

Computer Vision e Astrazioni sequenziali

Come dicevano nella puntata scorsa, un più efficace programma di analisi di linee fu messo a punto da G.R. Grape (G.R. Grape, Computer Vision through Sequential Abstraction, A.I. Note, Stanford University, Stanford, 1969) verso la fine degli anni Sessanta; a quanto pare l'autore fu solo parzialmente a conoscenza dei risultati dei suoi contemporanei, e giunse ai suoi risultati partendo da un progetto pratico, vale a dire quello di un occhio artificiale, condotto dalla università di Stanford. L'approccio al problema è lievemente diverso (anche se, ancora una volta, vengono affrontati poliedri convessi) e si basa sulla assimilazione di quanto visto ai modelli astratti, anche essi solidi (non viene, perciò più eseguita la scomposizione del modello in parti e formati elementari). Si tratta, quindi, di un gradino un po' più alto rispetto a quelli precedenti, ed il progetto, nell'insieme, può già essere inteso come un vero e proprio sistema di visione

Il sistema di Grape parte da una immagine televisiva digitalizzata, utilizzando, quindi, un approccio molto simile a quello di Shirai. Questo approccio porta, come abbiamo già visto diverse volte, ad una iniziale confusione (e sovrabbondanza) di dati disponibili, rappresentati in figura A. La presenza di un maggiore o minore numero di dati (linee) è dovuta, comunque, anche alla maggiore o minore intensità di illuminazione degli oggetti e della scena, oltre che, ovviamente, alla sensibilità del digitalizzatore-camera. È ovvio che sta proprio nell'approccio la maggiore possibilità di errore. INTERPRET, lo ricordate, non potrebbe, dalla figura A, trarre molto.

L'originalità del sistema di Grape sta nel fatto che esso, anche utilizzando tecniche proprie dei programmi di Roberts e Falk, esegue, tout court, collegamenti tra linee spezzate senza tener conto di linee nascoste. Il risultato è quello rappresentato in figura B, che rappresenta la prima interpretazione di quanto è stato digitalizzato e rappresentato in figura A.

Il vero problema per cui il pur efficiente algoritmo di ricerca delle linee di Shirai incontra difficoltà nel sistema di Grape sta nel fatto che esso non riesce, nel caso di prolungamenti e connessioni di linee, a stabilire quali di queste linee siano da considerare locali (specifiche, cioè dell'oggetto in primo piano) o globali. Tutto ciò in quanto la decisione di scelta, effettuata dalla mente umana che guarda ed analizza la figura B, parte, pur sempre da processi euristici coinvolgenti una generale conoscenza dei corpi poliedrici e della loro visione prospettica nello spazio. In questo caso il sistema di Shirai mostra i suoi limiti, in quanto segue certe sue prerogative essenziali per giungere a determinati risultati. Ad esempio, nel caso della forcina della figura C, Shirai collega immediatamente il segmento centrale con il vertice, senza poi preoccuparsi molto di dove andrà a parare l'altra parte del segmento. Il procedimento di Shirai non implica un «guardarsi intorno» in modo

da cercare, intelligentemente, una interpretazione alternativa a questa linea. Se la linea non «finisce» da nessuna parte, il programma tenta un'ultima soluzione disperata; esso esegue una ricerca circolare attorno all'estremo «orfano» del segmento, ed accresce il raggio del cerchio sempre più fino a trovare «qualcosa di soddisfacente». Non esiste, invece, l'idea di cercare prima da qualche altra parte, per cercare di collegare insieme frammenti di linea già trovati.

L'algoritmo di Shirai presenta ulteriori complicazioni, ma mano che l'input diviene sempre più imperfetto. Ad esempio, le due aree individuate dal cerchietto della figura D vengono interpretate, da questo programma, come una regione triangolare, e come un vertice ad L, ancorché imperfetti (cose, ambedue, che non sono). La cosa, giustificabile tenendo conto dei metodi di analisi del programma, è del tutto assurda, ed è giustificata appena dalla considerazione che solo una analisi generale della figura potrebbe dare un senso alla scena, ed assegnare le linee in questione a specifici oggetti rappresentati nella scena.

