

La visione secondo Shirai

Per la prima volta il programma di Shirai utilizza un approccio intelligente in presenza di una interpretazione dubbia o di un tentativo di collegamento tra linee presumibilmente confluenti: in termini spicci, il programma usa una parziale interpretazione basata sulla più ovvia risoluzione grafica della scena da completare. Ma non basta: man mano che la interpretazione avanza senza intoppi, il procedimento di riconoscimento diviene sempre più efficiente, visto che il programma, basato su un efficiente sistema esperto, fa tesoro delle conoscenze già acquisite, e diviene sempre più veloce e potente nel suo sviluppo col passare del tempo (scherzosamente, Shirai ebbe a dire, non molto lungi dalla verità, che il suo programma lavorava alla immagine come ad un cruciverba, la cui soluzione diveniva sempre più facile, avanzando nello sviluppo stesso). Il risultato è che il programma passa da un semplice lettore di linee iniziale ad un vero interprete di immagini, con una distinzione non ben netta tra le due fasi. Conseguentemente, la ricerca finale delle linee può essere orientata non solo ad una pedissequa interpretazione in bidimensionale, ma esistono tutti i presupposti per una efficace interpretazione tridimensionale della scena. Il programma basa la sua tecnica su una sofisticata ricerca della più probabile linea accettabile dal contesto, attivando, tra l'altro un efficace algoritmo di ricerca delle linee nascoste, cosa ignota ai programmi precedenti. Questa tecnica può rivelarsi, alla fine, erronea, in quanto capita spesso che il programma immagini linee senza verificare se queste abbiano senso con linee precedentemente localizzate.

Il programma di Shirai si dimostra particolarmente efficiente nella ricerca delle linee di contorno, le linee, cioè, rappresentanti gli spigoli del corpo separanti il corpo (bianco) dell'oggetto dal-

Poiché, come si è visto precedentemente, il problema della interpretazione degli oggetti poggia, in maniera estesa, sulla corretta individuazione e comprensione di tutte le linee visibili appartenenti al corpo stesso, e sulla individuazione univoca della corrispondenza linea oggetto, un nuovo programma vede la luce nel 1972 ad opera di Y. Shirai (Yoshiaki Shirai, «A context Sensitive Line Finder for recognition of Polyhedra», AI 4 (1973), altrimenti ripubblicato con numerose modifiche in PVC, 1974, pp. 93, 134, sotto il titolo «Analizing Intensity Array Using Knowledge About Scenes»).

lo sfondo (nero, o comunque più scuro). Secondo il programma ancora, tutti i successivi gradienti di colore sono «frontiere» separanti un corpo da un altro, e le frontiere tra superfici più chiare sono le «linee interne», separanti una faccia dall'altra nello stesso corpo. L'algoritmo di ricerca e lettura delle immagini cerca di individuare prima il contorno dell'oggetto, successivamente le frontiere, ed infine le linee interne, secondo un procedimento di ricerca euristica che possiede come base un po' tutte le tecniche dei programmi precedentemente descritti, ma che appare fortemente integrato da un processo di «fusione» intelligente delle immagini.

Per cercare le linee di contorno (le più semplici da individuare) il programma esegue una scansione dell'immagine alla ricerca di punti ad elevato contrasto, e li collega tra loro con linee verificate come linee di spigolo. Dopo di che esegue una scansione in piccola scala di questi punti, con una semplificazione elevata del problema in quanto vengono esaminate aree dell'immagine non più ampie di una array di 8x8 elementi (in un secondo tempo Shirai passò ad analizzare array di 12x12 e 16x16 punti, senza peraltro ottenere risultati di miglior valore). Utilizzando la rappresentazione ottenuta da questa interpretazione semplificata, il programma ritorna all'immagine iniziale e, ormai conscio dei punti da osservare, concentra la sua attenzione su di essi per tracciare un contorno più accurato dell'immagine. È un procedimento molto intuitivo, che parte da una individuazione molto semplicistica della scena per passare a fasi sempre più avanzate e complesse.

