

La specializzazione della conoscenza

I problemi messi in evidenza dalle tematiche e dalle tecniche di riconoscimento finora espresse hanno toccato problematiche piuttosto semplificate (riconoscimento di strutture poliedriche semplici); ma siamo ancora lontani da realtà applicative, come il riconoscimento di immagini non regolari. Ad esempio, come fa HAL a leggere sulle labbra di Bowman la sua condanna ed a prendere le relative contromisure? Dura vita per i robot, se non ci fosse stato un umano che li ha dotati del senso (anzi del programma) della vista! Certo è che prima di imparare a leggere sulle labbra, occorre riconoscere le labbra stesse, anche come facenti parte di una testa, una faccia. Stranamente (o forse neanche tanto) tecniche di riconoscimento della figura umana e della sua faccia non furono, come si può credere, affrontate come logico sviluppo delle tecniche di riconoscimento elementare finora descritte: probabilmente il fascino dell'idea di poter dotare una macchina della capacità di riconoscere (ed ubbidire?) un uomo fu tale che algoritmi di riconoscimento del viso umano furono approntati pressoché parallelamente a quelli finora analizzati finalizzati soprattutto ad oggetti; vediamo qualcosa sull'argomento!

I programmi finora nominati sono, come si può immaginare, del tutto didattici e speculativi, senza alcun aggancio con la realtà, come non esiste d'altro canto realtà formata da cubi, parallelepipedi, e spigoli illuminati secondo le nostre esigenze particolari. Ma un programma davvero utile deve avere necessità di individuare oggetti non standard (ve lo immaginate come è complesso riconoscere un foglio di carta appallottolato?), e deve possedere requisiti particolari di conoscenza, specialistici per ciascun oggetto che desidera riconoscere nel mondo (dominio) oggetto di riconoscimento; non solo, ma deve essere in grado di selezionare i processi euristici destinati al riconoscimento dell'oggetto stesso. Selezionare questi processi correttamente non è generalmente semplice se si considera, tra l'altro, che spesso, nel problema della conoscenza, interviene la lingua parlata, molto più ambigua della visione (Abelson, nel suo «About structure of knowledge» cita un esempio assurdo circa una macchina che, ipoteticamente interrogata sul muro di Berlino, cercherebbe informazioni dettagliate circa la chimica dei laterizi).

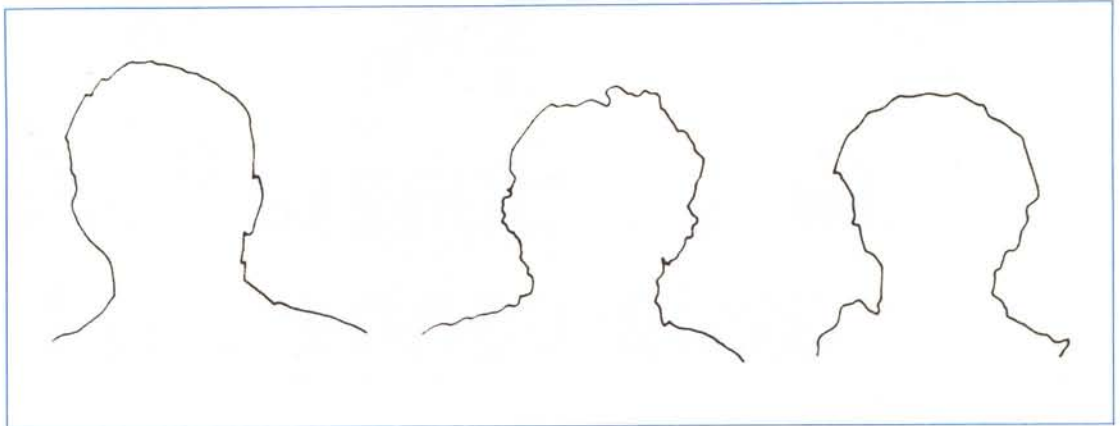
D'altro canto abbiamo visto, nelle puntate precedenti, che quanto è buono ed utile una volta (come ad esempio una forte illuminazione degli oggetti) potrebbe non esserlo in altre circostanze, così come, nel caso precedente, la conoscenza della chimica dei mattoni può o non può essere utile. Il problema e la soluzione, come al solito, stanno nel mezzo, nell'uso, cioè «intelligente» (non a caso) dei particolari in possesso per decidere i processi euristici da adottare, non dimenticando mai quello di cui si ha bisogno.

