



LABORATORIO DI VISIONE ARTIFICIALE

Tel. 0382/391372

Telefax 0382/412881

Tecniche e Strumenti per l'Elaborazione di
Immagini ed il Riconoscimento di Forme

Parte II: Livello Alto e Strumenti Disponibili

9.

29 Novembre 1989
c/o SPERONI Spa, Spessa Po

Luca Lombardi:

Sequenza di immagini: stima del movimento.

Dipartimento di Informatica e Sistemistica, Università
degli Studi di Pavia

Stima del movimento

Il flusso ottico.

Corrispondenze discrete.

Sequenze di Immagini

La stima del movimento introduce nella visione artificiale una componente nuova ed essenziale

il tempo

L'evoluzione temporale dell'ambiente e della scena comporta un enorme incremento dei dati necessari per elaborare gli *eventi* che hanno luogo nel campo visivo.

L'evoluzione temporale sulla scena non dipende semplicemente dal moto degli oggetti: anche in una scena stazionaria si hanno dei cambiamenti causati dal moto relativo dei sensori (o dell'osservatore).

Anzi proprio il moto relativo può essere sfruttato per ricostruire la natura tridimensionale della scena osservata.

Una fonte di informazione importante è rappresentata anche dalla variazione dei toni di grigio.

Data l'enorme quantità di informazioni da trattare solo negli ultimi anni l'analisi di sequenze di immagini ha avuto uno sviluppo significativo.

Stima del Movimento

Il problema della ricostruzione della struttura volumetrica e del moto da viste 2D ha una grande importanza non solo per le difficoltà teoriche che pone, ma anche per la dimensione delle sue applicazioni pratiche. Dal punto di vista teorico, analizzare una sequenza pone difficoltà di tipo nuovo rispetto al trattamento di immagini statiche, poichè oltre ad estrarre le parti di interesse da ogni singola immagine occorre integrarle e interpretarle in maniera coerente. Dal punto di vista delle applicazioni, si possono citare da un lato quelle in cui il movimento e la struttura sono usati a scopo di interpretazione, descrizione o controllo e dall'altro quelle in cui l'obiettivo è la rappresentazione efficiente del contenuto informativo della sequenza.

Stima del Movimento

In generale, la stima del movimento (e della struttura) 3D a partire dalle sue proiezioni 2D su una sequenza di immagini comprende tecniche di diverso livello: da quello del trattamento dei dati di luminosità per ricavare informazione sul moto nel piano di formazione delle immagini, all'interpretazione di questo per ricavarne il moto 3D e la struttura degli oggetti, fino all'integrazione con modelli di scena e conoscenze a priori.

Stima del Movimento

Occorre notare che questo problema, schematizzabile come il tentativo di ottenere informazione 3D da proiezioni 2D, appartiene alla categoria dei problemi **inversi o mal condizionati**, per i quali cioè non è garantita l'esistenza e l'unicità della soluzione.

Ma soprattutto tale soluzione risulta molto "sensibile" ad errori ed approssimazioni nei dati di partenza: a causa del processo di inversione, variazioni anche molto piccole nei dati si ripercuotono amplificate in errori significativi nei risultati, se non vengono presi adeguati accorgimenti correttivi.

Stima del Movimento

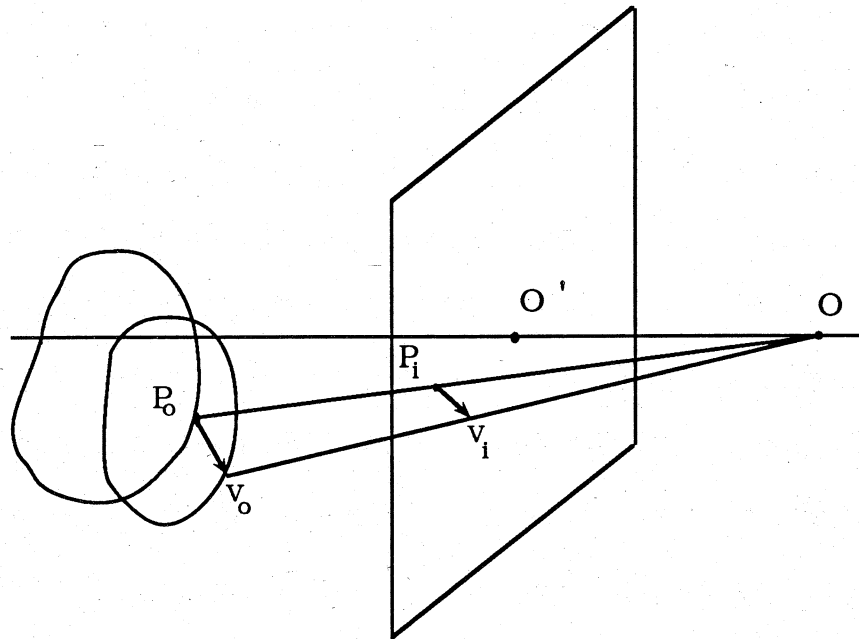
Si distinguono due classi di metodi fondamentali di stima del moto da una sequenza di viste.

Il movimento relativo tra gli oggetti della scena ed il sensore produce uno spostamento di parti delle immagini formate nel piano focale del sensore, spostamento che può essere caratterizzato con il moto di un insieme discreto di caratteristiche o elementi significativi ("features") o con quello della distribuzione di luminosità nelle immagini.

Una prima classe di metodi di stima del moto 3D è basata sul calcolo del flusso ottico ("optical flow"), ovvero sulle variazioni nel tempo di valori risultanti dall'applicazione di operatori locali all'immagine.

La seconda classe di metodi si basa sull'individuazione, in ogni immagine della sequenza, di un insieme di elementi significativi 2D, rappresentativi di caratteristiche 3D degli oggetti, come vertici o linee di separazione tra superfici, e sulla determinazione della corrispondenza tra tali elementi nelle varie immagini.

Flusso Ottico e 3D



Rappresentazione schematica della relazione tra velocità $\mathbf{v}_0$ di un punto dell'oggetto (P_0) rispetto al sensore e quella della sua proiezione sul piano immagine (P_i).

O è il centro ottico del sistema, i vettori velocità $\mathbf{v}_0$ e $\mathbf{v}_i$ sono le derivate dei vettori OP_0 e OP_i , e coincidono con gli spostamenti per $\delta t=1$.

