

 **COLLANA
INFORMATICA**

**NUOVA
EDIZIONE**

FrancoAngeli

Visione computazionale

Tecniche di ricostruzione
tridimensionale

Andrea Fusiello

Informazioni per il lettore

Questo file PDF è una versione gratuita di sole 20 pagine ed è leggibile con



La versione completa dell'e-book (a pagamento) è leggibile con Adobe Digital Editions. Per tutte le informazioni sulle condizioni dei nostri e-book (con quali dispositivi leggerli e quali funzioni sono consentite) consulta [cliccando qui](#) le nostre F.A.Q.



Collana di informatica – Nuova serie

diretta da Arrigo L. Frisiani

Comitato scientifico:

Giovanni Adorni (Università di Genova), Luigi Benedicenti (University of Regina), Aurelio Boari (Università di Bologna), Giacomo Bucci (Università di Firenze), Virginio Cantoni (Università di Pisa), Paolo Ciancarini (Università di Bologna), Gianni Conte (Università di Parma), Paolo Corsini (Università di Pisa), Fabio Crestani (Università della Svizzera Italiana), Rita Cucchiara (Università di Modena e Reggio Emilia), Valeria De Antonellis (Università di Brescia), Gianluca Foresti (Università di Udine), Alfonso Fuggetta (Politecnico di Milano), Andrea Fusiello (Università di Verona), Salvatore Gaglio (Università di Palermo), Marco Gori (Università di Siena), Enrico Grosso (Università di Sassari), Giovanni Guida (Università di Brescia), Giuseppe Iazeolla (Università di Roma "Tor Vergata"), Sebastiano Impedovo (Università di Bari), Pieter Kritzing (University of Cape Town), Massimo Maresca (Università di Padova), Paolo Maresca (Università di Napoli Federico II), Giuseppe Mastronardi (Politecnico di Bari), Antonino Mazzeo (Università di Napoli Federico II), Massimo Melucci (Università di Padova), Marco Mezzalama (Politecnico di Torino), Stefano Mizzaro (Università di Udine), Alfredo Petrosino (Università di Napoli "Parthenope"), Antonio Puliafito (Università di Messina), Gabriella Sanniti di Baja (CNR - Istituto di Cibernetica), Nello Scarabottolo (Università di Milano), Fabrizio Sebastiani (CNR - Istituto di Scienza e Tecnologie dell'Informazione), Giovanni Semeraro (Università di Bari), Alberto Sillitti (Libera Università di Bolzano), Giancarlo Succi (Libera Università di Bolzano), Carlo Tasso (Università di Udine), Genoveffa Tortora (Università di Salerno), Marco Vanneschi (Università di Pisa), Mario Vento (Università di Salerno), Alessandro Verri (Università di Genova), Lorenzo Vita (Università di Catania), Renato Zaccaria (Università di Genova).

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: www.francoangeli.it e iscriversi nella home page al servizio "Informatemi" per ricevere via e-mail le segnalazioni delle novità.

 **COLLANA
INFORMATICA**

**NUOVA
EDIZIONE**

FrancoAngeli

Visione computazionale

Tecniche di ricostruzione
tridimensionale

Andrea Fusiello

Codice MATLAB©, materiale aggiuntivo ed errata corrige si trovano in rete all'indirizzo http://www.francoangeli.it/Area_multimediale/Elenco_Libri.asp

Progetto grafico di copertina di *Alessandro Petrini*

2a edizione. Copyright © 2013, 2018 by FrancoAngeli s.r.l., Milano, Italy

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito www.francoangeli.it.

Indice

Prefazione	13
Ringraziamenti	15
1 Introduzione	17
1.1 Il prodigio della visione	17
1.2 Visione computazionale di basso livello	18
2 Formazione dell'immagine	20
2.1 Introduzione	20
2.2 Geometria della formazione dell'immagine	20
2.3 Immagini digitali	22
2.4 Lenti sottili	23
2.5 Fotometria della formazione dell'immagine	26
3 Calibrazione della fotocamera	29
3.1 Introduzione	29
3.2 Modello stenopeico della fotocamera	30
3.2.1 Modello semplificato	30
3.2.2 Modello generale	33
3.2.3 Dissezione della MPP	37
3.2.4 Espressioni alternative per la proiezione prospettica .	39
3.3 Calibrazione della fotocamera	40
3.3.1 Metodo DLT	41
3.3.2 Raffinamento non lineare	42