Ovviamente, questo processo di ricerca e verifica di contorni dipende dalla affidabilità del contrasto tra superfici chiare e scure, vale a dire, come avevamo fatto notare in premessa, tra oggetti e sfondo. Ma, a differenza di quanto

accadeva con i programmi precedenti, questa ricerca viene considerata solo come una sgrossatura del problema, vale a dire viene utilizzata per indicare al programma «dove guardare». Il procedimento di individuazione della immagine, che per i programmi precedenti talora finiva qui, per Shirai è appena cominciato.

Consideriamo la figura A: il primo passo del programma è quello di individuare le «frontiere dell'oggetto con l'esterno»; questo porta, ovviamente (cosa che era negata ai programmi precedenti), a riconoscere anche punti come [b] situati in una concavità della immagine. E l'utilità di riconoscere le concavità è rappresentata non solo dal fatto in sé, ma anche dalla considerazione che esse possono costituire la linea di giunzione di due oggetti appoggiati lungo una faccia tra di loro. Questo, in altri termini, porta alla considerazione che linee situate in corrispondenza di superfici concave possono essere considerate come linee di frontiera tra differenti oggetti. Poiché le linee di contorno individuano spigoli del corpo, sono necessariamente linee di frontiera, o solamente segmenti. L'algoritmo di riconoscimento prosegue quindi in questo modo: le linee [ab] e [pb] vengono ambedue analizzate per stabilire cosa rappresentano nel corpo stesso (si noti come [ab] termini in una «forcina», da tutti i programmi finora descritti considerata come rappresentativa di limiti di corpi diversi). A questo punto il primo contorno è individuato, ed il programma esegue una ricerca circolare attorno al punto [c] per individuare una linea per chiudere la figura. Individua, così, la linea [cd], interpretata immediatamente come una linea di frontiera; ancora una ricerca in circolo, e si arriva a [d], e successivamente, ad [e]. Ma qui arriva il bello! [cde] pare proprio la giunzione di una forcina e d'altro canto non esiste

figura che possa essere rappresentativa di un corpo con la sola forma [cde]. Deve trattarsi, allora, certamente di una «fork», e da qui parte la ricerca della linea [df], che finirebbe nel nulla se in [f] non ci fosse, ancora, qui, una «fork» che porta da una parte in [g], dall'altra in [s], di nuovo, ambedue punti di frontiera del corpo stesso. Tutto il corpo resta così individuato, soprattutto nella non facile individuazione delle frontiere tra i due corpi lungo una superficie [bcdeg].

Questo esempio è ben piccola cosa in relazione a quanto effettivamente può fare il programma di Shirai; nel nostro metodo abbiamo nominato solo tre dei dieci procedimenti euristici su cui si basa il programma (ricerca della linea di frontiera, analisi in circolo, analisi in flashback dei risultati). Ma si è già visto come siano apparse le prime forme di ricerca intelligente, generalizzate, nel nostro esempio, dalla ricerca circolare, attorno ad un punto, di linee di frontiera; e così come l'uso di questo o quest'altro algoritmo varia, intelligente-

mente, man mano che l'interpretazione va avanti e si evolve, allo stesso modo i risultati sono continuamente testati alla ricerca della verifica tra essi e la conoscenza più generale maturata precedentemente; il risultato più appariscente di tutto questo va e vieni è rappresentato dal fatto che con questo sistema vengono immediatamente scartati i particolari poco significativi che possono ingenerare confusione o quelli imperfetti nell'input.