Flusso Ottico

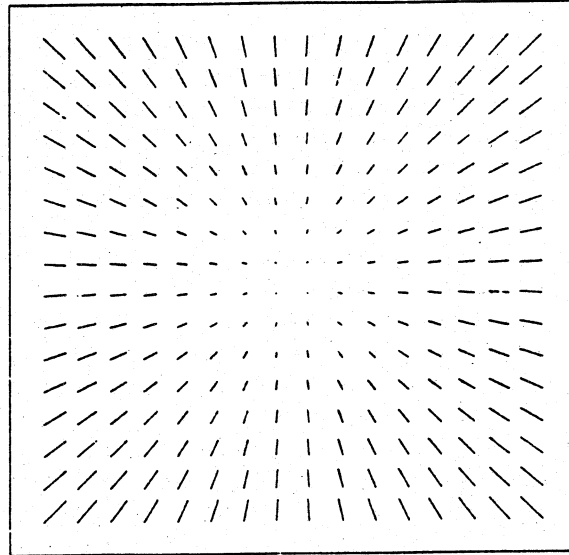
Consideriamo un oggetto in movimento relativo rispetto ad un sensore come una telecamera: ad un certo istante, se P_0 ha una velocità $\mathbf{v}_0$ relativa alla telecamera, al suo spostamento $\mathbf{v}_0 dt$ nell'intervallo di tempo dt , corrisponderà uno spostamento $\mathbf{v}_i dt$ di P_i , essendo $\mathbf{v}_0$ e $\mathbf{v}_i$ i vettori derivata temporale dei vettori che congiungono il centro ottico O del sistema con P_0 e P_i . Usando il legame tra OP_0 e OP_i , risulta assegnato un vettore spostamento (o velocità) ad ogni punto P_i dell'immagine: l'insieme di questi vettori costituisce il campo di spostamenti ("motion field") effettivo. In realtà, però, l'informazione disponibile è solo la variazione temporale della distribuzione di luminosità 2D: il flusso ottico è proprio definito come il campo vettoriale di tali velocità e non sempre coincide con la proiezione del movimento 3D.

Flusso Ottico

Per chiarire la differenza tra quest'ultima e il moto apparente della luminosità 2D, e il significato del termine "problema mal condizionato", consideriamo una sfera perfetta che ruota davanti al sistema di acquisizione: l'immagine acquisita non presenta variazioni mentre la sfera ruota, per cui il flusso ottico è nullo ovunque, mentre il campo di spostamenti effettivo è diverso da zero.

Il caso duale è rappresentato da una sfera fissa sottoposta ad una sorgente di illuminazione mobile, che produce quindi ombre variabili nel tempo: il flusso ottico è diverso da zero, a fronte di un motion field nullo ovunque.

Flusso Ottico



Esempio ideale di flusso ottico corrispondente al campo di velocità effettivo, relativa ad un oggetto che si allontana con velocità uniforme dal centro ottico.

Calcolo del flusso ottico

Vediamo ora le equazioni di base che permettono il calcolo del flusso ottico.

Indicheremo con $f(\mathbf{x}, \mathbf{y}, t)$ l'intensità luminosa nell'immagine nel punto $(\mathbf{x}, \mathbf{y})$ al tempo t .

Assumeremo come ipotesi di partenza che tale intensità vari solamente a causa del moto 3D, e quindi sia la stessa al tempo $t+\delta t$ nel punto $(\mathbf{x}+\delta \mathbf{x}, \mathbf{y}+\delta \mathbf{y})$ dell'immagine, sarà

$$f(\mathbf{x}+\delta \mathbf{x}, \mathbf{y}+\delta \mathbf{y}, t+\delta t) = f(\mathbf{x}, \mathbf{y}, t)$$

con

$$\delta \mathbf{x} = \mathbf{u} \delta t$$

e

$$\delta \mathbf{y} = \mathbf{v} \delta t,$$

definendo con u e v le componenti della velocità (cioè del vettore del flusso ottico nel punto considerato) nelle direzioni x e y , cioè $u = dx/dt$, $v = dy/dt$.

Equazione fondamentale

Supponendo che δt , δx e δy siano molto piccole, si ottiene l'equazione fondamentale:

$$f(x,y,t) + f_x \delta x + f_y \delta y + f_t \delta t + \varepsilon = f(x,y,t)$$

dove f_x , f_y e f_t indicano le derivate parziali dell'intensità rispetto alle variabili spaziali e temporale, ed ε contiene termini di ordine superiore al primo in δx , δy e δt . Dividendo per δt e facendo il limite per δt che tende a zero si ottiene

$$f_x u + f_y v + f_t = 0$$

o riscritta in forma vettoriale

$$\mathbf{v} \cdot \mathbf{grad} f = -f_t$$

Tale equazione viene detta l'equazione di vincolo del flusso ottico. Tale equazione non permette di calcolare le due incognite u e v . Permette solo di ottenere l'informazione relativa alla componente del flusso ottico nella direzione del gradiente

$$(f_t / (f_x^2 + f_y^2)^{1/2}),$$

ma non in quella perpendicolare, cioè lungo le linee a intensità luminosa costante. Inoltre $\mathbf{v}$ non può essere calcolata nei punti dove il gradiente è nullo.

Calcolo del flusso ottico

Per il calcolo del flusso ottico occorre quindi introdurre un ulteriore vincolo, derivato da conoscenze a priori (limitazioni sul tipo di moto) o dal fatto che di solito il campo di spostamento varia dolcemente in gran parte dell'immagine. Un tipo di soluzione è ottenuta allora minimizzando contemporaneamente lo scostamento da zero del primo membro della equazione precedente e le variazioni del flusso ottico, misurate dal suo gradiente. Occorre cioè minimizzare rispetto a u e v il funzionale

$$\int \int \{ (f_x u + f_y v + f_t)^2 + \lambda [(u_x^2 + u_y^2) + (v_x^2 + v_y^2)] \} dx dy$$

dove λ è un parametro che governa il peso relativo dei due errori.

Limitazioni

Una limitazione dell'approccio di base descritto sta nell'assunzione di regolarità del flusso ottico e di continuità della luminosità nell'immagine. Se il flusso ha discontinuità, come capita in corrispondenza dei bordi degli oggetti occorre evitare di estendere con regolarità la soluzione da una regione all'altra.

Ciò è in contrasto con il fatto che proprio le discontinuità del flusso ottico possono essere utilizzate per segmentare l'immagine, viceversa se si disponesse del risultato della segmentazione si otterrebbe una migliore stima del flusso ottico stesso.

Una soluzione proposta consiste nell'inserire la segmentazione all'interno della procedura iterativa di stima: ad ogni iterazione si individuano le zone dove il flusso varia bruscamente e si impedisce all'iterazione successiva di collegare con continuità la soluzione attraverso la discontinuità.

Limitazioni

Una seconda limitazione proviene dall'aver trascurato i termini superiori al primo ordine nella equazione, ciò rende insoddisfacente l'algoritmo in presenza di contorni e di angoli nell'immagine, per ovviare a tale problema sono stati proposti metodi che considerano anche le derivate seconde dell'espansione in serie di Taylor, tuttavia la minimizzazione del funzionale risultante richiede la soluzione di un sistema di due equazioni non lineari che contengono i cubi delle incognite. Queste equazioni assumono tuttavia forma lineare in u e v nei punti dell'immagine dove esistono angoli nella distribuzione di luminosità, vale a dire dove la direzione del gradiente di luminosità cambia bruscamente. La procedura consiste quindi nell'individuare queste zone con un opportuno operatore, valutare la soluzione in esse e usare questi valori come stima iniziale di un algoritmo iterativo che estende la soluzione del sistema non lineare nelle zone circostanti.

Altri Metodi

Esistono numerosi metodi alternativi per la stima del flusso ottico, e in particolare per i vincoli da affiancare all'equazione mostrata.