3.3.3	Fattorizzazione della MPP	43
3.3.4	Distorsione radiale	45
4	Geometria epipolare	50
4.1	Introduzione	50
4.2	Geometria epipolare	51
4.3	Matrice fondamentale	52
4.4	Calcolo della matrice fondamentale	54
4.4.1	Algoritmo dei sette punti	56
4.4.2	Precondizionamento	56
4.4.3	Raffinamento non lineare	58
5	Omografie tra immagini	63
5.1	Introduzione	63
5.2	Omografia indotta da un piano	63
5.3	Calcolo dell'omografia con DLT	66
5.4	Parallasse planare	68
5.5	Metodo di Sturm-Maybank-Zhang per la calibrazione	70
5.6	Calibrazione attraverso H_∞	73
6	Orientamento relativo	76
6.1	Introduzione	76
6.2	Matrice essenziale	77
6.2.1	Calcolo della matrice essenziale	78
6.2.2	Fattorizzazione della matrice essenziale	80
6.3	Metodo non lineare	83
7	Ricostruzione da due immagini	87
7.1	Triangolazione	87
7.1.1	Caso normale	87
7.1.2	Metodo linear-eigen	88
7.1.3	Raffinamento non lineare	90
7.2	Ambiguità della ricostruzione	91
7.3	Ricostruzione euclidea	93
7.4	Orientamento assoluto	94
7.4.1	Analisi procrustiana ortogonale	95
7.5	Ricostruzione proiettiva	97
7.6	Promozione euclidea da parametri interni noti	99

8	Orientamento esterno	104
8.1	Introduzione	104
8.2	Metodo lineare di Fiore	105
8.3	Metodo non lineare	106
8.4	Metodo diretto	108
9	Punti salienti	113
9.1	Introduzione	113
9.2	Operatore di Harris e Stephens	114
9.2.1	Accoppiamento e tracciamento	117
9.3	Scale Invariant Feature Transform (SIFT)	122
9.3.1	Spazio-scala	122
9.3.2	Filtraggio LoG	124
9.3.3	Rivelatore	125
9.3.4	Descrittore	128
9.3.5	Accoppiamento	130
10	Campo di moto	133
10.1	Introduzione	133
10.2	Il campo di moto	133
10.3	Ricostruzione dal campo di moto	136
10.3.1	Metodi lineari	138
10.4	Il flusso ottico	141
10.4.1	Metodo di Lucas-Kanade	142
11	Stereopsi: calcolo della disparità	147
11.1	Introduzione	147
11.2	Accoppiamento di finestre	150
11.2.1	Metriche di accoppiamento	151
11.2.2	Compromesso affidabilità - accuratezza	155
11.2.3	Indicatori di affidabilità	158
11.2.4	Occlusioni	159
11.3	Accoppiamento globale	160
12	Metodi attivi	165
12.1	Introduzione	165
12.2	Illuminazione strutturata	166
12.2.1	Stereo attivo	167
12.2.2	Triangolazione attiva	167

12.3 Stereo fotometrico	172
12.3.1 Dalle normali alle coordinate	176
12.4 Sensori a tempo di volo	177
13 Ricostruzione proiettiva multivista e autocalibrazione	180
13.1 Introduzione	180
13.2 Ricostruzione proiettiva multivista	181
13.2.1 Metodo di Sturm-Triggs	181
13.3 Autocalibrazione	183
13.3.1 Promozione euclidea	185
13.3.2 Metodo di Mendonça-Cipolla	187
13.4 Fattorizzazione di Tomasi-Kanade	191
14 Ricostruzione euclidea multivista	195
14.1 Introduzione	195
14.1.1 Tassonomia	198
14.2 Approcci basati sui punti	198
14.2.1 Ricostruzione incrementale	198
14.2.2 Ricostruzione ad albero	200
14.2.3 Compensazione per modelli indipendenti	201
14.3 Approccio basato sui sistemi di riferimento	202
14.3.1 Sincronizzazione delle rotazioni	204
14.3.2 Sincronizzazione delle traslazioni	206
14.3.3 Localizzazione da rilevamento	208
14.4 Compensazione per stelle proiettive	210
14.4.1 Matrice jacobiana della CSP	213
14.4.2 Sistema ridotto	216
15 Registrazione 3D	220
15.1 Introduzione	220
15.2 Iterative Closest Point	221
15.3 Metodi per l'allineamento di molte viste	224
15.3.1 Analisi procrustiana generalizzata	224
15.3.2 Metodi senza corrispondenze date	226
16 Ricostruzione volumetrica	230
16.1 Introduzione	230
16.2 Metodi in spazio-oggetto	232
16.2.1 Ricostruzione da sagome	233