In questo modo, se viene individuato un frammento di linea (vale a dire un segmento che, in una sua parte, non risulta collegato con altre parti significative) il programma tenta di interpretare se questo frammento, idealmente prolungato, può collegarsi con un altro frammento, si da rappresentare un unico logico (in questa occasione l'algoritmo somiglia molto al modulo COMPLETE del programma di Falk, anche se risulta di questo molto più potente e più generalizzato nelle applicazioni). Alcune volte l'algoritmo, in assenza di risultato,

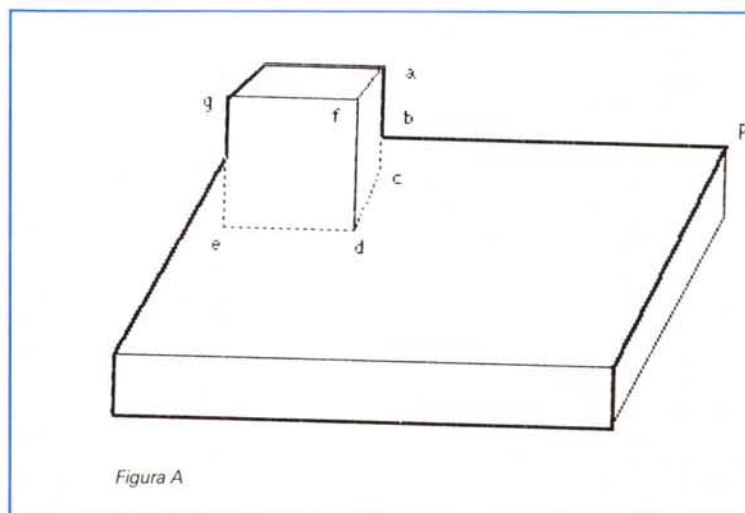


Figura A

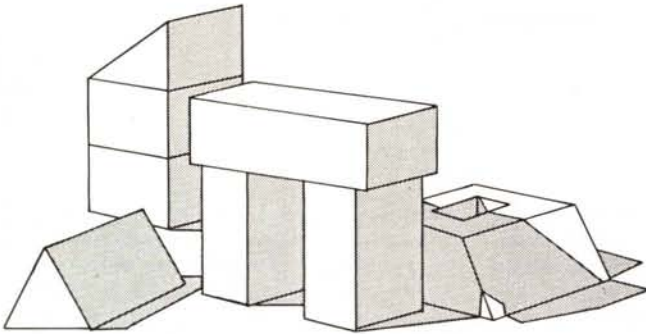


Figura B
Da P.H. Winston,
«The psychology of
Computer Vision»,
Mac Graw-Hill, 1975.

si modifica autonomamente in modo da essere più «sensibile» così da poter interpretare anche linee di frontiera tra superfici a tenue variazione di colore come separazioni tra corpo e corpo e tra corpo e background. Allo stesso modo può essere più o meno resa sensibile la ricerca circolare attorno ad un punto.

Un ulteriore vantaggio del programma di Shirai è quello che esso ignora le linee irrilevanti, evitando (v. figg. B e C) che linee di penombra siano considerate come separazioni tra corpi o facce, in questo essendo molto più efficiente del più evoluto programma visto in precedenza, quello di Falk, ancora una volta, il cui modulo INTERPRET riesce sì a completare le linee frammentarie o scansite male, ma non può ignorare linee lette, a cui, in ogni caso «deve» dare una interpretazione in termini di spigoli o frontiere tra oggetti diversi. Ancora una volta, in maniera molto spiritosa Margaret Boden, nel suo volume «Artificial intelligence and natural man», opera già citata diverse volte e che è fonte inesauribile di spunti di questi articoli, paragona questa facoltà del programma alla capacità dell'uomo di separare le parole della conversazione col suo interlocutore dal rumore di fondo della musica e dal tintinnio dei bicchieri in un party. Tutto il superfluo nella figura viene definito da Shirai come «noise», rumore di fondo, fruscio, in stretta analogia con l'equivalente sonoro; Shirai, pur ammettendo

che una troppa spinta tecnica di utilizzo del taglio del «rumore» può nuocere alla capacità del programma di riconoscere le immagini, o, almeno, di dare a loro senso compiuto, risolve, ancora una volta il problema con la sua tecnica di va e vieni; se il «taglio» è stato troppo spinto, allora si procede all'inverso, tentando un recupero di linee precedentemente ritenute superflue e scartate, in modo da tentare la ricostruzione dell'immagine, oltre tutto in termini intelligenti.

