

Αοι



Arcangelo Distante  
Cosimo Distante

**Visione computazionale  
Dall'energia all'immagine**

Volume I

*Prefazione di*  
Virginio Cantoni





Aracne editrice

[www.aracneeditrice.it](http://www.aracneeditrice.it)  
[info@aracneeditrice.it](mailto:info@aracneeditrice.it)

Copyright © MMXVIII  
Giacchino Onorati editore S.r.l. – unipersonale

[www.giacchinoonoratieditore.it](http://www.giacchinoonoratieditore.it)  
[info@giacchinoonoratieditore.it](mailto:info@giacchinoonoratieditore.it)

via Vittorio Veneto, 20  
00020 Canterano (RM)  
(06) 4551463

ISBN 978-88-255-1528-2

*I diritti di traduzione, di memorizzazione elettronica,  
di riproduzione e di adattamento anche parziale,  
con qualsiasi mezzo, sono riservati per tutti i Paesi.*

*Non sono assolutamente consentite le fotocopie  
senza il permesso scritto dell'Editore.*

I edizione: maggio 2018

*Ai miei genitori ed alla mia famiglia, Maria e Maria Grazia – AD*

*Ai miei genitori, a mia moglie Giovanna e, ai miei figli Francesca e Davide – CD*



# Indice

<b>Prefazione</b> di Virginio Cantoni	15
<b>Introduzione</b>	17
<b>1 Processo di formazione dell'Immagine</b>	<b>25</b>
1.1 Introduzione . . . . .	25
1.2 Dall'energia all'immagine . . . . .	26
1.3 Energia elettromagnetica, Fotoni e Luce . . . . .	27
1.3.1 Caratteristica delle onde elettromagnetiche . . . . .	28
1.4 L'energia delle onde elettromagnetiche . . . . .	31
1.5 Sorgenti di onde elettromagnetiche . . . . .	32
1.6 Interazione luce-materia . . . . .	33
1.7 Fotoni . . . . .	34
1.8 Propagazione delle onde elettromagnetiche nella materia . . . . .	37
1.9 Lo spettro delle radiazioni elettromagnetiche . . . . .	38
1.10 La Luce . . . . .	41
1.10.1 Propagazione della Luce . . . . .	42
1.10.2 Riflessione e Rifrazione . . . . .	44
1.11 La fisica della luce . . . . .	49
1.12 Energia di un'onda elettromagnetica . . . . .	50
1.13 Riflettanza e trasmittanza . . . . .	53
1.13.1 Angolo di Brewster . . . . .	61
1.13.2 Riflessione interna . . . . .	63
1.14 Radiazioni termiche . . . . .	65
<b>2 Modello Radiometrico</b>	<b>77</b>
2.1 Introduzione . . . . .	77
2.2 Sorgenti luminose ed aspetti radiometrici . . . . .	79
2.3 Funzione di distribuzione di riflettanza bidirezionale BRDF . . . . .	83
2.3.1 Modello Lambertiano . . . . .	87
2.3.2 Modello di riflettanza speculare . . . . .	90
2.3.3 Modello di riflettanza composto Lambertiano-Speculare . . . . .	91

2.3.4	Modello di Phong . . . . .	92
2.4	Equazione fondamentale nel processo di formazione dell'immagine	93
<b>3</b>	<b>Il Colore</b>	<b>99</b>
3.1	Introduzione . . . . .	99
3.1.1	La teoria della percezione del colore . . . . .	101
3.2	Il sistema visivo umano . . . . .	107
3.3	Fenomeni visivi: sensibilità al contrasto . . . . .	116
3.4	Fenomeni visivi: contrasto simultaneo . . . . .	118
3.5	Fenomeni visivi: bande di Mach . . . . .	119
3.6	Fenomeni visivi: Daltonismo . . . . .	120
3.7	I colori della natura . . . . .	121
3.8	Costanza del colore . . . . .	126
3.9	Colorimetria . . . . .	129
3.9.1	Metamerismo e Leggi di Grassmann . . . . .	133
3.9.1.1	Il Metamerismo . . . . .	135
3.9.1.2	Leggi di Grassmann . . . . .	138
3.10	Metodo della sintesi additiva . . . . .	139
3.10.1	Curve tristimolo di eguale radianza . . . . .	143
3.10.1.1	Esempio: Calcolare le componenti (R,G,B) normalizzate . . . . .	145
3.10.2	Coordinate di cromaticità . . . . .	146
3.11	Rappresentazione 3D del colore RGB . . . . .	147
3.12	Coordinate di colore XYZ . . . . .	148
3.13	Diagramma di cromaticità - <b>RGB</b> . . . . .	151
3.14	Diagramma di cromaticità - Spazio <b>XYZ</b> . . . . .	154
3.14.0.1	Esempio: Calcolare le posizioni dei primari RGB nel diagramma di cromaticità xy . . . . .	156
3.14.1	Analisi della trasformazione da RGB al sistema XYZ . . . . .	158
3.15	Rappresentazione geometrica del colore . . . . .	160
3.16	Spazio di colore HSI . . . . .	163
3.17	Il colore nella elaborazione delle immagini . . . . .	166
3.18	Dallo spazio RGB allo spazio HSI . . . . .	167
3.18.1	Dallo spazio RGB $\rightarrow$ HSI . . . . .	168
3.18.2	Dallo spazio HSI $\rightarrow$ RGB . . . . .	170
3.19	Spazio di colore HSV e HLS . . . . .	170
3.20	Spazio del colore CIE 1964 UCS . . . . .	172
3.21	Spazio del colore CIE 1976 $L^*a^*b^*$ . . . . .	175
3.22	Spazio del colore CIE 1976 $L^*u^*v^*$ . . . . .	177
3.23	Spazio del colore CIELab LCh e CIELuv LCh . . . . .	177
3.24	Spazio del colore YIQ . . . . .	178
3.25	Metodo della sintesi sottrattiva . . . . .	179
3.26	Tecnologie di riproduzione del colore . . . . .	185
3.27	Riepilogo e Conclusioni . . . . .	191