L'originalità e l'efficienza nel programma di Grape sono affidate alla analisi della evidenza delle linee che, per quanto ambiguo sia il contesto, possiedono un inequivocabile senso, in accordo a più globali considerazioni. Tanto per fare un esempio, sempre nella figura D, il sistema di Grape individua correttamente il significato dei vertici compresi nei circoli, e li interpreta come appartenenti a corpi diversi, non solo, ma riesce ad interpretare come materiale di scarto parti non significative per la struttura degli oggetti principali, tutto ciò è possibile in quanto il sistema «legge» ed astrae dall'immagine modelli (cubi, parallelepipedi, prismi) solo completi o completabili.

Tutto ciò è possibile in quanto il sistema Grape possiede una gerarchia interna di rappresentazione e di analisi dei modelli. Ad esempio, sempre nella figura D consideriamo il contorno di

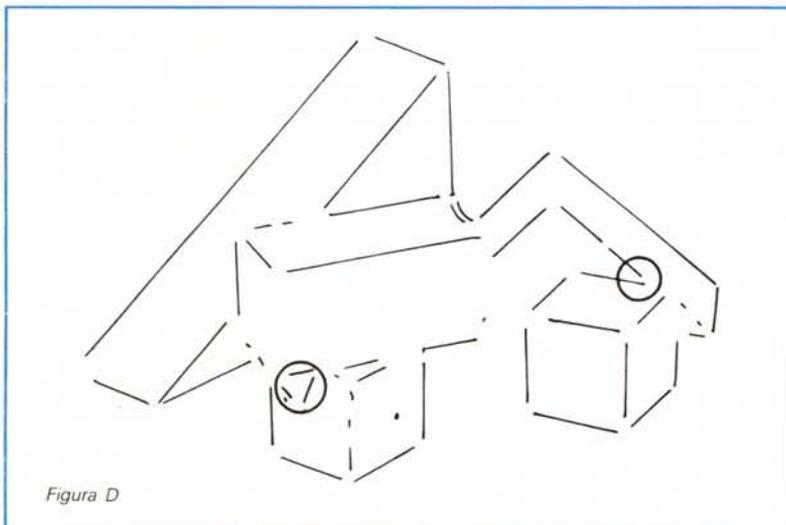
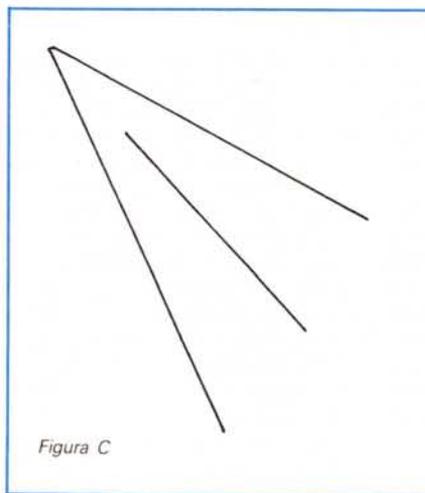
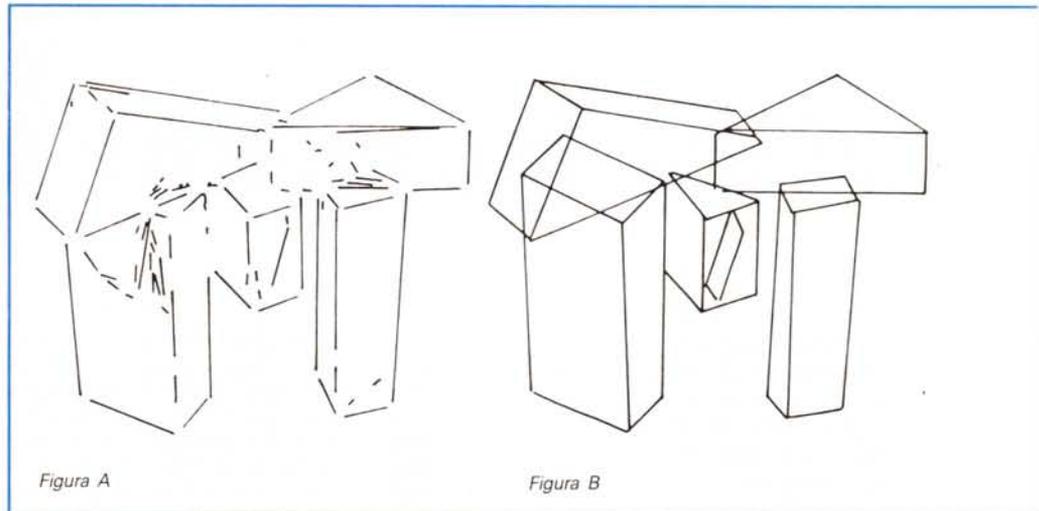
Tutte le figure del presente articolo sono tratte da «G.R. Grape, Computer Vision Through...».
Opera citata nel testo.