16.2.2 Ricostruzione da fotocoerenza	237
16.3 Stereo multivista	240
16.4 Marching cubes	244
17 Sintesi di immagini	249
17.1 Introduzione	249
17.2 Trasformazioni parametriche	250
17.2.1 Mosaici	250
17.2.2 Stabilizzazione dell'immagine	256
17.2.3 Fotopiano	257
17.3 Rettificazione epipolare	258
17.4 Trasformazioni non parametriche	263
17.4.1 Trasferimento con la profondità	264
17.4.2 Interpolazione con disparità	264
17.4.3 Trasferimento epipolare	265
17.4.4 Trasferimento con parallasse	266
17.5 Trasformazione geometrica dell'immagine	267
A Nozioni di algebra lineare	275
A.1 Introduzione	275
A.2 Prodotto scalare	275
A.3 Norma matriciale	276
A.4 Matrice inversa	277
A.5 Determinante	277
A.6 Matrici ortogonali	279
A.7 Forme lineari e quadratiche	279
A.8 Rango	280
A.9 Autovalori ed autovettori	282
A.10 Decomposizione ai valori singolari	284
A.11 Prodotto esterno	289
A.12 Prodotto di Kronecker	291
A.12.1 Prodotto di Khatri-Rao	292
B Altre nozioni matematiche	294
B.1 Derivate di funzioni vettoriali e matriciali	294
B.2 Rotazioni	297
B.2.1 Differenziale della rotazione	298
B.3 Minimi quadrati non lineari	300
B.3.1 Metodo di Gauss-Newton	301

B.3.2	Metodo di Levenberg-Marquardt	302
B.4	Grafi	304
C	Nozioni di geometria proiettiva	306
C.1	Introduzione	306
C.2	Proiezione prospettica	307
C.3	Coordinate omogenee	309
C.4	Equazione della retta	310
C.5	Trasformazioni	311
D	Regressione	314
D.1	Introduzione	314
D.2	Minimi quadrati	315
D.3	Regressione robusta	316
D.3.1	Campioni anomali e robustezza	316
D.3.2	M-stimatori	317
D.3.3	Minima Mediana dei Quadrati	319
D.3.4	RANSAC	322
E	Codice Matlab	325
	Listati	327
	Acronimi	329
	Indice analitico	331

A mio padre

Prefazione

«Di solito, ci si convince meglio con le ragioni trovate da se stessi che non con quelle venute in mente ad altri.»

B. Pascal

Il testo si concentra sullo studio di teorie e tecniche computazionali per stimare le proprietà geometriche di oggetti solidi a partire dalle immagini. La scelta degli argomenti riflette una visione personale della disciplina e certamente è possibile riscontrare omissioni. La trattazione che ho privilegiato è una visita in profondità dei metodi della visione computazionale, piuttosto che una visita in ampiezza. Questo vuol dire che il testo non ha la presunzione di fornire una panoramica sui metodi esistenti; invece, ne sono stati selezionati alcuni e descritti ad un livello tale da consentirne l'implementazione. Infatti per molti di essi viene riportato il listato MATLAB (sorgenti disponibili sul web).

In questa seconda edizione quasi tutti i capitoli sono stati rimaneggiati, con l'aggiunta di nuovo materiale e lo spostamento di quello esistente. Le principali modifiche hanno riguardato:

- l'introduzione dei listati MATLAB degli algoritmi;
- la trattazione dettagliata dei metodi non-lineari, inclusa la compensazione per stelle proiettive, con i calcoli delle derivate;
- l'introduzione delle tecniche di "sincronizzazione" nel capitolo 14;
- la rimozione del capitolo 15 della prima edizione, che trattava i metodi basati su chiaroscuro, tessitura e sfocatura.

Il testo è sufficientemente modulare, il docente è libero di comporre il percorso che meglio si adatta al programma del corso, tuttavia suggerisco di includere sempre almeno i capitoli: 3, 4, 6, 7, 8, 14, che contengono un nucleo di concetti geometrici fondamentali, sia per la visione computazionale che per la fotogrammetria.

I prerequisiti principali sono le nozioni di algebra lineare che vengono richiamate in appendice A, con la maggior parte delle quali lo studente di informatica o ingegneria dell'informazione dovrebbe comunque avere già familiarità. Le nozioni matematiche riportate nelle altre appendici possono essere sviluppate durante il corso all'occorrenza.

Ringraziamenti

Questo testo deriva dalle dispense che ho preparato per corsi istituzionali o interventi seminariali negli ultimi 20 anni. I primi capitoli nascono, in forma embrionale, nel 1997 e si sono poi evoluti ed ampliati fino alla versione attuale. Ringrazio gli studenti dell'Università di Udine e dell'Università di Verona che, lungo questi anni, hanno segnalato errori, mancanze, e parti poco chiare.

Il testo ha beneficiato inoltre delle puntuali correzioni suggerite da Federica Arrigoni, Guido Maria Cortelazzo, Fabio Crosilla, Riccardo Gherardi, Luca Magri, Francesco Malapelle, Samuele Martelli e Roberto Toldo, che ringrazio sentitamente.

I crediti per le figure sono riconosciuti nelle rispettive didascalie.