<b>4</b>	<b>Sistema ottico</b>	<b>195</b>
4.1	Introduzione . . . . .	195
4.2	Riflessione della luce su specchi sferici . . . . .	197
4.3	Rifrazione della luce su superfici sferiche . . . . .	202
4.4	Lenti sottili . . . . .	208
4.4.1	Diagramma dei raggi principali per lenti sottili . . . . .	210
4.4.2	Ingrandimento ottico: Microscopio e Telescopio . . . . .	217
4.5	Aberrazioni ottiche . . . . .	221
4.5.1	Parametri di un sistema ottico . . . . .	225
<b>5</b>	<b>Digitalizzazione e visualizzazione immagine</b>	<b>229</b>
5.1	Introduzione . . . . .	229
5.2	Il sistema ottico umano . . . . .	231
5.3	Sistemi di acquisizione immagini . . . . .	234
5.4	Rappresentazione dell'immagine digitale . . . . .	240
5.5	Risoluzione e frequenza spaziale . . . . .	241
5.6	Modello geometrico della formazione dell'immagine . . . . .	245
5.7	Formazione dell'immagine con un sistema ottico reale . . . . .	247
5.8	Risoluzione del sistema ottico . . . . .	255
5.8.1	Funzione di Modulazione del Contrasto - MTF . . . . .	259
5.9	Funzione di Trasferimento Ottico (OTF) . . . . .	263
5.10	Campionamento . . . . .	265
5.11	Quantizzazione . . . . .	275
5.12	Sistemi di Acquisizione Immagini Digitali - SAID . . . . .	280
5.12.1	Campo di vista - FOV ( <i>Field Of View</i> ) . . . . .	283
5.12.2	Focale $f$ del sistema ottico . . . . .	284
5.12.3	Risoluzione spaziale dell'ottica . . . . .	285
5.12.4	Dimensione e risoluzione spaziale del sensore . . . . .	285
5.12.5	Risoluzione temporale del sensore . . . . .	286
5.12.6	Profondità di campo e di fuoco . . . . .	286
5.12.6.1	Calcolo del diametro di CoC . . . . .	288
5.12.7	Calcolo della profondità di campo . . . . .	290
5.12.8	Calcolo della iperfocale . . . . .	292
5.12.9	Profondità di fuoco . . . . .	292
5.12.10	Macchina fotografica . . . . .	294
5.12.11	Telecamera . . . . .	297
5.12.12	Termocamere . . . . .	300
5.12.13	Telecamere ToF . . . . .	301
5.13	Microscopia . . . . .	310
5.14	Telescopio . . . . .	312
5.15	La funzione MTF di un sistema di acquisizione . . . . .	313

<b>6</b>	<b>Proprietà dell'immagine digitale</b>	<b>317</b>
6.1	Immagine digitale binaria . . . . .	317
6.2	Vicinanza di pixel . . . . .	318
6.3	Metrica delle immagini . . . . .	319
6.3.1	Distanza Euclidea $D_E$ . . . . .	319
6.3.2	4-Distanza ( <i>city block</i> ) . . . . .	320
6.3.3	8-Distanza ( <i>chessboard</i> ) . . . . .	320
6.4	Trasformata distanza . . . . .	321
6.5	Percorso ( <i>path</i> ) . . . . .	324
6.6	Connettività . . . . .	324
6.7	Regione . . . . .	325
6.7.1	Componenti connesse . . . . .	326
6.7.2	Foreground background ed holes . . . . .	326
6.7.3	Oggetto . . . . .	327
6.7.4	Contorni . . . . .	327
6.7.5	Bordi ( <i>Edges</i> ) . . . . .	328
6.8	Proprietà topologiche dell'immagine . . . . .	329
6.8.1	Numero di Eulero . . . . .	329
6.8.2	Involucro convesso ( <i>convex hull</i> ) . . . . .	330
6.8.3	Area, perimetro e compattezza . . . . .	330
6.9	Proprietà indipendente dalla posizione del pixel . . . . .	331
6.9.1	Istogramma . . . . .	332
6.10	Proprietà dipendente dalla correlazione tra pixel . . . . .	333
6.10.1	L'immagine come processo stocastico . . . . .	333
6.10.2	Misura di correlazione . . . . .	334
6.11	Qualità dell'immagine . . . . .	335
6.11.1	Rumore dell'immagine . . . . .	337
6.11.2	Rumore Gaussiano . . . . .	337
6.11.3	Rumore salt-and-pepper . . . . .	337
6.11.4	Rumore impulsivo . . . . .	338
6.11.5	Gestione del rumore . . . . .	338
6.12	Informazioni percettive dell'immagine . . . . .	339
6.12.1	Contrasto . . . . .	339
6.12.2	Acutezza . . . . .	339
<b>7</b>	<b>Organizzazione dei dati</b>	<b>341</b>
7.1	Dati nei diversi livelli di elaborazione . . . . .	341
7.2	Strutture dei dati . . . . .	342
7.2.1	Matrici . . . . .	342
7.2.2	Matrice di co-occorrenza . . . . .	343
7.2.2.1	Algoritmo per il calcolo della matrice di co-occorrenza . . . . .	344
7.3	Codifica dei contorni (Chain Code) . . . . .	348
7.4	Codifica Run-Length . . . . .	349