Si comincia, quindi, a vedere la trama del programma di Shirai, e le sue limitazioni; se, per continuare con l'analogia col party, il nostro interlocutore ha una voce flebile o la musica è troppo alta, anche l'eccezionale capacità di separazione del nostro cervello ha scarso effetto; nello stesso modo il programma di Shirai ha, sotto questo punto di vista, limiti non facilmente valicabili. A questo problema pone rimedio un analizzatore di linee molto avanzato, realizzato da G.R. Grape (G.R. Grape, «Computer Vision Through Sequential Abstraction», Stanford, 1969), in connessione con il gruppo di lavoro intorno alla realizzazione di un occhio artificiale alla Università di Stanford; il principio, che illustreremo più precisamente nella prossima puntata, utilizza modelli astratti di schemi preordinati (ancora una volta si lavora solo su poliedri e poligoni convessi), per cercare di trarre un senso da una serie caotica di linee in input (si noti come il programma sia precedente come realizzazione al modello di Shirai, che, comunque, pare che sia stato al corrente degli studi del collega). Il procedimento, in sé, non possiede nulla di nuovo se non nel fatto che applica, ancora una volta, certe tecniche di analogia già viste in precedenza, ma con un approccio più raffinato ed efficiente.

MC

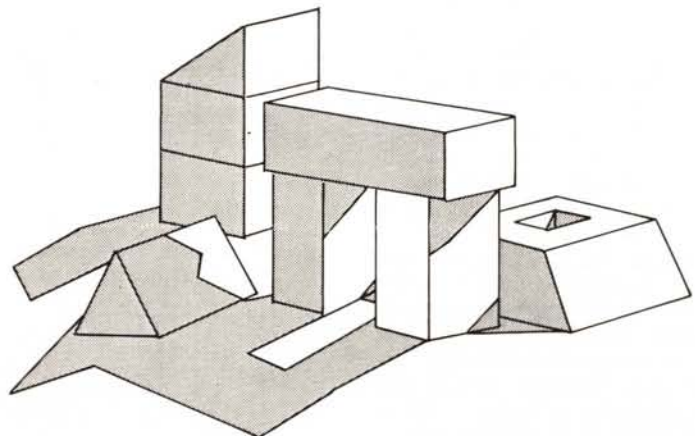
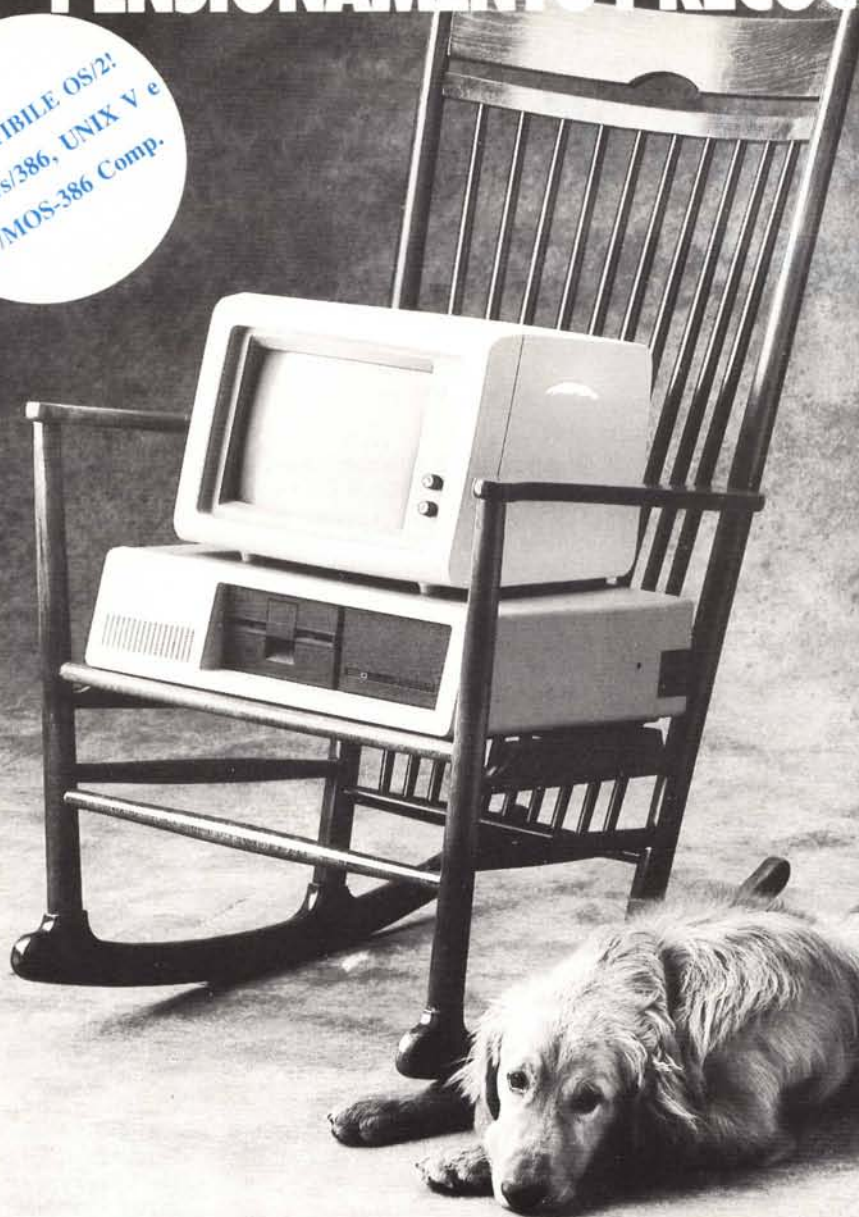


