



La visione

Seconda parte

I primi tentativi di codifica delle informazioni visive e di recupero e riconoscimento degli oggetti dalle loro immagini risalgono al 1940. Nel 1950 fu sviluppata la teoria del riconoscimento tramite pattern, griglia, di informazioni elementari (si veda la puntata scorsa) che tuttora resta la tecnica più efficace per l'analisi delle immagini. Purtroppo, in questo campo, non si sono avuti grossi progressi.

A dispetto dei notevoli progressi nel campo dell'informatica, la nostra conoscenza dei principi della percezione visuale rimane piuttosto primitiva. I ripetuti tentativi di progettazione e costruzione di modelli destinati alla interpretazione di una scena arbitraria sono risultati inficiati da prestazioni modeste, elevato limite d'errore, scarsa capacità di discernimento e flessibilità ridotta od assente, tanto da far concludere che, almeno per adesso, il riconoscimento generalizzato delle immagini è ancora di là da venire. A dimostrazione di ciò i progressi ottenuti negli ultimi dieci anni, sebbene talora promettenti, sono del tutto frammentari e poco affidabili.

I prolungati studi sulla visione finora svolti hanno comunque accertato che la soluzione del problema passa attraverso due stadi: l'analisi visiva del particolare il cui fine è di fornire informazioni efficaci provenienti da una immagine a mosaico; ogni parte dell'immagine (v. figura A) è analizzata allo stesso modo; nella fase successiva (fusione), si tenta di raggiungere lo scopo di riconoscere gli oggetti da informazioni utili da essi deducibili. Come si vede il processo iniziale di fusione, è da una parte formato da procedure parallele, dall'altro è locale (insieme identici di analisi di piccole porzioni di immagini). Processi di più alto livello passano poi attraverso di essi per estrarre una simbolica e per quanto possibile univoca descrizione della scena.

Vediamo in che modo è possibile ricavare da una immagine la maggiore informazione possibile e, per questo, rifacciamoci alla array figurata della puntata precedente. Ogni casella può

essere ulteriormente suddivisa in quattro parti (utile in questo caso una array tridimensionale) che evidenzerebbe, nella sua struttura, altrettanti parametri fondamentali della minima parte d'immagine considerata. Ogni cella conterrebbe così non più un solo dato, ma quattro quantità: illuminazione, riflessione, profondità ed orientazione, quest'ultima riferita alla superficie prevalente dell'immagine.

Ogni cella, caratterizzata da questi quattro vettori contribuisce, nel suo piccolo e per suo conto, a realizzare «l'immagine intrinseca» che, tra l'altro, traduce i valori di una realtà e di quantità fisiche in una serie di vettori che pur dipendendo tutti dalla quantità di luce, non sono del tutto comparabili tra di loro. In questo modo, e variando opportunamente i livelli di approssimazione delle quantità prima descritte, è possibile realizzare un numero pressoché infinito di combinazioni, necessarie a descrivere le ancora più infinite variabili della immagine sede.

Ma, e poteva mancare, il problema non si esaurisce qui. L'immagine intrinseca non è né facile né rapida da analizzare. Da una parte esiste ancora il problema di ricezione ed organizzazione di tale messe di dati, dall'altra, ovviamente, rimane ancora il problema dell'analisi di questi perché, ricomposti ed elaborati, diano risultati plausibili in termini di rappresentazione della immagine primitiva. In effetti la visione umana è estremamente avanzata in quanto l'analisi di quanto viene trasmesso dall'occhio viene eseguito, dal cervello umano, in parallelo, mentre l'analisi dei dati contenuti nella matrice di cui prima avviene in modo del tutto seriale. Questo comporta impiego di un certo lasso di tempo. È pertanto lecito affermare che una macchina vedente non potrà ignorare questo problema, e non potrà fare a meno di un computer capace di analisi di dati in parallelo funzionante in tempo reale.

Nella semplice analisi che abbiamo finora descritta esistono diverse peccate, sempre per il solito motivo che è

sempre arbitrario assegnare interpolazioni fisiche all'insieme di informazioni ottiche provenienti dalla immagine; lo spettro dell'ambiguità non trova posto migliore; infatti mai come in questo caso non esistono corrispondenze biunivoche.