7.4.1	Codifica Run Length code per immagini a livello di grigio ed a colori . . . . .	351
7.5	Organizzazione topologica dei dati - Grafi . . . . .	352
7.5.1	Region Adiacency Graph (RAG) . . . . .	352
7.5.2	Caratteristiche del RAG . . . . .	352
7.5.3	Algoritmo per costruire RAG . . . . .	353
7.5.4	Organizzazione relazionale . . . . .	353
7.6	Struttura gerarchica dei dati . . . . .	353
7.6.1	Pyramids . . . . .	355
7.6.1.1	M-piramide (Matrice-Piramide) . . . . .	355
7.6.2	Quadtree . . . . .	355
7.6.2.1	T-Piramide . . . . .	358
7.6.3	Piramide Gaussiana e Laplaciana . . . . .	358
7.6.4	Octree . . . . .	360
7.6.5	Operazioni su Quadtree e Octree . . . . .	361
<b>8</b>	<b>Rappresentazione e Descrizione delle Forme</b>	<b>363</b>
8.1	Introduzione . . . . .	363
8.2	Rappresentazione esterna degli oggetti . . . . .	364
8.2.0.1	Chain Code . . . . .	365
8.2.0.2	Approssimazione poligonale - perimetro . . . . .	367
8.2.0.3	Approssimazione poligonale - splitting . . . . .	368
8.2.0.4	Approssimazione poligonale - merging . . . . .	369
8.2.0.5	Approssimazione contorno con segmenti curvi . . . . .	369
8.2.0.6	Impronta (Signature) . . . . .	370
8.2.1	Rappresentazione mediante involucro convesso . . . . .	372
8.2.2	Rappresentazione mediante scheletrizzazione . . . . .	372
8.2.2.1	Algoritmo di assottigliamento (thinning) . . . . .	376
8.3	Descrizione delle forme . . . . .	378
8.3.1	Descrittori elementari di forme . . . . .	379
8.3.1.1	Perimetro . . . . .	379
8.3.1.2	Area . . . . .	380
8.3.1.3	Compattezza . . . . .	381
8.3.1.4	Asse maggiore . . . . .	382
8.3.1.5	Asse minore . . . . .	383
8.3.1.6	Rettangolo di base . . . . .	383
8.3.1.7	Rettangolarità . . . . .	383
8.3.1.8	Eccentricità . . . . .	384
8.3.1.9	Elongazione . . . . .	384
8.3.1.10	Riepilogo dei descrittori elementari di forme . . . . .	384
8.3.2	Momenti . . . . .	384
8.3.2.1	Invarianza alla traslazione . . . . .	386
8.3.2.2	Invarianza alla scala . . . . .	386
8.3.2.3	Invarianza all'orientamento dell'oggetto . . . . .	387
8.3.2.4	Invarianza per traslazione, rotazione e scala . . . . .	392