Humpty Dumpty, nell'incantato suo mondo, confida ad Alice di non saperla riconoscere, ed anzi la prega di indicargli qualcosa di effettivo per consentirgli una differenziazione da altre persone, dicendole: «Sei così eguale agli altri: la tua faccia è la stessa di quella degli altri, hai due occhi, con un naso nel mezzo, ed una bocca al di sotto. È sempre così; se tu avessi gli occhi allineati col naso, o che so, la bocca sopra di essi, questo mi potrebbe essere di molto aiuto».

Sebbene il discorso di Humpty Dumpty sia assurdo, e notevolmente semplicistico, esso già presume un elevatissimo livello della conoscenza, come riconoscimento di occhi, bocca, naso e, cosa ben più importante, la loro implicita connessione e conoscenza delle relazioni che tra questi «oggetti», elementi, intercorrono. Un programma destinato al riconoscimento delle persone deve essere così ben articolato da aver non solo conoscenza di visi (riconoscendo, cioè le interconnessioni e le relazioni, abbastanza univoche, che intercorrono tra diversi oggetti; in altre parole, più semplici, di come un viso si presenta effettivamente), ma anche di avere la possibilità in questi visi, che sono formati tutti degli stessi «pezzi», di distinguere anche le particolarità che permettono ad HAL di riconoscere Alice o Martin Bowman.

A questo scopo tenta di soddisfare un programma di M.D. Kelly («Edge Detection in Pictures by Computer Using Planning». MI-6, pagg. 397-410). Il programma usa un algoritmo particolare rappresentato da un cercatore di linee sensibile al profilo della testa di una persona. La prima fase di riconoscimento è rappresentata dalla estrazione,

Figura A
Da Bernard Meitner
e Donald Michie,
«Machine Intelligen-
ce», 6,
Edinburgh University
Press, Edimbourg,
1971.



dall'immagine, di un accurato profilo della testa della persona osservata, ottenuta da una immagine della persona stessa su diversi sfondi. L'output del programma è costituito da linee (fig. A) rappresentanti il profilo della testa, nudo e crudo, ignorando del tutto particolari dello sfondo o della testa stessa.

L'algoritmo di analisi della immagine è simile a quello del programma di Shirai, e si dimostra discretamente «intelligente» se si considera che il riconoscimento avviene attraverso una opportuna pianificazione della lettura dell'immagine stessa. All'inizio viene prodotta una più piccola immagine per riduzione dall'originale; questa porta, ovviamente, ad una perdita voluta di molti particolari poco significativi senza sacrificare eccessivamente la leggibilità dell'immagine stessa. Dopo di ciò si tenta di individuare su questa figura un profilo di base, che viene poi utilizzato per pianificare la ricerca di particolari sulla immagine principale, più grande, in base ad un semplice principio: le linee presenti della figura di output vengono riferite a linee simili sulla immagine principale; se c'è concordanza (vale a dire che, sulla immagine principale, non ci sono discontinuità che poi sono sparite sulla rappresentazione ridotta, e che rappresenterebbero, a tutti gli effetti, parti diverse della figura) fra le due immagini, il programma prosegue nella sua ricerca, altrimenti torna indietro per un successivo tentativo.