Infatti, nella semplice osservazione di basso livello, a mezzo di celle elementari afferenti ognuna certe informazioni di base di una piccola parte di immagine, esistono certe proprietà di base dell'immagine, fondamentali per il riconoscimento della stessa, che non possono essere rilevate con tale sistema. Ad esempio non è possibile definire, nell'immagine intrinseca, una qualche orientazione generale della figura. Ancora, una macchina vedente non potrebbe, attraverso la visione di grado modesto, riconoscere proporzioni o distinguere tra l'immagine di un oggetto e quella prodotta da un filmato o da una foto.

Una soluzione appena parziale del problema potrebbe essere quella di limitare il riconoscimento solo ad oggetti ben definiti. Il problema viene certamente ristretto in margini più angusti, ma è ovvio che tutto il lavoro d'analisi si riduce in maniera drastica ed è possibile quindi dedicare maggiori risorse alle tecniche di analogia e corrispondenza immagine-oggetto. È quello che generalmente si fa con una macchina guidata attraverso una telecamera; l'operatore osserva solo quello che è necessario perché l'auto stessa non vada a sbattere e segua sempre la strada assegnata. Ma, per risolvere il problema di una più ampia visione non è possibile far riferimento ai singoli particolari senza poi fonderli in procedure di analisi visiva di livello più elevato. In altre parole, più che raccogliere informazioni particolareggiate circa la scena, come avveniva nella matrice di cui prima, occorre che la macchina vedente raccolga informazioni circa il «significato» della più complessa immagine che le singole parti formano.

La regola che governa l'assunzione delle immagini a livello più elevato è regolata dal principio euristico:

«Le caratteristiche osservabili in una immagine sono dovute alle proiezioni predominanti della corrispondente immagine in 3 dimensioni».

Ad esempio la visione rettangolare della figura B può essere, nel solo campo della figura bidimensionale, fornita da una qualsiasi delle figure di sinistra. Un altro esempio è rappresentato dalla tessitura di una superficie che, a ben guardare, non è altro che una serie di orientazioni di superficie. Si chiama «gradiente di tessitura» un cambio regolare e sistematico nella misura e nello spessore degli elementi concorrenti a creare la tessitura. Nella

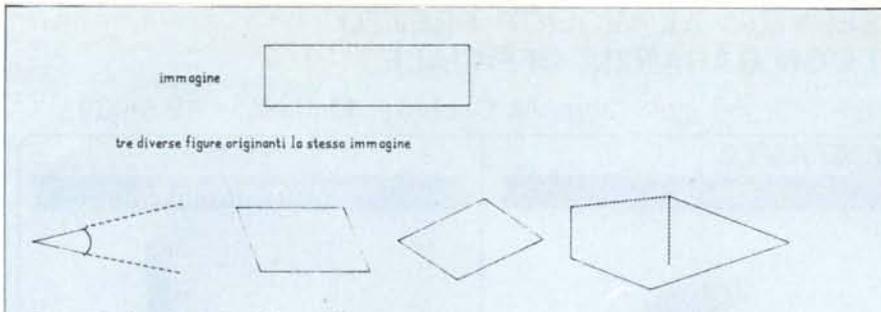


Figura A - Interpretazione di una visione.

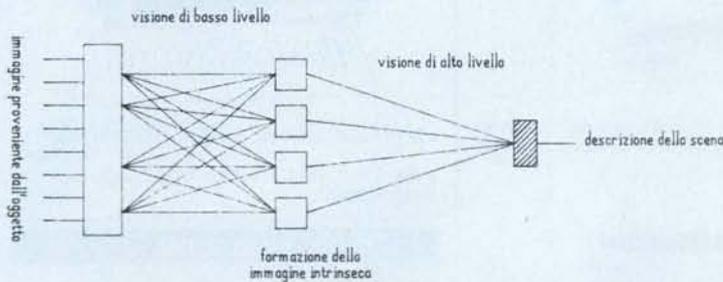


Figura B - Sviluppo degli studi di analisi nella formazione di una immagine.