8.3.2.5	Momenti del contorno . . . . .	394
8.3.3	Momenti basati su funzioni basi ortogonali . . . . .	397
8.3.4	Descrittori di Fourier . . . . .	399
8.3.4.1	Invarianza alla traslazione . . . . .	400
8.3.4.2	Invarianza alla scala . . . . .	401
8.3.4.3	Invarianza alla rotazione . . . . .	401
8.3.4.4	Invarianza rispetto al punto di partenza . . . . .	401
8.3.4.5	Descrittori FD con la funzione distanza . . . . .	402
8.3.4.6	Normalizzazione dei descrittori di Fourier . . . . .	403
8.3.4.7	Descrittori di Fourier region-based . . . . .	404
8.3.4.8	Considerazioni implementative . . . . .	405
<b>9</b>	<b>Miglioramento dell'Immagine</b>	<b>407</b>
9.1	Introduzione ai livelli computazionali . . . . .	407
9.2	Miglioramento della qualità dell'immagine - Image Enhancement	409
9.2.1	Istogramma . . . . .	410
9.2.2	Funzione densità di probabilità e funzione cumulativa dell'immagine . . . . .	410
9.2.3	Manipolazione del contrasto . . . . .	412
9.2.3.1	Lineare con unico tratto . . . . .	413
9.2.3.2	Lineare con più tratti . . . . .	413
9.2.3.3	Non lineare . . . . .	415
9.2.3.4	Negativa e inversa . . . . .	416
9.2.3.5	Non monotona . . . . .	416
9.2.3.6	Manipolazione bit-plane . . . . .	416
9.2.3.7	Divisione dei livelli di intensità (Intensity Slice)	417
9.2.4	Trasformazione gamma . . . . .	418
9.3	Modifica dell'istogramma . . . . .	420
9.3.1	Equalizzazione dell'istogramma . . . . .	421
9.3.1.1	Esempio 1: Operatore puntuale lineare . . . . .	424
9.3.1.2	Esempio 2: Operatore puntuale non lineare . . . . .	425
9.3.1.3	Esempio 3: Casi singolari . . . . .	426
9.3.2	Equalizzazione adattiva dell'istogramma (AHE) . . . . .	426
9.3.3	Equalizzazione adattiva dell'istogramma con contrasto limitato-CLAHE . . . . .	428
9.4	Modifica con modello predefinito dell'istogramma . . . . .	430
9.5	Operazioni puntuali omogenee . . . . .	432
9.6	Operazioni puntuali non omogenee . . . . .	433
9.6.1	Operatore puntuale per correggere l'errore radiometrico	433
9.6.2	Operatore locale statistico . . . . .	434
9.7	Miglioramento della qualità per le immagini a colori . . . . .	435
9.7.1	Immagini a colori naturali . . . . .	435
9.7.2	Immagini a pseudo-colore . . . . .	435
9.7.3	Immagini a falso colore . . . . .	436
9.8	Miglioramento qualità delle immagini multispettrali . . . . .	437

9.9	Verso gli Operatori locali e globali . . . . .	438
9.9.1	Filtraggio numerico spaziale . . . . .	440
9.9.1.1	Linearità . . . . .	440
9.9.1.2	Risposta impulsiva e PSF-Point Spread Function	441
9.9.1.3	Invarianza spaziale . . . . .	443
9.9.1.4	Convoluzione . . . . .	444
9.9.1.5	Proprietà della convoluzione . . . . .	445
9.9.1.6	Riepilogo . . . . .	446
9.10	Convoluzione spaziale . . . . .	447
9.10.1	Convoluzione spaziale monodimensionale . . . . .	447
9.10.1.1	Gestione dei bordi nella convoluzione 1D . . .	451
9.10.2	Convoluzione spaziale bidimensionale . . . . .	452
9.10.2.1	Estensione della risposta all'impulso . . . . .	454
9.10.2.2	Gestione dei bordi nella convoluzione 2D . . .	455
9.10.2.3	Complessità computazionale . . . . .	455
9.10.2.4	Peculiarità della maschera di convoluzione . .	456
9.10.2.5	Maschera simmetrica . . . . .	458
9.10.2.6	Maschera simmetrica circolare . . . . .	459
9.10.2.7	Separabilità della convoluzione . . . . .	459
9.10.2.8	Complessità computazionale della convoluzione 2D separabile . . . . .	461
9.11	Filtraggio nel dominio delle frequenze . . . . .	462
9.11.1	La Trasformata Discreta di Fourier DFT . . . . .	462
9.11.1.1	Magnitudo, Angolo di fase e Spettro di potenza	463
9.11.1.2	La DFT nella elaborazione delle immagini . .	464
9.11.1.3	Separabilità della DFT . . . . .	469
9.11.1.4	Algoritmo FFT . . . . .	469
9.11.2	Risposta in frequenza di un sistema lineare . . . . .	470
9.11.3	Teorema di convoluzione . . . . .	471
9.11.3.1	Convoluzione in frequenza . . . . .	472
9.11.3.2	Estensione a segnali 2D . . . . .	473
9.11.3.3	Scelta del dominio di filtraggio . . . . .	473
9.12	Operatori locali: smoothing . . . . .	475
9.12.1	Media aritmetica . . . . .	476
9.12.2	Filtro media . . . . .	476
9.12.3	Filtri non lineari . . . . .	480
9.12.4	Filtro mediano . . . . .	481
9.12.5	Filtro di minimo e massimo . . . . .	481
9.12.6	Filtro di smoothing Gaussiano . . . . .	482
9.12.6.1	Simmetria circolare . . . . .	483
9.12.6.2	Proprietà di separabilità del filtro Gaussiano .	483
9.12.6.3	Progettazione del filtro Gaussiano discreto . .	484
9.12.7	Filtri binomiali . . . . .	486
9.12.8	Analisi computazionale dei filtri di smoothing . . . . .	488
9.13	Filtraggio passa-basso nel dominio di Fourier . . . . .	489