Un caso interessante è quello degli algoritmi ricorsivi per la stima dello spostamento in immagini televisive, usato a scopo di riduzione di ridondanza basata sulla predizione del movimento.

Esistono poi metodi gerarchici in cui si opera inizialmente su una sequenza di immagini a bassa risoluzione, (ad esempio sottocampionate, o ad un livello alto di una rappresentazione piramidale) per ottenere una stima del flusso ottico che viene poi propagata ai livelli di miglior risoluzione per affinarla.

Altri Metodi

Un altro metodo è basato sull'uso di un'equazione di vincolo riferita non alla luminosità soltanto, ma a diverse funzioni g_k

$$\frac{\partial g_k}{\partial x} u + \frac{\partial g_k}{\partial y} v + \frac{\partial g_k}{\partial t} = 0 \quad k=1,2, \dots, n$$

dove le g_k sono misure di contrasto, o di entropia, dell'immagine.

Va poi ricordato che sono state messe a punto tecniche per stimare il flusso ottico relativo a una scena composta da più oggetti in movimento. Esse cercano di segmentare il campo di velocità in regioni corrispondenti a oggetti che seguono leggi del moto diverse, mediante raggruppamento di punti aventi vettori simili e con l'ausilio anche di informazione sui contorni.

Calcolo del flusso ottico 3D

La disponibilità del campo vettoriale di velocità in due dimensioni permette di stimare il movimento e la struttura spaziale dell'oggetto mediante un'opportuna formulazione delle relazioni tra i punti della superficie nello spazio e quelli nel piano immagine, e tra le associate velocità.

Tali relazioni sono quelle indotte dalla proiezione prospettica e richiedono alcune assunzioni, tipicamente la variazione dolce del flusso ottico e la regolarità della superficie 3D, per poterne troncare ai primi termini gli sviluppi in serie di Taylor.

Le equazioni finali sono, nella forma generale, piuttosto complicate: nonlineari nelle caratteristiche della superficie (pendenze e curvature) e del moto, con coefficienti costituiti dalle componenti u e v del moto 2D e dalle loro derivate spaziali prime e seconde.

Calcolo del flusso ottico 3D

Problemi legati al metodo:

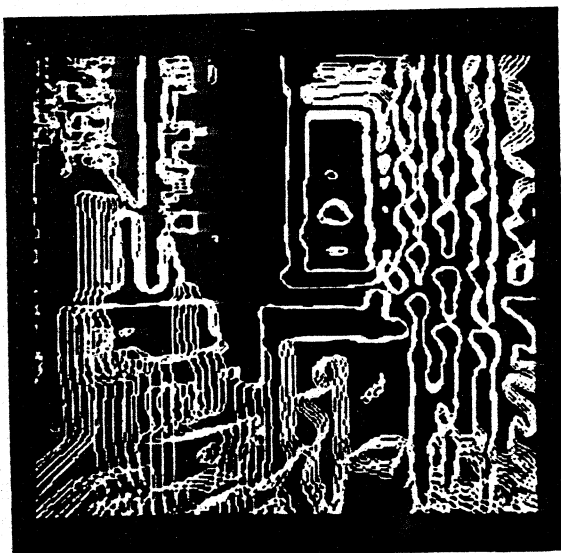
- ◆ Spesso la soluzione delle equazioni non è unica .
- ◆ La velocità di traslazione e le informazioni di struttura sono ricavabili solo a meno di un fattore scala.
- ◆ La curvatura della superficie è ricavabile solo se il movimento comprende una componente di traslazione parallela al piano immagine.

Data la complessità del sistema generale e le precedenti considerazioni, sono stati proposti e usati numerosi metodi basati su questa impostazione, ma con limitazioni sul tipo di moto (solo traslatorio o solo rotatorio) o delle superfici degli oggetti (ad esempio: solo planari): queste limitazioni sono peraltro rappresentative di casi ~~applicativi~~ di interesse industriale, ad esempio si può pensare all'esplorazione di un oggetto mediante il movimento controllato del sensore visivo.

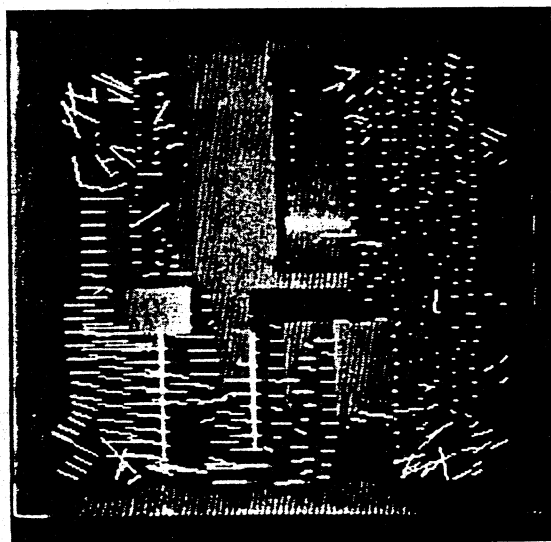
Calcolo del flusso ottico 3D

Si può notare che per una traslazione pura i vettori velocità si intersecano in un punto, detto fuoco di espansione (FOE) o di contrazione (FOC), a seconda che il sensore si avvicini o si allontani dall'oggetto; una pura rotazione li vincolerebbe a giacere su una sezione di conica.

La figura a) mostra sovrapposti i contorni estratti dalle cinque immagini, mentre in b) è rappresentato, sovrapposto ad una delle viste, un sottoinsieme dei vettori del flusso ottico.



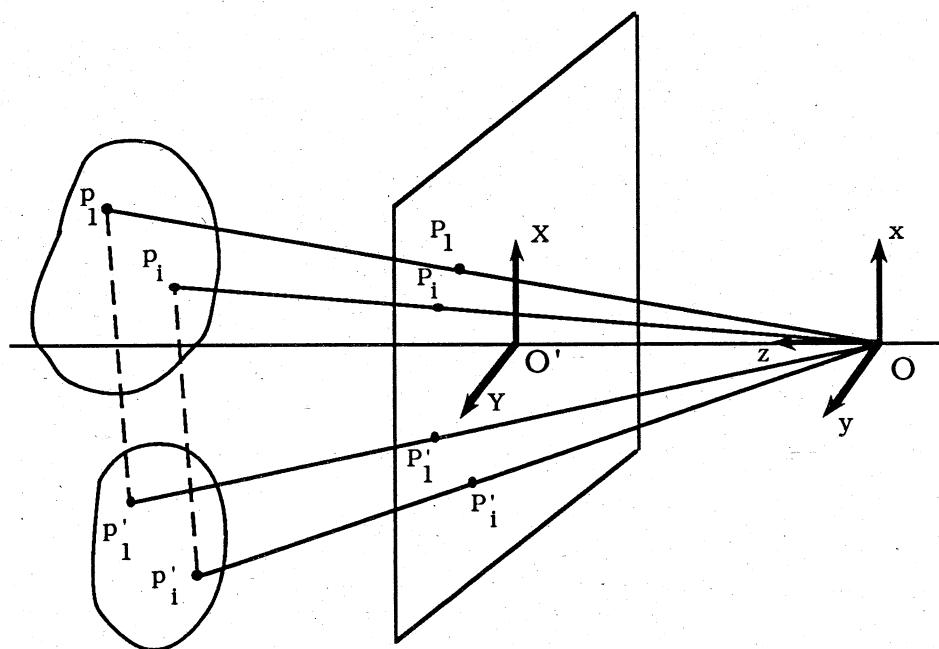
a)



b)

Stime basate su corrispondenze discrete

Si considereranno ora le tecniche che non ricorrono al campo vettoriale di spostamenti 2D nel piano immagine, ma si basano sull'estrazione di caratteristiche, o *features* in ogni immagine della sequenza e sulla determinazione della corrispondenza tra di esse. La situazione è schematizzata nella figura, dove le "caratteristiche" sono punti P_i individuati sul piano immagine (X,Y) ad un certo istante come proiezione dei punti p_i dell'oggetto (in coordinate x,y,z): all'istante successivo i punti si sono spostati rispettivamente in P'_i e p'_i .