1. Introduzione

1.1. Il prodigio della visione

Se ci fermiamo a riflettere in modo distaccato sulla visione come abilità sensoriale, dobbiamo convenire con Ullman che essa è prodigiosa:

«As seeing agents, we are so used to the benefits of vision, and so unaware of how we actually use it, that it took a long time to appreciate the almost miraculous achievements of our visual system. If one tries to adopt a more objective and detached attitude, by considering the visual system as a device that records a band of electromagnetic radiation as an input, and then uses it to gain knowledge about surrounding objects that emit and reflect it, one cannot help but be struck by the richness of information this system provides.» (Ullman, 1996)

La visione artificiale nasce come branca dell'intelligenza artificiale negli anni Settanta del secolo scorso e si sviluppa successivamente come disciplina autonoma con proprie metodologie, paradigmi e problemi. Nell'approccio moderno essa non si propone di replicare la visione umana. Il tentativo sarebbe infatti probabilmente destinato al fallimento per l'intrinseca differenza tra i due *hardware*. Il paragone che viene spesso evocato è quello del volo: gli sforzi per replicare il volo animale nella storia sono tutti falliti, mentre gli aeroplani, con un approccio completamente diverso, hanno risolto il problema in modo più che soddisfacente, superando in alcuni aspetti gli uccelli stessi.

1.2. Visione computazionale di basso livello

La *visione computazionale* (o *visione artificiale* o *computer vision*) è la disciplina dell'informatica che si occupa dell'estrazione di informazioni dalle immagini. Le informazioni possono essere di natura numerica (per esempio coordinate spaziali) o simbolica (per esempio identità e relazioni tra oggetti). Semplificando, potremmo dire che si tratta di scoprire *cosa* è presente nella scena e *dove*. Essa ha profonde connessioni con la robotica, la *pattern recognition*, l'elaborazione delle immagini e la fotogrammetria.

Seguendo Ullman (1996) si usa distinguere tra visione di *basso livello* e di *alto livello*. La prima si occupa di estrarre determinate proprietà fisiche dell'ambiente visibile, come profondità, forma tridimensionale, contorni degli oggetti. I processi di visione di *basso livello* sono tipicamente paralleli, spazialmente uniformi e relativamente indipendenti dal problema e dalla conoscenza a priori associata a particolari oggetti.

Viceversa, la visione di *alto livello* si occupa dell'estrazione di proprietà delle forme, di relazioni spaziali, di riconoscimento e classificazione di oggetti. I processi di alto livello sono di solito applicati ad una porzione dell'immagine, dipendono dall'obiettivo della computazione e dalla conoscenza a priori associata agli oggetti.

Anche tralasciando i problemi di alto livello legati alla percezione e al riconoscimento di oggetti, il solo compito di ricostruire un modello geometrico della scena è difficile. Di questo ci occuperemo in questo libro: studieremo metodi computazionali (algoritmi) che mirano ad ottenere una rappresentazione della struttura solida (*sterèos*) del mondo tridimensionale esperito attraverso proiezioni bidimensionali di esso, le immagini.

Questo approccio al problema si può agevolmente inquadrare all'interno del *ricostruzionismo*, la teoria sviluppata da Marr alla fine degli anni Settanta:

«In the theory of visual processes, the underlying task is to reliably derive properties of the world from images of it; the business of isolating constraints that are both powerful enough to allow a process to be defined and generally true of the world is a central theme of our inquiry.»
(Marr, 1982)

Questo, naturalmente, non è il solo paradigma possibile in visione. Aloimonos e Shulman (1989), per esempio, sostengono che la descrizione del mondo non debba essere generica, ma dipendente dall'obiettivo.

La visione computazionale (di basso livello) si può efficacemente descrivere come *l'inverso della grafica al calcolatore*, nella quale, dati:



Figura 1.1. – Relazione tra visione e grafica

- la descrizione geometrica della scena,
 - la descrizione fotometrica della scena (fonti luminose e proprietà delle superfici),
 - la descrizione completa dell'apparato di acquisizione (fotocamera),
- il calcolatore produce l'immagine "sintetica" vista dalla fotocamera.

La riduzione dimensionale operata dalla proiezione (geometria) e la molteplicità delle cause che concorrono a determinare la luminosità ed il colore rendono il problema inverso *sottovincolato* e di non banale soluzione.

Il sistema visivo umano sfrutta molteplici "indizi visivi" (*visual cues*) per risolvere il problema. Alcuni di questi indizi sono: la disparità binoculare, il moto, il chiaroscuro (o *shading*), la sfocatura, le tessiture.

Bibliografia

- Aloimonos J.; Shulman D. (1989). *Integration of Visual Modules. An Extension to the Marr Paradigm*. Academic Press, Waltham, MA.
- Marr D. (1982). *Vision*. Freeman, San Francisco, CA.
- Ullman D. (1996). *High-level Vision*. The MIT Press, Cambridge, MA.