9.13.1 Filtro ideale passa-basso . . . . .	490
9.13.2 Filtro di Butterworth . . . . .	493
9.13.3 Filtro Gaussiano passa-basso . . . . .	493
9.13.4 Filtro trapezoidale . . . . .	498
9.13.5 Riepilogo risultati dei filtri di smoothing . . . . .	499
<b>Bibliography</b>	<b>502</b>
<b>Ringraziamenti</b>	<b>509</b>

# Prefazione

di Virginio Cantoni

Monumentale. Come altro definire l'opera dei due autori? Chiariamo subito che il termine non riguarda l'ampiezza e la qualità dell'opera – oltre 1250 pagine ricche di immagini curate e ben organizzate – ma soprattutto la completezza della trattazione e la descrizione dell'evolversi esplosivo di una disciplina che oggi, finalmente, va occupando il punto focale e critico di tanti settori vitali per la società moderna toccando praticamente tutti i settori applicativi con bioimmagini, remote sensing, assistenza telematica, accessibilità tattile, guida automatica, data mining and video annotation solo per citarne alcune applicazioni già a livello almeno preindustriale. I mattoni su cui si è venuta costruendo l'attuale visione computazionale sono presentati in dettaglio. Modelli diversi partendo dai principi fisici della formazione del segnale visivo, considerando i diversi metodi di acquisizione, i loro limiti e le strategie per migliorarne la qualità alla generazione sulla base della conoscenza delle caratteristiche di trasduzione e conversione in immagine digitale. Segue la trattazione di quella parte della disciplina che un tempo è stata schematicamente chiamata da Michael Duff la visione intermedia. La fase di elaborazione che viene dopo la percezione e prima dell'elaborazione di più "alto" livello con obiettivi concettualmente più elevati allora considerati prossimi al ragionamento umano con l'obiettivo di interpretare il contenuto semantico del segnale visivo. In modo che definirei esaustivo sono quindi presentati i molteplici approcci di rappresentazione e manipolazione dell'informazione visiva in forma digitale: strutture dati (dalla matrice 2D alle strutture piramidali o multigrid); trasformazioni lineari filtraggi di tutti i tipi (nello spazio, in frequenza, ecc.); proprietà peculiari, invarianze, descrittori e rilevatori di varia natura e precisione. Sono quindi introdotti in modo dettagliato i vari approcci alla segmentazione cioè l'identificazione, la localizzazione e la forma delle componenti salienti dell'immagine che da questo momento perde le caratteristiche strutturali di partenza. Infine con il terzo volume si affronta l'analisi del contenuto informativo, lo studio dei componenti salienti, di fatto la semantica del segnale percepito. Accanto ai metodi di successo che si sono affermati specificatamente nel riconoscimento di forme draconianamente descritti in letteratura come approcci statistici, sintattici e strutturali uno spazio significativo è dato all'approccio basato sulle reti neurali che avviate da McCulloch e Pitts (vorrei però ricordare che espressamente nel loro lavoro citato parlano del modello di neurone descritto nella tesi di laurea da Eduardo Caianiello, poi

fondatore del gruppo italiano di riconoscimento di forme) che oggi sono alla base del deep learning. Questa evoluzione della machine learning sta portando l'apprendimento delle macchine verso traguardi impensati raggiungendo risultati efficienti in settori della visione computerizzata che con i vecchi approcci erano proibitivi. Infine voglio sottolineare il costante riferimento che gli autori portano verso le soluzioni della visione umana, forse non sono più i tempi di David Marr nei quali la neurofisiologia poteva offrire modelli di elaborazione del segnale visivo facilmente applicabili per le soluzioni automatiche ma ai livelli di elaborazione più alti le soluzioni antropomorfe ancora possono ispirare nuovi modelli concettuali!

*Virginio Cantoni*

IAPR e IEEE Fellow

Dipartimento di Ingegneria Industriale e dell'Informazione

Università di Pavia

# Introduzione

Negli ultimi 20 anni diverse ricerche interdisciplinari nei settori della *fisica, dell'informatica e della cibernetica, dell'Elaborazione numerica di Segnali ed Immagini, delle tecnologie elettriche ed elettroniche*, hanno portato allo sviluppo di Sistemi Intelligenti.

I cosiddetti *Sistemi Intelligenti (o Agenti Intelligenti)* rappresentano la frontiera tuttora più avanzata e innovativa della ricerca in campo informatico ed elettronico, in grado di influenzare direttamente la qualità della vita, la competitività e le modalità di produzione delle imprese, di monitorare e valutare l'impatto ambientale, di rendere più efficienti le attività di servizio e di gestione delle amministrazioni pubbliche, e la sicurezza delle persone.