Il programma segue, nella sua ricerca, solo linee sottili e possiede inoltre tecniche di controllo della «qualità» del-

le linee seguite; tanto per intenderci, se la persona di cui si legge l'immagine sta sullo sfondo di una porta od una finestra, il programma, pur analizzando il profilo della porta stessa, ne scarta immediatamente la sua validità e compatibilità con l'immagine principale, in quanto è stato istruito a comprendere che linee rette così lunghe hanno poco a che vedere con un profilo umano. A questo punto il programma esegue un backtracking, rendendosi conto che è su una pista sbagliata; ciononostante lo stesso programma è abile a riconoscere irregolarità nella forma principale, come ad esempio il contorno dei capelli.

Come abbiamo già detto, il programma di Kelly non andò oltre il riconoscimento del contorno della testa umana (ad esempio, il programma non è capace di leggere il profilo di un leone o di un elefante, perché il programma stesso non ha alternative per assegnare il profilo che sta leggendo a teste diverse da quella umana). Così, ad esempio, l'algoritmo utilizzato non sarebbe capace di interpretare come umano il profilo del diavolo, provvisto di piccole corna, e viceversa ne riconoscerebbe l'appartenenza al nostro genere in mancanza, così come accade con Louis Cheepher nel bel film di Alan Parker.

Come accadeva col programma di riconoscimento dei poliedri messo a punto da Shirai, i processi di individuazione ed accettazione delle linee si basano su tecniche procedurali inserite nel programma stesso. Ad esempio, se, in una fase di ricostruzione della immagine viene letta una linea orizzontale, al

di sopra di un'altra che è già stata riconosciuta come la sommità della testa il programma scarta la linea stessa, non senza aver prima tentato di riconoscere la linea come una irregolarità dei capelli (viene, in effetti verificato che questa linea non sia collegata, attraverso una curva, al profilo della testa stessa, come potrebbe accadere con una ciocca di capelli disordinata sulla testa di una donna, l'esempio è importante in quanto evidenzia come, attraverso la scalatura della immagine, spariscono le particolarità come i singoli capelli al di sopra del profilo generale, cosa sempre possibile). È attraverso una tecnica di questo genere che Satana non viene riconosciuto, quando provvisto di corna, perché, al massimo, potrà accadere che, in condizioni di illuminazione cattiva, il profilo delle corna sarà confuso con quello dei capelli, specie se Satana si sarà adeguato alla moda dei capelli lunghi.

Il primo passo compiuto dal programma di Kelly nel creare la fotografia in piccola scala è la ricerca di tre linee principali, che individuano la sommità e le parti laterali della testa, facendo attenzione, ovviamente, a che la linea orizzontale sia superiore a quelle verticali (ad esempio, non verrebbe letto il viso di una persona sospesa all'ingiù ad un trapezio). Il programma tenta quindi di connettere insieme le tre linee per creare un profilo di una testa, ancorché appena abbozzato, verificando se la linea orizzontale possa, sull'originale, essere assimilata ad una curva rappresentante la sommità della testa stessa di un uomo. Una volta che è stata riconosciuta

Figura B

$$4 \frac{1}{2} + 3 \frac{3}{4} = 5 \frac{1}{4}$$

$$4 \frac{1}{2} + 3 \frac{3}{4} = 5 \frac{1}{4}$$

la metà superiore della testa stessa, il programma inizia la sua analisi delle aree inferiori, alla ricerca di due curve concave (rappresentanti i lati del collo) seguite da due curve convesse (le spalle).