base del cubo in basso a destra della figura. Prolungando acconciamente le linee viene analizzato e rilevato un profilo a freccia che, a destra e sinistra, si collega a due profili ad L, che superiormente, confluiscono ancora in due frecce. Riunendo queste parti è possibile giungere ad un probabile modello, che suggerisce un modello a forma di cubo. A questo punto il sistema tenta un completamento della figura estraendo dalla rappresentazione tutte le linee significative e necessarie per il completamento del poliedro stesso.

Se, come in questo caso, l'analisi ha effetto positivo (in effetti viene utilizzata una semplice ma efficiente struttura top-down) tutti i vertici visibili dal centro vengono esclusi dalla successiva analisi.

Il sistema di partire da analisi a basso livello (linee, vertici) per giungere alla assimilazione con prototipi, parte a continue interazioni tra alti e bassi livelli d'analisi.

Tutto ciò è molto più elastico ed efficiente di quanto avevamo visto nei casi precedenti: ad esempio, l'unico vero sistema di analisi top-down analizzati nelle puntate precedenti, quello di



Folk, con i suoi moduli consequenziali, non consentiva analisi e verifiche all'indietro, ciononostante anche il sistema Grape presenta qualche problema, connesso con la sua esigenza di semplificazione dei problemi. Ad esempio la «smania» di semplificazione e di riduzione a modelli, propri di questo sistema, porta ad una non sempre accettabile semplificazione del modello di base, anche in funzione del fatto che la presenza delle linee nascoste a posteriori, ed anche, riflessioni, distorsioni prodotte dalla telecamera, portano a presenza di «rumore di fondo» che, sebbene il sistema ne preveda la pulizia, non riesce ad essere escluso del tutto quando, ad esempio, interferisce con entità nette e precise, come linee e vertici di poliedri. Il problema è sempre lì; solo in un mondo (reale ad immaginario), che offre una perfetta visuale, una visione perfetta è possibile. Poiché ciò si verifica ben di rado, è solo questione di «intelligenza» del mezzo che analizza se l'interpretazione è più o meno fedele all'originale; abbiamo così da una parte, il più semplice dei programmi di lettura delle immagini (Roberts?) e dall'altra la mente umana; che l'approccio e la soluzione di Grape siano più efficienti lo dimostra il fatto che la tecnica di interpretazione umana è più prossima al secondo polo che non al primo!

Viva la visione chiara e ben definita, quindi! Ma come la mettiamo se, come appare più che probabile, gli oggetti presenti nelle immagini possiedono delle ombre?.

Lo vedremo nella prossima puntata.