Figura C - Ibid. come in fig. B.

SALVA IL TUO PC DA UN PENSIONAMENTO PRECOGE.

COMPATIBILE OS/2!
Windows/386, UNIX V e
PC/MOS-386 Comp.



ORDINA OGGI LA TUA NUOVA MOTHERBOARD BRAINSTORM 386 ST

80386 ST/MB - LIT. 3.800.000

Dai al tuo PC nuova linfa vitale! Con la nostra motherboard 386 il PC, PXC/XT o compatibile sarà un degno rivale in velocità dei migliori sistemi 386 in circolazione. Addirittura più veloce. Questo per il Megabyte di RAM ad alta velocità e lo zoccolo per il coprocessore 80387 per velocità esplosive mai raggiunte sinora. Per lasciare il pensionamento fuori della porta 80386 ST/MB è compatibile con il PC/AT (BIOS e I/O) e vi permette di usare la nuova generazione di DOS, l'OS/2 e il PC/MOS 386. Abbiamo anche incluso due slot di espansione a 16 bit per le più recenti schede di espansione. Nessuna scheda acceleratrice potrebbe darti tanta versalità.

Hauppauge!



Con la potenza del 386 e vera compatibilità software AT, il tuo lavoro, il Desktop Publishing e le tue applicazioni ingegneristiche avranno una sferzata di nuova produttività.

Specifiche Tecniche:

- 16 Mhz 80386 - 1 Megabyte di interleaved RAM a 100 nsec
- I/O e BIOS compatibile AT per il supporto dell'OS/2 - sette slot espansione a 8 bit - due slot espansione a 16 bit - uno slot di espansione RAM a 32 bit (max 12 megabyte) - coprocessore matematico 80387 opzionale.

Per maggiori dettagli e informazioni chiamaci oggi!

 **gesin trade**

Importato e distribuito in Italia da: **GESIN TRADE srl - Via Augusto Valenziani, 5
00187 Roma - Tel. 06/486459/4752792**