Lo studio di un sistema intelligente, indipendentemente dal campo d'impiego, può essere semplificato in tre componenti essenziali:

1. la *prima* interagisce con l'ambiente per l'acquisizione dei dati del dominio di interesse, utilizzando sensori adeguati (per l'acquisizione di Segnali ed Immagini);
2. la *seconda* analizza ed interpreta i dati rilevati dalla prima componente utilizzando anche tecniche di apprendimento per costruire/aggiornare rappresentazioni adeguate della realtà anche complessa nella quale il sistema opera (Visione Computazionale);
3. la *terza* sceglie le azioni più appropriate per raggiungere gli obiettivi assegnati al sistema intelligente (scelta dei Modelli Decisionali Ottimali) interagendo con le prime due componenti, e con gli operatori umani, nel caso di soluzioni applicative basate su paradigmi cooperativi uomo-macchina (le attuali evoluzioni dell'automazione compresa quella industriale).

In questo scenario di avanzamento della conoscenza per lo sviluppo di Sistemi Intelligenti, s'inquadra il contenuto informativo di questo manoscritto nel quale sono riportate le esperienze di ricerca pluriennali e di docenza degli autori, e degli approfondimenti scientifici esistenti in letteratura. In particolare, il manoscritto articolato in tre parti (volumi), tratta gli aspetti del sottosistema sensoriale per percepire l'ambiente in cui un sistema intelligente è immerso e capace di agire anche autonomamente.

Il *primo volume* descrive l'insieme dei processi fondamentali della visione artificiale che dall'energia portano alla formazione dell'immagine digitale. So-

no analizzati i fenomeni di propagazione della luce (cap.1 e 2), la teoria di percezione del colore (cap.3), l'impatto del sistema ottico (cap.4), gli aspetti di trasduzione da energia luminosa (il flusso ottico) a segnale elettrico (dei fotorecettori), e gli aspetti di trasduzione del segnale elettrico (con valori continui) in valori discreti (pixel), ovvero la conversione del segnale da analogico a digitale (cap.5). Questi primi 5 capitoli sintetizzano il processo di acquisizione della scena 3D, in forma simbolica, rappresentata numericamente dai pixel dell'immagine digitale (proiezione 2D della scena 3D).

Il cap.6 descrive le proprietà geometriche, topologiche, qualità e informazioni percettive dell'immagine digitale. Sono definite le metriche, le modalità di aggregazione e di correlazione tra pixel, utili per definire strutture simboliche della scena di più alto livello rispetto al pixel. L'organizzazione dei dati per i diversi livelli di elaborazione è descritta nel cap.7 mentre nel cap.8 è riportata la rappresentazione e descrizione delle strutture omogenee della scena.

Con il cap.9 inizia la descrizione degli algoritmi di elaborazione immagine, per il miglioramento delle qualità visive dell'immagine, basati su operatori puntuali, locali e globali. Sono riportati algoritmi operanti nel dominio spaziale e nel dominio delle frequenze evidenziando con degli esempi le differenze significative tra i vari algoritmi anche dal punto di vista del carico computazionale.

Il *secondo volume* inizia con il capitolo che descrive gli algoritmi di estrazione dei contorni basati su operatori locali nel dominio spaziale e su tecniche di filtraggio nel dominio delle frequenze.

Nel cap.2 sono presentate le trasformate lineari fondamentali che hanno immediata applicazione nel campo dell'elaborazione dell'immagine, in particolare, per estrarre le caratteristiche essenziali contenute nelle immagini. Tali caratteristiche, che sintetizzano efficacemente il carattere informativo globale dell'immagine, sono poi utilizzate per gli altri processi di elaborazione dell'immagine: classificazione, compressione, descrizione, ecc. Le trasformate lineari sono anche utilizzate, come operatori globali, per migliorare le qualità visive dell'immagine (*enhancement*), per attenuare il rumore (*restoration*), oppure per ridurre la dimensionalità dei dati (*data reduction*).

Nel cap.3 sono descritte le trasformazioni geometriche delle immagini, necessarie in diverse applicazioni della visione artificiale, sia per correggere eventuali distorsioni geometriche introdotte durante l'acquisizione (per esempio, immagini acquisite mentre gli oggetti oppure i sensori sono in movimento, come nel caso di acquisizioni da satellite e/o aereo), oppure per introdurre voluti effetti geometrici visivi. In tutte e due i casi, l'operatore geometrico deve essere in grado di riprodurre nel modo più fedele possibile l'immagine con lo stesso contenuto informativo iniziale attraverso il processo di ricampionamento dell'immagine.

Nel cap.4 *Ricostruzione dell'immagine degradata (image restoration)* sono descritte un insieme di tecniche che eseguono correzioni quantitative sull'immagine per compensare le degradazioni introdotte durante il processo di acquisizione e trasmissione. Tali degradazioni sono rappresentate dall'effetto nebbia o sfocamento causato dal sistema ottico e dal moto dell'oggetto o dell'osservato-

re, dal rumore causato dal sistema opto-elettronico e dalla risposta non lineare dei sensori, dal rumore casuale dovuto alla turbolenza atmosferica o, più in generale, dal processo di digitalizzazione e trasmissione. Mentre le tecniche di *enhancement*, tendono a ridurre in misura qualitativa, le degradazioni presenti nell'immagine, migliorandone la qualità visiva anche quando non si ha nessuna conoscenza del modello di degradazione, le tecniche di *restoration* sono utilizzate invece per eliminare o attenuare in modo quantitativo le degradazioni presenti nell'immagine, partendo anche dalle ipotesi di conoscenza dei modelli di degradazione.