Ovviamente a questa ricerca segue quella dei particolari più piccoli, come il naso, la bocca, gli occhi, il profilo delle orecchie, ecc. Ciononostante siamo ancora lontani dalla possibilità di riconoscimento delle persone, anche perché non è certo possibile dire al programma: «Per riconoscere Alice, sappi che possiede occhi tirati all'insù, un naso sottile, una bocca impertinente come un bocciolo di rosa». Inoltre non è possibile costruire un programma che, accanto al riconoscimento di una persona, sia abile nel riconoscere profili di porte, sedie, finestre, presenti sullo sfondo. La ragione di ciò è rappresentata dal fatto che non esiste, purtroppo, un linguaggio generale destinato a descrivere la struttura di un disegno potente e rapido abbastanza da comunicare alla macchina i più accurati e particolari concetti della figura stessa. Il problema della interpretazio-

ne del significato e del riconoscimento di un profilo come quello di una determinata persona fu, almeno in linea teorica affrontato fin dal 1960 da M.L. Minsky («Steps toward Artificial Intelligence», CT, pag. 422, e, ancora dello stesso autore «A framework of representing knowledge» IJCAI - 4, pag. 211; un successivo, più approfondito tentativo di applicare il principio di Minsky fu attuato da T.G. Evans, «A program for the solution of geometric-analogy Intelligence Test Question») i cui lavori sono rimasti, per lungo tempo, una delle pietre miliari del campo della visione attraverso la macchina. Minsky affrontò in maniera formale il concetto di «sopra e sotto» nella visione, applicando le sue esperienze, però, non al riconoscimento di persone ma a quello degli oggetti. Ad esempio, la figura B mostra un classico esempio (estratto peraltro dalla bibliografia citata), in cui il concetto di «sopra» e «sotto» è fondamentale per la lettura della prima delle tre frazioni dove l'ambiguità dello scritto può essere risolta da una persona, o da un program-

ma a conoscenza delle regole matematiche, ma non da un lettore generico di caratteri, che leggerà, nella figura a sinistra, più probabilmente 412 o 4'2 o qualcos'altro di incomprensibile (per la corretta interpretazione dell'esempio si ricordi che nella comune notazione matematica americana è ammesso l'uso di frazioni senza il numeratore, che viene sottinteso pari ad 1; n.d.a.).

A questo punto appare chiara una cosa; che i programmi di riconoscimento non possono essere (come con humor dice lo stesso D.H. Waltz (D.L. Waltz, «Understanding Line Drawing of Scene with Shadow»)) solo «iconici», vale a dire legati alla assimilazione ed al confronto di immagini con altre precostituite, ma devono essere «semantici», capaci di eseguire analogie al di fuori di regole precostituite, devono cioè essere «intelligenti» (altrimenti a che pro il titolo della nostra rubrica). Come fare? Occorre riunire, come sempre, il lavoro di molti e di essi trarre il meglio per costruire una base di supporto a nuove tecniche. A risentirci!

MC



NEWEL srl
hardware software telematica

20155 MILANO - Via Mac Mahon, 75
tel. 02/32.34.92 - tel. 02/32.70.226

Newel è

**COMMODORE POINT
LIBRERIA JACKSON
RIVENDITORE AMSTRAD
AMIGA SPECIALIST
SOFT CENTER
RIVENDITORE LAGO
STAR
MANNESMAN
ACORN (NEW) ARCHIMEDES
ATARI DISTRIBUTOR
OC 118 COMPATIBLE DRIVE 64/14
MODEMPHONE
LIBRERIA TECNICA
OLIVETTI PC
PRODEST & SOFT
J - SOFT RIVENDITORE**

**NASHUA
GMC
3 M
POSSO BOXES
DURACELL
SUPERPILA
SONY AUDIO
PDM AUDIO
TASKER CAVETTERIA
PC COMPATIBILE DE ALER
ACCESSORI E SOFT
RIPARATORE, di COMMODORE, SINCLAIR,
PC, MONITOR, MODEMS, STAMPANTI ETC.
PRODUTTORE di SOFTWARE IN ITALIANO
*e tanto tanto altro....***

**NEGOZIO AL PUBBLICO
E VENDITA PER CORRISPONDENZA**

CASH & CARRY

NUOVO ORARIO 1988

DAL MARTEDI
AL VENERDI
9-1230 / 15-19

SABATO
10 - 13 / 14 - 18

LUNEDI CHIUSO