Corrispondenze discrete

Queste variazioni di posizione dei punti corrispondenti in (X,Y) sono i dati disponibili per tentare di ricavare il movimento dei punti p_i in (x,y,z) tra i due istanti, sempre che il problema ammetta soluzione.

Occorre notare che, oltre ad essere mal condizionato, il problema è intrinsecamente ambiguo: la soluzione, quando esiste, può solo essere ottenuta a meno di una costante, tipicamente un fattore scala sulla traslazione. Ciò è comprensibile notando, ad esempio, che in base ai dati (le proiezioni sul piano immagine) non risultano distinguibili i seguenti due moti: quello di un oggetto di date dimensioni che subisce una certa rototraslazione a distanza d dal sensore, e quello dell'oggetto di dimensioni doppie che si muove con velocità di traslazione doppia e di rotazione identica al precedente, a distanza $2d$ dal sensore.

Corrispondenze discrete

Il caso precedente differisce da quanto descritto riguardo al flusso ottico essenzialmente per il numero di punti considerati e la densità del flusso ottico.

In altri casi gli elementi di cui si assume nota la corrispondenza in (X,Y) non sono punti, ma ad esempio linee o regioni, per cui è meno giustificato il termine "discreto" usato per definire questi metodi, e più sfumata la differenza rispetto all'uso del flusso ottico 2D.

Corrispondenze discrete

E' difficile descrivere in modo sistematico i tutti possibili metodi di soluzione proposti in letteratura.

Possono comunque essere classificati secondo una varietà di parametri, quali:

- ◆ il tipo di osservabili considerato (punti, linee, angoli)
- ◆ eventuali vincoli spaziali tra gli osservabili (ad esempio coplanarità)
- ◆ il tipo di proiezione (ortografica, prospettica, sferica)
- ◆ i parametri del moto ricercati (direttamente quelli di rototraslazione, o parametri intermedi)
- ◆ la linearità o meno del sistema di equazioni (in dipendenza anche da alcune scelte precedenti)
- ◆ la rigidità o meno del moto (implicitamente o esplicitamente assunta nella formulazione)
- ◆ la limitazione o meno a piccoli spostamenti, e analogamente l'uso di un numero limitato di viste o di una lunga sequenza.

Calcolo delle corrispondenze

L'ipotesi su cui si basano i metodi di stima del movimento da features estratte in un'immagine e "inseguite" nella sequenza è che siano stati innanzitutto individuati elementi caratteristici in una vista, e che sia stata stabilita la corrispondenza tra di essi nella sequenza. In realtà, stabilire e mantenere le corrispondenze tra frame successivi è complicato, sia perchè gli elementi da inseguire non necessariamente permangono in tutte le immagini della sequenza, sia perchè è proprio il movimento una delle proprietà che aiutano a individuare le corrispondenze.

In ogni caso, le tecniche di più comune impiego per individuare le features sono basate sull'analisi delle caratteristiche locali della distribuzione di luminosità, come la tessitura, oppure su caratteristiche strutturali delle transizioni, quali angoli, tratti di contorno rettilinei.

Corrispondenze

Per risolvere poi il problema delle corrispondenze si usano misure di somiglianza opportunamente formulate. Una di queste è la correlazione locale, su intorni 1D e 2D, usando sia l'intensità luminosa che l'uscita di preelaborazioni. Altre misure sono quelle di somiglianza strutturale, cioè tra attributi significativi associati agli elementi estratti da ciascuna immagine. Nell'individuare e mantenere le corrispondenze sono anche introdotti vincoli di regolarità sulla velocità con cui il moto o la posizione può variare. In alternativa, gli errori e le ambiguità dovute al rumore e alle occlusioni possono essere corretti con verifiche di consistenza a posteriori.

Metodi di stima

Consideriamo inizialmente il caso in cui le caratteristiche siano dei punti, dove (x,y,z) sono le coordinate di un punto p della superficie di un oggetto in moto rigido relativo al sensore (assunto nell'origine O del sistema di coordinate assolute), e (X,Y) quelle della sua proiezione P sul piano immagine. Il problema di stimare la velocità 3D, o lo spostamento nell'unità di tempo, dell'insieme di punti p_i conoscendo le proiezioni in istanti successivi dei punti P_i è retto da due equazioni di base, che descrivono la proiezione (qui assunta prospettica, come nella maggior parte dei casi reali) e il moto rigido.

La proiezione è descritta da

$$X = F x/z$$

$$Y = F y/z$$

$$X' = F x'/z'$$

$$Y' = F y'/z'$$

Metodi di stima

Il moto rigido tra due istanti successivi è rappresentabile da una matrice di rotazione R , ortonormale, di dimensioni 3×3 e con elementi funzione di tre parametri indipendenti (quali due coseni direttori di un asse e l'angolo di rotazione attorno a questo), e da un vettore traslazione T di componenti T_x , T_y e T_z . Il moto di un corpo rigido ha perciò sei gradi di libertà e si può scrivere come

$$\mathbf{p}' = R\mathbf{p} + \mathbf{T}$$

dove $\mathbf{p}'$ e $\mathbf{p}$ sono vettori di componenti (x', y', z') e (x, y, z) . A partire da queste assunzioni di base si ottengono numerose tecniche, che conducono a equazioni lineari o non lineari, in incognite che sono o le posizioni successive dei punti p (da cui il moto) o i parametri del moto (come R e T).

Conclusioni

Esiste una grande varietà di metodi di stima del movimento 3D da sequenze, nessuno dei quali può dirsi del tutto soddisfacente, se non per applicazioni specifiche. I metodi differenziali, basati sul flusso ottico, richiedono l'ipotesi di continuità del moto e tendono ad esaltare gli effetti del rumore a causa degli operatori derivativi. Le tecniche basate sulle corrispondenze discrete valgono per spostamenti qualsiasi ma pongono la difficoltà di individuare e mettere in corrispondenza gli osservabili con la precisione necessaria per ottenere stime affidabili. E' quindi intensa l'attività di ricerca su miglioramenti ed estensioni delle tecniche di base, ad esempio con gli obiettivi di regolarizzare le soluzioni, eliminare l'assunzione di moto rigido, progettare sistemi integrati che usino anche informazioni sulla struttura tridimensionale o altre conoscenze a priori.



