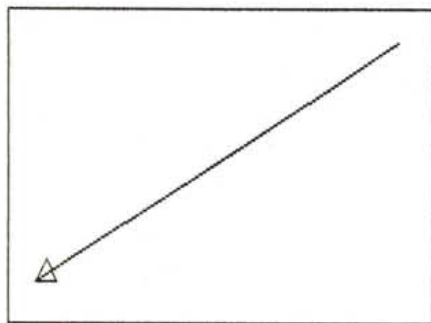


Figura 3c.

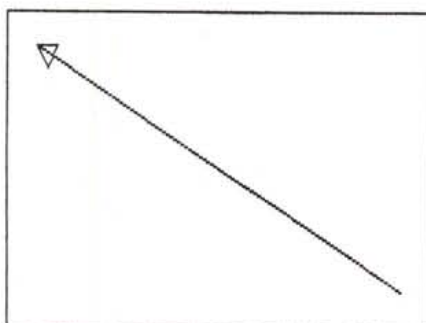


Figura 3d.

visiva, risulta quindi molto scarsa.

Ma anche così le dimensioni in bit di una sequenza di immagini in movimento sono troppo grandi per essere immagazzinate per una durata sufficiente su un CD-ROM, elaborate da un PC in tempo reale a 24 o 25 fotogrammi al secondo, o essere trasmesse via cavo, sempre in tempo reale. Questo problema viene risolto con la *compressione* delle informazioni digitali. In pratica, attraverso particolari calcoli matematici, si riduce il numero di bit necessari a descrivere un'immagine, e si può anche riprodurre un fotogramma registrando solo i pixel differenti da quello precedente. Inoltre, sempre con i sistemi attuali, la restituzione del movimento non avviene con la cadenza del cinema e della TV, ma a frequenze più basse, da 5 a 15 (nei casi migliori) fotogrammi al secondo. Se non ci accorgiamo del movimento a scatti lo dobbiamo proprio alla bassa qualità dell'immagine. Si aggiunga che gli algoritmi di compressione più efficaci sono quelli detti «a perdita di informazioni», che eliminano i dati che si ritengono inessenziali per la ricostruzione dell'immagine. I sistemi di compressione più recenti assicurano comunque immagini gradevoli, ma non c'è dubbio che il livello effettivo dell'informazione visiva ottenuta con questi metodi è molto più basso di quello offerto dai tradizionali sistemi ottici su pellicola.

Come sfruttare le immagini?

Tutto questo, in fondo, non è che una premessa per le poche osservazioni che seguono. Perché, viste le scarse possibilità di restituzione visiva che caratterizzano i sistemi informatici attuali, ottenere una buona comunicazione visiva dello schermo di un PC è decisamente problematico. Basta dare un'occhiata a qualcuno dei CD-ROM multi-

mediali oggi in vendita per rendersi conto di quanto la parte visiva sia di qualità scadente, in particolare per quanto riguarda i filmati. Bassa risoluzione, movimento spesso a scatti (e questo a volte dipende anche dal fatto che l'hardware «non ce la fa»); nel complesso le immagini fisse sono migliori, ma solo nei CD nati come repertori di immagini o in alcuni videogiochi vengono sfruttate a fondo le capacità grafiche delle macchine. In particolare, è ancora impossibile vedere un filmato in tempo reale a pieno schermo (*full motion, full video*), anche se forse è solo questione di mesi la disponibilità degli algoritmi di decompressione che consentiranno la qualità televisiva sul PC (dei sistemi di compressione e decompressione delle informazioni parleremo in dettaglio nel prossimo futuro).

Il dato di fondo è questo: lo schermo di un PC non può riprodurre la suggestione della proiezione di un film o di una serie di diapositive; con un videoproiettore dell'ultima generazione, che è in grado di mantenere la risoluzione della scheda video, la situazione migliora per un puro fatto psicologico, dovuto alla visione in una sala buia e purché la distanza non sia eccessiva. Ma la capacità di restituzione delle informazioni è quella che è, non resta che sfruttarla nel modo migliore.

E quindi, nel progettare la parte visiva di un prodotto multimediale, bisogna prima di tutto chiedersi se il contenuto di ogni immagine fissa, o di ogni filmato, contenga elementi di tale rilievo da poter essere restituiti in misura efficace anche nelle dimensioni e alla risoluzione della presentazione finale.

Facciamo un esempio: al cinema vediamo un paesaggio (campo lungo), in cui l'unico elemento di rilievo è un castello che occupa circa un quarto dell'inquadratura. Dalla porta del castello escono due persone che camminano

verso la macchina da presa e parlando tra loro. Nell'audio si sentono le loro voci. Sono ben visibili alcuni particolari del paesaggio, per esempio alcuni alberi su una collina e qualche nuvola nel cielo. La camera «stringe» in una lenta zoomata, e quando i due sono in primo piano a figura intera (campo medio) si intuiscono le espressioni e diventa chiaro il sincronismo tra suono e immagine. L'inquadratura si stringe ancora, fino al primo piano dei volti. A questo punto la rispondenza tra immagini e suono è assoluta e le espressioni sono molto suggestive.

Ora immaginiamo la stessa scena riprodotta da un CD-ROM su un quarto dello schermo del nostro PC. All'inizio vediamo un paesaggio privo di dettagli significativi, con la massa scura del castello e alcuni «scarabocchietti» di pixel sulla linea dell'orizzonte (sono gli alberi); dopo un po' notiamo dei pixel scuri che si muovono, e solo l'audio, che riporta le voci, ci fa supporre che si tratti di persone. Le due figure si ingrandiscono, ma solo quando arriviamo al primo piano possiamo riconoscere i movimenti della bocca che corrispondono (più o meno) alle parole. Tuttavia le espressioni non possono avere la stessa suggestione che offrono i due visi a tutto schermo in una proiezione al cinema.

Conclusione: di tutta la sequenza, solo la parte finale si presta ad essere inserita con sufficiente efficacia in un prodotto multimediale, sempre considerando i mezzi tecnici disponibili oggi. La prima parte della sequenza è praticamente inutilizzabile, perché sul piccolo schermo perde ogni significato.

In sintesi

Un personal computer, allo stato attuale della tecnologia, presenta una serie di limitazioni nella riproduzione delle immagini, e in particolare di quelle in movimento.

Il problema è nella grande quantità di dati necessari per ricostruire un fotogramma o una sequenza di fotogrammi, che comporta un grande impiego di memoria e di potenza di elaborazione; si ricorre quindi alla «semplificazione» e alla «compressione», che producono una perdita di informazioni. Dunque è necessario valutare ogni volta se gli elementi significativi di un'immagine possono essere mantenuti nella restituzione digitale sul video.

ME