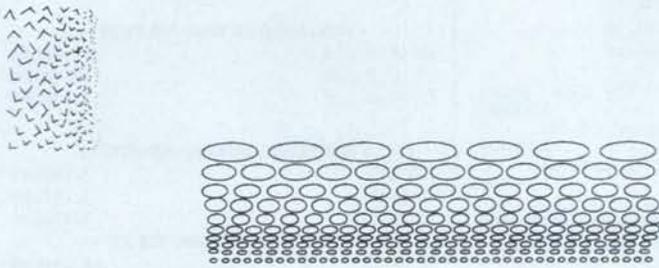


Figura C - Esempi di pattern simulanti illusione ottica.

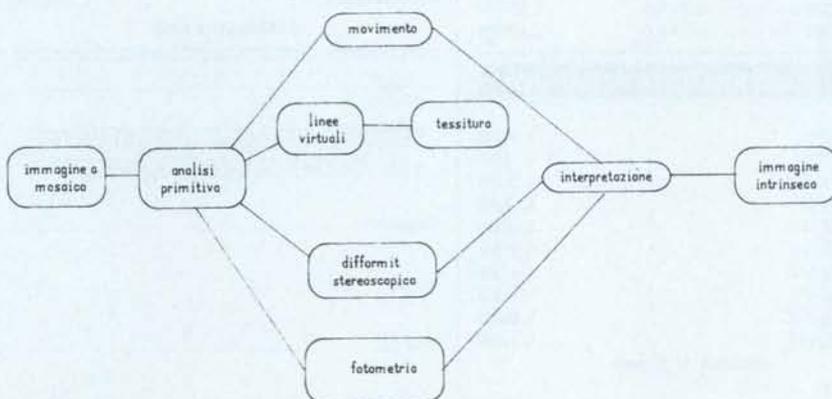


Figura D - Processo di formazione di una immagine intrinseca.

figura C si vede come la tessitura delle immagini determini l'effetto ottico di fuga dell'immagine dell'osservatore. È il classico esempio di come l'occhio, addirittura umano, possa essere tratto in inganno, in assenza di altri particolari. Immaginarsi come ciò può avvenire con maggiore facilità in una macchina costruita per analizzare un numero finito, ancorché elevato di situazioni.

La figura D, ricavata, come le altre, dalla bibliografia già più volte nominata in queste pagine, mostra uno schema di organizzazione del processo di analisi di visione a basso livello. Partendo dall'immagine prodotta dal flusso luminoso, si passa alla prima bozza, rappresentata, grossolanamente, da una descrizione dei bordi e di altre caratteristiche principali dell'immagine. A questo punto l'analisi della rappresentazione principale, ancora piuttosto bruta, può arricchirsi parallelamente attraverso altre metodologie di parsing. Vengono eseguite, senza precedenza relativa, analisi delle linee virtuali (nascoste dall'immagine stessa) della tessitura (significativa, come abbiamo visto, per individuare l'orientazione dell'oggetto); parallelamente attraverso l'analisi del movimento (corrispondenza di due immagini distanziate nel tempo), della riflessione, della stereoscopia (piccole differenze fra due vedute della stessa immagine, attraverso due punti di vista leggermente diversi, consentono di rilevare la profondità e l'orientazione di superfici) è possibile infine individuare, nell'immagine primaria, linee e regioni, corrispondenti a limiti ed aree, fisicamente e logicamente congruenti, contribuenti da una parte a formare l'immagine finale, dall'altra a consentire tentativi di individuazione degli oggetti concorrenti all'immagine. Si è pervenuti, così, all'«immagine intrinseca».

Ricordiamo che abbiamo escluso dal problema la percezione dei colori: ma già quanto abbiamo finora evidenziato è degno di studi piuttosto complessi, che attualmente hanno raggiunto i primi frutti, soprattutto di immagini statiche e a distanza prefissata. Se si considera che uno dei problemi cui si va incontro è anche quello della «scalatura» dell'immagine stessa si vede come anche una avanzata semplificazione del problema non può fare a meno di un numero enorme di variabili.