LABORATORIO DI VISIONE ARTIFICIALE

Tel. 0382/391372

Telefax 0382/412881

Tecniche e Strumenti per l'Elaborazione di
Immagini ed il Riconoscimento di Forme

Parte II: Livello Alto e Strumenti Disponibili

9.

29 Novembre 1989
c/o SPERONI Spa, Spessa Po

Luca Lombardi:

Sequenza di immagini: stima del movimento.

Dipartimento di Informatica e Sistemistica, Università
degli Studi di Pavia

Stima del movimento

Il flusso ottico.

Corrispondenze discrete.

Sequenze di Immagini

La stima del movimento introduce nella visione artificiale una componente nuova ed essenziale

il tempo

L'evoluzione temporale dell'ambiente e della scena comporta un enorme incremento dei dati necessari per elaborare gli *eventi* che hanno luogo nel campo visivo.

L'evoluzione temporale sulla scena non dipende semplicemente dal moto degli oggetti: anche in una scena stazionaria si hanno dei cambiamenti causati dal moto relativo dei sensori (o dell'osservatore).

Anzi proprio il moto relativo può essere sfruttato per ricostruire la natura tridimensionale della scena osservata.

Una fonte di informazione importante è rappresentata anche dalla variazione dei toni di grigio.

Data l'enorme quantità di informazioni da trattare solo negli ultimi anni l'analisi di sequenze di immagini ha avuto uno sviluppo significativo.

Stima del Movimento

Il problema della ricostruzione della struttura volumetrica e del moto da viste 2D ha una grande importanza non solo per le difficoltà teoriche che pone, ma anche per la dimensione delle sue applicazioni pratiche. Dal punto di vista teorico, analizzare una sequenza pone difficoltà di tipo nuovo rispetto al trattamento di immagini statiche, poichè oltre ad estrarre le parti di interesse da ogni singola immagine occorre integrarle e interpretarle in maniera coerente. Dal punto di vista delle applicazioni, si possono citare da un lato quelle in cui il movimento e la struttura sono usati a scopo di interpretazione, descrizione o controllo e dall'altro quelle in cui l'obiettivo è la rappresentazione efficiente del contenuto informativo della sequenza.

Stima del Movimento

In generale, la stima del movimento (e della struttura) 3D a partire dalle sue proiezioni 2D su una sequenza di immagini comprende tecniche di diverso livello: da quello del trattamento dei dati di luminosità per ricavare informazione sul moto nel piano di formazione delle immagini, all'interpretazione di questo per ricavarne il moto 3D e la struttura degli oggetti, fino all'integrazione con modelli di scena e conoscenze a priori.

Stima del Movimento

Occorre notare che questo problema, schematizzabile come il tentativo di ottenere informazione 3D da proiezioni 2D, appartiene alla categoria dei problemi **inversi o mal condizionati**, per i quali cioè non è garantita l'esistenza e l'unicità della soluzione.

Ma soprattutto tale soluzione risulta molto "sensibile" ad errori ed approssimazioni nei dati di partenza: a causa del processo di inversione, variazioni anche molto piccole nei dati si ripercuotono amplificate in errori significativi nei risultati, se non vengono presi adeguati accorgimenti correttivi.

Stima del Movimento

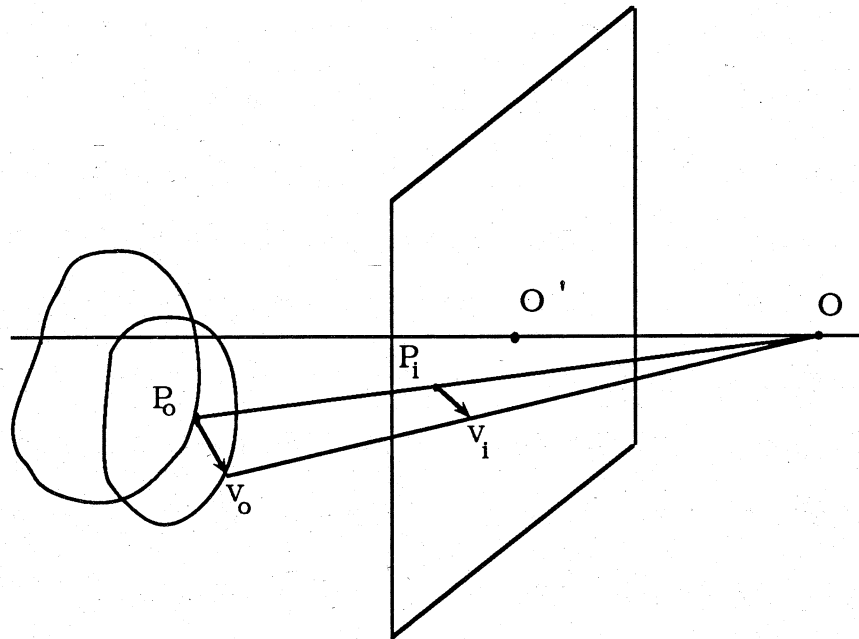
Si distinguono due classi di metodi fondamentali di stima del moto da una sequenza di viste.

Il movimento relativo tra gli oggetti della scena ed il sensore produce uno spostamento di parti delle immagini formate nel piano focale del sensore, spostamento che può essere caratterizzato con il moto di un insieme discreto di caratteristiche o elementi significativi ("features") o con quello della distribuzione di luminosità nelle immagini.

Una prima classe di metodi di stima del moto 3D è basata sul calcolo del flusso ottico ("optical flow"), ovvero sulle variazioni nel tempo di valori risultanti dall'applicazione di operatori locali all'immagine.

La seconda classe di metodi si basa sull'individuazione, in ogni immagine della sequenza, di un insieme di elementi significativi 2D, rappresentativi di caratteristiche 3D degli oggetti, come vertici o linee di separazione tra superfici, e sulla determinazione della corrispondenza tra tali elementi nelle varie immagini.

Flusso Ottico e 3D



Rappresentazione schematica della relazione tra velocità $\mathbf{v}_0$ di un punto dell'oggetto (P_0) rispetto al sensore e quella della sua proiezione sul piano immagine (P_i).

O è il centro ottico del sistema, i vettori velocità $\mathbf{v}_0$ e $\mathbf{v}_i$ sono le derivate dei vettori OP_0 e OP_i , e coincidono con gli spostamenti per $\delta t=1$.