Il cap.5, *Segmentazione delle immagini*, descrive diversi algoritmi di segmentazione, ovvero il processo di dividere l'immagine in regioni omogenee, dove tutti i pixel che corrispondono ad un oggetto della scena sono raggruppati insieme. Il raggruppamento dei pixel in regioni è basato in relazione ad un criterio di omogeneità che li distingue tra loro. Sono riportati algoritmi di segmentazione basati su criteri di similarità degli attributi dei pixel (colore, tessitura, ecc.) oppure basati su criteri geometrici di prossimità spaziale dei pixel (distanza Euclidea, ecc.). Tali criteri non sono sempre validi, e in diverse applicazioni è necessario integrare altre informazioni in relazione alla conoscenza a priori del contesto applicativo (dominio dell'applicazione). Il raggruppamento dei pixel, in quest'ultimo caso, si basa confrontando le regioni ipotizzate con le regioni modellate a priori.

Il cap.6 *Rilevatori e descrittori di punti di interesse*, descrive gli algoritmi più utilizzati per rilevare automaticamente le strutture significative (noti come punti di interesse, corner, feature) presenti nell'immagine corrispondenti a parti fisiche stabili della scena. L'abilità di tali algoritmi è quella di rilevare ed identificare parti fisiche della stessa scena in modo ripetibile, anche quando le immagini sono acquisite in condizioni di variabilità di illuminazione e cambiamento del punto di osservazione con eventuale cambiamento del fattore di scala.

*Il terzo volume* descrive gli algoritmi di visione artificiale che rilevano gli oggetti della scena, tentano la loro identificazione, la ricostruzione 3D, il loro assetto e localizzazione rispetto all'osservatore, e l'eventuale loro movimento.

Il cap.1 *Riconoscimento degli oggetti*, descrive gli algoritmi fondamentali della visione artificiale per riconoscere automaticamente gli oggetti della scena caratteristica essenziali di tutti i sistemi di visione degli organismi viventi. Mentre un osservatore umano esegue il riconoscimento anche di oggetti complessi, apparentemente in modo agevole e tempestivo, per una macchina di visione il processo di riconoscimento risulta difficile, necessita di un notevole tempo di calcolo e non sempre i risultati sono ottimali. Fondamentali per il processo di riconoscimento degli oggetti, diventano gli algoritmi di selezione ed estrazione delle feature. In varie applicazioni si può avere una conoscenza a priori della popolazione degli oggetti da classificare poichè si conoscono i pattern (feature significative) campioni dai quali si possono estrarre informazioni utili per la decisione di associare (*decision making*) ciascun individuo della popolazione ad una determinata classe. Questi pattern campioni (*training set*)

sono utilizzati dal sistema di riconoscimento per apprendere le informazioni significative sulla popolazione (estrazione dei parametri statistici, caratteristiche rilevanti, ecc.). Il processo di riconoscimento confronta le feature degli oggetti incogniti con le feature dei pattern modelli, allo scopo di identificarne in modo univoco la classe di appartenenza. Diversi sono stati negli anni i settori disciplinari (apprendimento automatico, analisi d'immagini, riconoscimento oggetti, ricerca dell'informazione, bioinformatica, biomedicina, analisi di dati intelligente, data mining, ...) ed i settori applicativi (robotica, telerilevamento, visione artificiale,...) per i quali diversi ricercatori hanno proposto diversi metodi di riconoscimento e sviluppato differenti algoritmi basati su diversi modelli di classificazione. Sebbene gli algoritmi proposti hanno un'univoca finalità, si differenziano per la proprietà attribuita alle classi di oggetti (i cluster) e per il modello con cui sono definiti tali classi (connettività, distribuzione statistica, densità,...). La diversità di discipline, specialmente tra quelle dell'estrazione automatica dei dati (*data mining*) e dell'apprendimento automatico (*machine learning*), ha portato a sottili differenze, soprattutto nell'utilizzo dei risultati e nelle terminologie, talvolta contraddittorie forse causate dai diversi obiettivi. Per esempio, nel data mining l'interesse dominante è l'estrazione automatica dei raggruppamenti, nella classificazione automatica è fondamentale il potere discriminante delle classi di appartenenza dei pattern. Gli argomenti di questo capitolo si sovrappongono tra aspetti legati al machine learning e quelli del riconoscimento basato su metodi statistici. Per semplicità gli algoritmi descritti sono ripartiti in base ai metodi di classificazione degli oggetti in supervisionati (basati su modelli deterministici, statistici, neurali, e non metrici quali i modelli sintattici e alberi decisionali) e non-supervisionati, ovvero metodi che non utilizzano nessuna conoscenza a priori per estrarre le classi di appartenenza dei pattern.