L'immagine primaria, articolata secondo il mosaico di informazioni prodotte dall'analisi di basso livello porta ad una messe di informazioni che occorre inserire in schemi logici omogenei. Il compito della macchina è quello di codificare tali blocchi di dati, per trarne l'informazione più prossima alla realtà.

Gruppo Distributori Associati

IL MIGLIOR SERVIZIO AL MIGLIOR PREZZO
PRODOTTI CON GARANZIA UFFICIALE

Sede: S. Martino Siccomario (Pavia) - S.S. dei giovi (ang. via Gabba) - ☎ 0382 - 49.94.39

COMPUTERS:

OLIVETTI



- **M24** - 640K RAM, 2 FDD 360K, VIDEO, TASTIERA. L. 3.450.000
- **M24** - 640K RAM, 1 UNITÀ FDD 360 K, 1 HD 20 Mb, VIDEO, TASTIERA. L. 4.450.000
- **M24 SP** - 640K RAM, 1 FDD 360K, 1 HD 20 Mb, VIDEO, TASTIERA. L. 5.500.000
- **M28** - 512K RAM, 1 FDD 1.2 Mb, 1 HD 20 Mb, VIDEO, TASTIERA. L. 6.990.000
- **M28** - 512K RAM, 1 FDD 1.2 Mb, 1 HD 20 Mb, VIDEO, TASTIERA, STREAMER DA 20 Mb. L. 7.980.000
- **M28** - 512K RAM, 1 FDD 1.2 Mb, 1 HD 40 Mb, VIDEO, TASTIERA, STREAMER DA 20 Mb. L. 9.100.000
- Box espansione 1 floppy da 369K autoalimentato × M28 L. 862.000
- Streamer interno da 20 Mb × M28 L. 1.912.000
- Streamer da 20 Mb in box esterno autoalimentato × M28 L. 2.378.000

EPSON



- **PC** - 256 K RAM, 2 FDD 360K, VIDEO, TASTIERA L. 2.079.000
- **PC/C** - 256 K RAM, 2 FDD 360K, VIDEOGRAFICO a colori, TASTIERA L. 2.620.000
- **PC/HD** - 256 K RAM, 1 FDD 360K, HD 20 Mb, VIDEO, TASTIERA L. 3.235.000
- **PC/HDC** - 256 K RAM, 1 FDD 360K, HD 20 Mb, VIDEOGRAFICO a colori, TASTIERA L. 3.775.000
- **PC+** - 640 K RAM, 2 FDD 360K, 8086, VIDEO GRAFICO, TASTIERA L. 2.850.000
- **PC+/HD** - 640 K RAM, 1 FDD 360K, 1 HD 20 Mb, VIDEO GRAFICO, TASTIERA L. 3.990.000
- **PC/AXM** - 80286 6-8-10 Mhz, 640K, 1 FDD 1.2 Mb, VIDEO, TASTIERA L. 3.990.000
- **PC/AX HD2** - 80286 6-8-10 Mhz, 640K, 1 FDD 1.2 Mb, 1 HD 20 Mb, VIDEO, TASTIERA L. 5.300.000
- **PC/AX HD4** - 80286 6-8-10 Mhz, 640K, 1 FDD 1.2 Mb, 1 HD 40 Mb, VIDEO, TASTIERA L. 6.550.000

12 MESI DI GARANZIA INTEGRALE

AMIGA



- **AMIGA** - 512 K RAM, VIDEO a colori, TASTIERA, MOUSE L. 2.050.000
 - **DRIVE 3,5"** - ESTERNO TECNOLOGIA TOSHIBA L. 349.000
 - **Dischetti** da 3,5" doppia faccia (min. 50 pezzi) L. 3.800
- Oltre 700 programmi disponibili con arrivi settimanali dalla Germania e dagli Stati Uniti!
Ogni sabato pomeriggio dimostrazione in sede.
(Richiedeteci la lista programmi).