Flusso Ottico

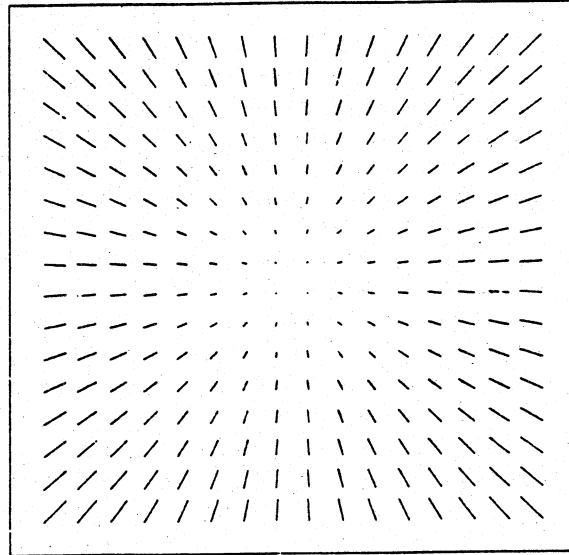
Consideriamo un oggetto in movimento relativo rispetto ad un sensore come una telecamera: ad un certo istante, se P_0 ha una velocità $\mathbf{v}_0$ relativa alla telecamera, al suo spostamento $\mathbf{v}_0 dt$ nell'intervallo di tempo dt , corrisponderà uno spostamento $\mathbf{v}_i dt$ di P_i , essendo $\mathbf{v}_0$ e $\mathbf{v}_i$ i vettori derivata temporale dei vettori che congiungono il centro ottico O del sistema con P_0 e P_i . Usando il legame tra OP_0 e OP_i , risulta assegnato un vettore spostamento (o velocità) ad ogni punto P_i dell'immagine: l'insieme di questi vettori costituisce il campo di spostamenti ("motion field") effettivo. In realtà, però, l'informazione disponibile è solo la variazione temporale della distribuzione di luminosità 2D: il flusso ottico è proprio definito come il campo vettoriale di tali velocità e non sempre coincide con la proiezione del movimento 3D.

Flusso Ottico

Per chiarire la differenza tra quest'ultima e il moto apparente della luminosità 2D, e il significato del termine "problema mal condizionato", consideriamo una sfera perfetta che ruota davanti al sistema di acquisizione: l'immagine acquisita non presenta variazioni mentre la sfera ruota, per cui il flusso ottico è nullo ovunque, mentre il campo di spostamenti effettivo è diverso da zero.

Il caso duale è rappresentato da una sfera fissa sottoposta ad una sorgente di illuminazione mobile, che produce quindi ombre variabili nel tempo: il flusso ottico è diverso da zero, a fronte di un motion field nullo ovunque.

Flusso Ottico



Esempio ideale di flusso ottico corrispondente al campo di velocità effettivo, relativa ad un oggetto che si allontana con velocità uniforme dal centro ottico.

Calcolo del flusso ottico

Vediamo ora le equazioni di base che permettono il calcolo del flusso ottico.

Indicheremo con $f(\mathbf{x}, \mathbf{y}, t)$ l'intensità luminosa nell'immagine nel punto $(\mathbf{x}, \mathbf{y})$ al tempo t .

Assumeremo come ipotesi di partenza che tale intensità vari solamente a causa del moto 3D, e quindi sia la stessa al tempo $t+\delta t$ nel punto $(\mathbf{x}+\delta \mathbf{x}, \mathbf{y}+\delta \mathbf{y})$ dell'immagine, sarà

$$f(\mathbf{x}+\delta \mathbf{x}, \mathbf{y}+\delta \mathbf{y}, t+\delta t) = f(\mathbf{x}, \mathbf{y}, t)$$

con

$$\delta \mathbf{x} = \mathbf{u} \delta t$$

e

$$\delta \mathbf{y} = \mathbf{v} \delta t,$$

definendo con u e v le componenti della velocità (cioè del vettore del flusso ottico nel punto considerato) nelle direzioni x e y , cioè $u = dx/dt$, $v = dy/dt$.

Equazione fondamentale

Supponendo che δt , δx e δy siano molto piccole, si ottiene l'equazione fondamentale:

$$f(x,y,t) + f_x \delta x + f_y \delta y + f_t \delta t + \varepsilon = f(x,y,t)$$

dove f_x , f_y e f_t indicano le derivate parziali dell'intensità rispetto alle variabili spaziali e temporale, ed ε contiene termini di ordine superiore al primo in δx , δy e δt . Dividendo per δt e facendo il limite per δt che tende a zero si ottiene

$$f_x u + f_y v + f_t = 0$$

o riscritta in forma vettoriale

$$\mathbf{v} \cdot \mathbf{grad} f = -f_t$$

Tale equazione viene detta l'equazione di vincolo del flusso ottico. Tale equazione non permette di calcolare le due incognite u e v . Permette solo di ottenere l'informazione relativa alla componente del flusso ottico nella direzione del gradiente

$$(f_t / (f_x^2 + f_y^2)^{1/2}),$$

ma non in quella perpendicolare, cioè lungo le linee a intensità luminosa costante. Inoltre $\mathbf{v}$ non può essere calcolata nei punti dove il gradiente è nullo.

Calcolo del flusso ottico

Per il calcolo del flusso ottico occorre quindi introdurre un ulteriore vincolo, derivato da conoscenze a priori (limitazioni sul tipo di moto) o dal fatto che di solito il campo di spostamento varia dolcemente in gran parte dell'immagine. Un tipo di soluzione è ottenuta allora minimizzando contemporaneamente lo scostamento da zero del primo membro della equazione precedente e le variazioni del flusso ottico, misurate dal suo gradiente. Occorre cioè minimizzare rispetto a u e v il funzionale

$$\int \int \{ (f_x u + f_y v + f_t)^2 + \lambda [(u_x^2 + u_y^2) + (v_x^2 + v_y^2)] \} dx dy$$

dove λ è un parametro che governa il peso relativo dei due errori.

Limitazioni

Una limitazione dell'approccio di base descritto sta nell'assunzione di regolarità del flusso ottico e di continuità della luminosità nell'immagine. Se il flusso ha discontinuità, come capita in corrispondenza dei bordi degli oggetti occorre evitare di estendere con regolarità la soluzione da una regione all'altra.

Ciò è in contrasto con il fatto che proprio le discontinuità del flusso ottico possono essere utilizzate per segmentare l'immagine, viceversa se si disponesse del risultato della segmentazione si otterrebbe una migliore stima del flusso ottico stesso.

Una soluzione proposta consiste nell'inserire la segmentazione all'interno della procedura iterativa di stima: ad ogni iterazione si individuano le zone dove il flusso varia bruscamente e si impedisce all'iterazione successiva di collegare con continuità la soluzione attraverso la discontinuità.

Limitazioni

Una seconda limitazione proviene dall'aver trascurato i termini superiori al primo ordine nella equazione, ciò rende insoddisfacente l'algoritmo in presenza di contorni e di angoli nell'immagine, per ovviare a tale problema sono stati proposti metodi che considerano anche le derivate seconde dell'espansione in serie di Taylor, tuttavia la minimizzazione del funzionale risultante richiede la soluzione di un sistema di due equazioni non lineari che contengono i cubi delle incognite. Queste equazioni assumono tuttavia forma lineare in u e v nei punti dell'immagine dove esistono angoli nella distribuzione di luminosità, vale a dire dove la direzione del gradiente di luminosità cambia bruscamente. La procedura consiste quindi nell'individuare queste zone con un opportuno operatore, valutare la soluzione in esse e usare questi valori come stima iniziale di un algoritmo iterativo che estende la soluzione del sistema non lineare nelle zone circostanti.

Altri Metodi

Esistono numerosi metodi alternativi per la stima del flusso ottico, e in particolare per i vincoli da affiancare all'equazione mostrata.

Un caso interessante è quello degli algoritmi ricorsivi per la stima dello spostamento in immagini televisive, usato a scopo di riduzione di ridondanza basata sulla predizione del movimento.

Esistono poi metodi gerarchici in cui si opera inizialmente su una sequenza di immagini a bassa risoluzione, (ad esempio sottocampionate, o ad un livello alto di una rappresentazione piramidale) per ottenere una stima del flusso ottico che viene poi propagata ai livelli di miglior risoluzione per affinarla.

Altri Metodi

Un altro metodo è basato sull'uso di un'equazione di vincolo riferita non alla luminosità soltanto, ma a diverse funzioni g_k

$$\frac{\partial g_k}{\partial x} u + \frac{\partial g_k}{\partial y} v + \frac{\partial g_k}{\partial t} = 0 \quad k=1,2, \dots, n$$

dove le g_k sono misure di contrasto, o di entropia, dell'immagine.

Va poi ricordato che sono state messe a punto tecniche per stimare il flusso ottico relativo a una scena composta da più oggetti in movimento. Esse cercano di segmentare il campo di velocità in regioni corrispondenti a oggetti che seguono leggi del moto diverse, mediante raggruppamento di punti aventi vettori simili e con l'ausilio anche di informazione sui contorni.

Calcolo del flusso ottico 3D

La disponibilità del campo vettoriale di velocità in due dimensioni permette di stimare il movimento e la struttura spaziale dell'oggetto mediante un'opportuna formulazione delle relazioni tra i punti della superficie nello spazio e quelli nel piano immagine, e tra le associate velocità.

Tali relazioni sono quelle indotte dalla proiezione prospettica e richiedono alcune assunzioni, tipicamente la variazione dolce del flusso ottico e la regolarità della superficie 3D, per poterne troncare ai primi termini gli sviluppi in serie di Taylor.

Le equazioni finali sono, nella forma generale, piuttosto complicate: nonlineari nelle caratteristiche della superficie (pendenze e curvature) e del moto, con coefficienti costituiti dalle componenti u e v del moto 2D e dalle loro derivate spaziali prime e seconde.

Calcolo del flusso ottico 3D

Problemi legati al metodo:

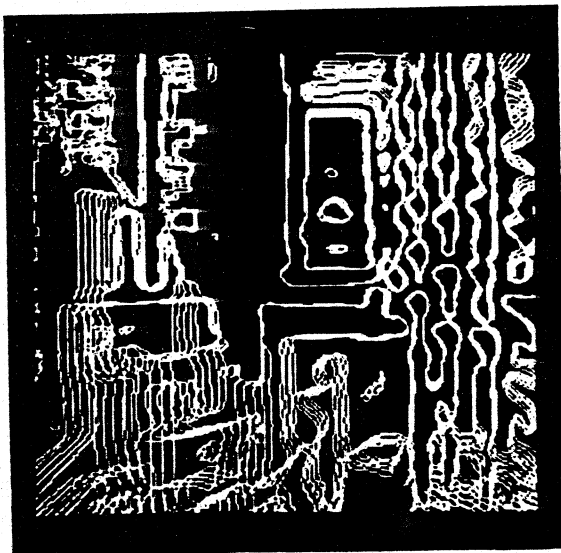
- ♦ Spesso la soluzione delle equazioni non è unica .
- ♦ La velocità di traslazione e le informazioni di struttura sono ricavabili solo a meno di un fattore scala.
- ♦ La curvatura della superficie è ricavabile solo se il movimento comprende una componente di traslazione parallela al piano immagine.

Data la complessità del sistema generale e le precedenti considerazioni, sono stati proposti e usati numerosi metodi basati su questa impostazione, ma con limitazioni sul tipo di moto (solo traslatorio o solo rotatorio) o delle superfici degli oggetti (ad esempio: solo planari): queste limitazioni sono peraltro rappresentative di casi ~~applicativi~~ di interesse industriale, ad esempio si può pensare all'esplorazione di un oggetto mediante il movimento controllato del sensore visivo.

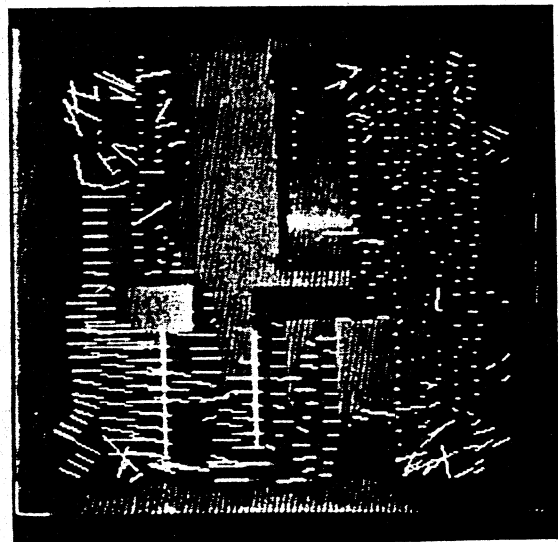
Calcolo del flusso ottico 3D

Si può notare che per una traslazione pura i vettori velocità si intersecano in un punto, detto fuoco di espansione (FOE) o di contrazione (FOC), a seconda che il sensore si avvicini o si allontani dall'oggetto; una pura rotazione li vincolerebbe a giacere su una sezione di conica.

La figura a) mostra sovrapposti i contorni estratti dalle cinque immagini, mentre in b) è rappresentato, sovrapposto ad una delle viste, un sottoinsieme dei vettori del flusso ottico.



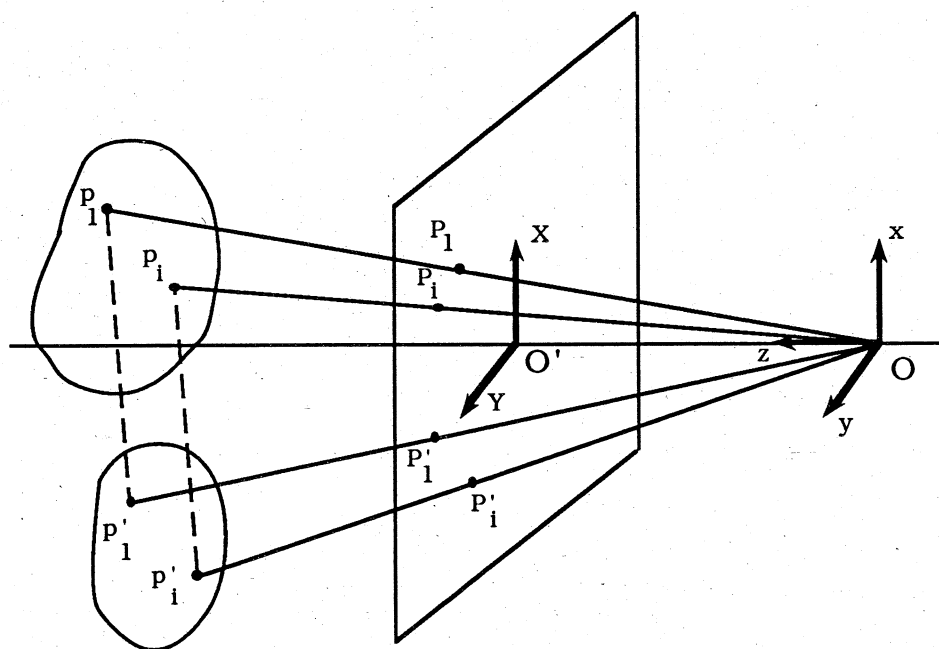
a)



b)

Stime basate su corrispondenze discrete

Si considereranno ora le tecniche che non ricorrono al campo vettoriale di spostamenti 2D nel piano immagine, ma si basano sull'estrazione di caratteristiche, o *features* in ogni immagine della sequenza e sulla determinazione della corrispondenza tra di esse. La situazione è schematizzata nella figura, dove le "caratteristiche" sono punti P_i individuati sul piano immagine (X,Y) ad un certo istante come proiezione dei punti p_i dell'oggetto (in coordinate x,y,z): all'istante successivo i punti si sono spostati rispettivamente in P'_i e p'_i .