STAMPANTI:

EPSON



- **LX 86 F/T** - 80 colonne, 120 CPS, grafica, parallela. L. 580.000



- **FX 800** - 80 colonne, 200 CPS, grafica, parallela. L. 750.000
- **EX 800** - 80 colonne, 250 CPS, grafica, parallela. L. 970.000



- **FX 1000** - 136 colonne, 200 CPS, grafica, parallela. L. 850.000
- **EX 1000** - 136 colonne, 250 CPS, grafica, parallela. L. 1.390.000



- **LQ 800 F/T** - 80 colonne, 180 CPS, grafica, parallela, 60 CPS NLQ L. 1.250.000
- **LQ 1000 F/T** - 136 colonne, 180 CPS, grafica, parallela L. 1.480.000
- **LQ 2500 F/T** - 136 colonne, 270 CPS, grafica, parallela L. 1.840.000

ACCESSORI PER STAMPANTI EPSON

- **KIT COLORE** per serie EX-800 / EX-1000 L. 150.000
- **KIT COLORE** per serie LQ-2500 L. 150.000
- **INTERFACCIA SERIALE** per serie FX-800 / FX-1000 L. 150.000
- **INSERITORE** fogli singoli per LX-86 L. 210.000
- **INSERITORE** fogli singoli per FX-800 L. 380.000
- **INSERITORE** fogli singoli per FX-1000 L. 470.000
- **INSERITORE** fogli singoli per EX-800 L. 380.000
- **INSERITORE** fogli singoli per EX-1000 L. 470.000

NASTRI ORIGINALI PER STAMPANTI

- **EPSON LX-86** L. 10.900
- **EPSON FX-800** L. 8.000
- **EPSON FX-1000** L. 10.500
- **EPSON FX-85** L. 8.000
- **EPSON FX-105** L. 10.500
- **EPSON EX-800** L. 18.500
- **EPSON EX-1000** L. 18.500
- **EPSON LQ-800** L. 16.000
- **EPSON LQ-2500** L. 24.000
- **EPSON LQ-1500** L. 11.000

confezioni da 12 pezzi

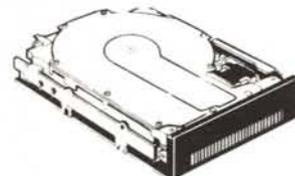
- **TALLY MT 130/140/180/290** L. 11.800
- **TALLY MT 80** L. 10.500
- **COMMODORE 803** L. 10.500

IBM COMPATIBILI



- **F2 - XT 512K**, 2 FDD 360K, TASTIERA, MONITOR PHILIPS L. 1.690.000
- **F20 - XT 512K**, 1 FDD 360K, 1 HD 20 Mb, TASTIERA, MONITOR PHILIPS L. 2.490.000

HARD-DISK



MONTAGGIO INTERNO PER PC/XT: •

- KIT 10 Mb SLIM L. 850.000
- KIT 20 Mb SLIM L. 980.000
- KIT 30 Mb SLIM L. 2.340.000

MONTAGGIO ESTERNO PER PC/XT: •

- KIT 10 Mb L. 1.450.000
- KIT 20 Mb L. 1.570.000
- KIT 33 Mb L. 2.850.000

MONTAGGIO INTERNO PER AT: •

- KIT 20 Mb L. 1.460.000
- KIT 30 Mb L. 1.725.000
- KIT 40 Mb L. 1.950.000

CARD-DISK

- **MINISCRIBE 20 Mb** L. 1.390.000

GARANZIA 12 MESI



DESIDERO RICEVERE GRATUITAMENTE
IL VOSTRO CATALOGO COMPLETO

Cognome

Nome

Indirizzo

Professione

Firma

ASSISTENZA TECNICA SPECIALIZZATA • VENDITA RATEALE O LEASING • VENDITA PER CORRISPONDENZA

CONDIZIONI DI VENDITA: Il pagamento potrà essere effettuato in forma anticipata a mezzo vaglia telegrafico o assegno circolare o in contrassegno tramite posta o corriere. Le spese sono a carico del destinatario.
La spedizione è prevista entro 15 gg.

I PREZZI SONO IVA ESCLUSA