Corrispondenze discrete

Queste variazioni di posizione dei punti corrispondenti in (X,Y) sono i dati disponibili per tentare di ricavare il movimento dei punti p_i in (x,y,z) tra i due istanti, sempre che il problema ammetta soluzione.

Occorre notare che, oltre ad essere mal condizionato, il problema è intrinsecamente ambiguo: la soluzione, quando esiste, può solo essere ottenuta a meno di una costante, tipicamente un fattore scala sulla traslazione. Ciò è comprensibile notando, ad esempio, che in base ai dati (le proiezioni sul piano immagine) non risultano distinguibili i seguenti due moti: quello di un oggetto di date dimensioni che subisce una certa rototraslazione a distanza d dal sensore, e quello dell'oggetto di dimensioni doppie che si muove con velocità di traslazione doppia e di rotazione identica al precedente, a distanza $2d$ dal sensore.

Corrispondenze discrete

Il caso precedente differisce da quanto descritto riguardo al flusso ottico essenzialmente per il numero di punti considerati e la densità del flusso ottico.

In altri casi gli elementi di cui si assume nota la corrispondenza in (X,Y) non sono punti, ma ad esempio linee o regioni, per cui è meno giustificato il termine "discreto" usato per definire questi metodi, e più sfumata la differenza rispetto all'uso del flusso ottico 2D.

Corrispondenze discrete

E' difficile descrivere in modo sistematico i tutti possibili metodi di soluzione proposti in letteratura.

Possono comunque essere classificati secondo una varietà di parametri, quali:

- ◆ il tipo di osservabili considerato (punti, linee, angoli)
- ◆ eventuali vincoli spaziali tra gli osservabili (ad esempio coplanarità)
- ◆ il tipo di proiezione (ortografica, prospettica, sferica)
- ◆ i parametri del moto ricercati (direttamente quelli di rototraslazione, o parametri intermedi)
- ◆ la linearità o meno del sistema di equazioni (in dipendenza anche da alcune scelte precedenti)
- ◆ la rigidità o meno del moto (implicitamente o esplicitamente assunta nella formulazione)
- ◆ la limitazione o meno a piccoli spostamenti, e analogamente l'uso di un numero limitato di viste o di una lunga sequenza.

Calcolo delle corrispondenze

L'ipotesi su cui si basano i metodi di stima del movimento da features estratte in un'immagine e "inseguite" nella sequenza è che siano stati innanzitutto individuati elementi caratteristici in una vista, e che sia stata stabilita la corrispondenza tra di essi nella sequenza. In realtà, stabilire e mantenere le corrispondenze tra frame successivi è complicato, sia perchè gli elementi da inseguire non necessariamente permangono in tutte le immagini della sequenza, sia perchè è proprio il movimento una delle proprietà che aiutano a individuare le corrispondenze.

In ogni caso, le tecniche di più comune impiego per individuare le features sono basate sull'analisi delle caratteristiche locali della distribuzione di luminosità, come la tessitura, oppure su caratteristiche strutturali delle transizioni, quali angoli, tratti di contorno rettilinei.

Corrispondenze

Per risolvere poi il problema delle corrispondenze si usano misure di somiglianza opportunamente formulate. Una di queste è la correlazione locale, su intorni 1D e 2D, usando sia l'intensità luminosa che l'uscita di preelaborazioni. Altre misure sono quelle di somiglianza strutturale, cioè tra attributi significativi associati agli elementi estratti da ciascuna immagine. Nell'individuare e mantenere le corrispondenze sono anche introdotti vincoli di regolarità sulla velocità con cui il moto o la posizione può variare. In alternativa, gli errori e le ambiguità dovute al rumore e alle occlusioni possono essere corretti con verifiche di consistenza a posteriori.

Metodi di stima

Consideriamo inizialmente il caso in cui le caratteristiche siano dei punti, dove (x,y,z) sono le coordinate di un punto p della superficie di un oggetto in moto rigido relativo al sensore (assunto nell'origine O del sistema di coordinate assolute), e (X,Y) quelle della sua proiezione P sul piano immagine. Il problema di stimare la velocità 3D, o lo spostamento nell'unità di tempo, dell'insieme di punti p_i conoscendo le proiezioni in istanti successivi dei punti P_i è retto da due equazioni di base, che descrivono la proiezione (qui assunta prospettica, come nella maggior parte dei casi reali) e il moto rigido.

La proiezione è descritta da

$$X = F x/z$$

$$Y = F y/z$$

$$X' = F x'/z'$$

$$Y' = F y'/z'$$

Metodi di stima

Il moto rigido tra due istanti successivi è rappresentabile da una matrice di rotazione R , ortonormale, di dimensioni 3×3 e con elementi funzione di tre parametri indipendenti (quali due coseni direttori di un asse e l'angolo di rotazione attorno a questo), e da un vettore traslazione T di componenti T_x , T_y e T_z . Il moto di un corpo rigido ha perciò sei gradi di libertà e si può scrivere come

$$\mathbf{p}' = R\mathbf{p} + \mathbf{T}$$

dove $\mathbf{p}'$ e $\mathbf{p}$ sono vettori di componenti (x', y', z') e (x, y, z) . A partire da queste assunzioni di base si ottengono numerose tecniche, che conducono a equazioni lineari o non lineari, in incognite che sono o le posizioni successive dei punti p (da cui il moto) o i parametri del moto (come R e T).

Conclusioni

Esiste una grande varietà di metodi di stima del movimento 3D da sequenze, nessuno dei quali può dirsi del tutto soddisfacente, se non per applicazioni specifiche. I metodi differenziali, basati sul flusso ottico, richiedono l'ipotesi di continuità del moto e tendono ad esaltare gli effetti del rumore a causa degli operatori derivativi. Le tecniche basate sulle corrispondenze discrete valgono per spostamenti qualsiasi ma pongono la difficoltà di individuare e mettere in corrispondenza gli osservabili con la precisione necessaria per ottenere stime affidabili. E' quindi intensa l'attività di ricerca su miglioramenti ed estensioni delle tecniche di base, ad esempio con gli obiettivi di regolarizzare le soluzioni, eliminare l'assunzione di moto rigido, progettare sistemi integrati che usino anche informazioni sulla struttura tridimensionale o altre conoscenze a priori.