Nel cap.2 *Reti neurali RBF, SOM e di Hopfield* sono descritte tre tipologie differenti di reti neurali: *Radial Basis Functions*-RBF, *Self-Organizing Maps*-SOM, e la rete di Hopfield. RBF utilizza un approccio differente nel disegno di una rete neurale basato sullo strato *hidden* (unico nella rete) composto da neuroni in cui vengono definite funzioni a base radiale, da qui il nome di *Radial Basis Functions*-RBF, e che effettua una trasformazione non-lineare dei dati di input forniti alla rete. Questi neuroni costituiscono delle basi per i dati di input (vettori). La ragione per cui si adopera una trasformazione non-lineare nello strato hidden, seguita da una lineare in quello di output, permette ad un problema di classificazione di pattern di operare in uno spazio a dimensione molto più grande (nella trasformazione non lineare dallo strato di input in quello hidden) ed è più probabile di essere linearmente separabile rispetto ad uno spazio a dimensione ridotta. Da questa osservazione deriva la ragione per cui lo strato hidden è generalmente più grande di quello di input (ovvero il numero di neuroni hidden è maggiore della cardinalità del segnale di input). La rete SOM invece ha un modello di apprendimento non supervisionato ed ha l'originalità di raggruppare autonomamente dati di input sulla base della loro similarità senza valutare l'errore di convergenza con informazioni esterne sui dati. Utile quando

sui dati non si ha una conoscenza esatta per classificarli. Si ispira alla topologia del modello della corteccia del cervello considerando la connettività dei neuroni ed in particolare il comportamento di un neurone attivato e l'influenza con i neuroni vicini che rinforzano i legami rispetto a quelli più lontani che man man diventano più deboli. Con la rete di Hopfield il modello di apprendimento è supervisionato e con la capacità di memorizzare informazione e di recuperarla attraverso contenuti anche parziali dell'informazione originale. Presenta una sua originalità basata su fondamenti fisici che ha rivitalizzato l'intero settore delle reti neurali. Alla rete è associata una funzione energia da minimizzare durante la sua evoluzione con una successione di stati, fino a raggiungere uno stato finale corrispondente al minimo della funzione energia. Questa sua caratteristica consente di essere utilizzata per risolvere e impostare un problema di ottimizzazione in termini di funzione obiettivo da associare ad una funzione energia.

Nel cap.3 *Analisi della Tessitura* sono riportati gli algoritmi che caratterizzano la tessitura presente nelle immagini. La tessitura risulta una componente importante per il riconoscimento degli oggetti. Nel campo dell'elaborazione delle immagini si è consolidato con il termine tessitura una qualunque disposizione geometrica e ripetitiva dei livelli di grigio (o di colore) di una immagine. In tale contesto la tessitura diventa una componente strategica aggiuntiva per risolvere il problema del riconoscimento oggetti, la segmentazione delle immagini ed i problemi di sintesi. Alcuni degli algoritmi descritti si basano sui meccanismi della percezione visiva umana della tessitura. Sono utili per lo sviluppo di sistemi per l'analisi automatica del contenuto informativo di un'immagine ottenendo un partizionamento dell'immagine in regioni con differenti tessiture.

Nel cap.4 *Paradigmi per la Visione 3D* sono riportati gli algoritmi che analizzano immagini 2D per ricostruire una scena tipicamente di oggetti 3D. Un sistema di visione 3D ha il problema fondamentale tipico dei problemi inversi, ossia da singole immagini 2D, che sono solo una proiezione bidimensionale del mondo 3D (acquisizione parziale), deve poter ricostruire la struttura 3D della scena osservata ed eventualmente definire una relazione tra gli oggetti. La ricostruzione 3D avviene partendo dalle immagini 2D che contengono solo informazioni parziali del mondo 3D (perdita di informazione dalla proiezione 3D→2D) e possibilmente utilizzando i parametri geometrici e radiometrici di calibrazione del sistema di acquisizione. Sono illustrati i meccanismi della visione umana basati anche sulla predizione e conoscenza a priori del mondo. Nel settore della visione artificiale, la tendenza attuale è quella di sviluppare sistemi 3D orientati per specifici domini ma con caratteristiche che vanno nella direzione di imitare alcune funzioni del sistema visivo umano. Sono descritti metodi di ricostruzione 3D che utilizzano più telecamere osservando la scena da più punti di vista, oppure sequenze d'immagini tempo-varianti acquisite da una singola telecamera. Sono descritte le teorie della visione, dalle leggi Gestalt al paradigma della visione di Marr ed i modelli computazionali della stereovisione.

Nel cap.5 *Shape from Shading*-(SfS) sono riportati gli algoritmi per ricostruire la forma (*shape*) della superficie 3D visibile utilizzando soltanto le in-

formazioni di variazione di luminosità (*shading*, ossia le sfumature di livello di grigio o di colore) presenti nell'immagine. Il problema inverso, di ricostruire la forma della superficie visibile dalle variazioni di luminosità presenti nell'immagine è noto come il problema dello *Shape from Shading*. La ricostruzione della superficie visibile non deve intendersi rigorosamente come ricostruzione 3D della superficie. Infatti, da un singolo punto di osservazione della scena, un sistema di visione monoculare non può stimare una misura di distanza tra osservatore e oggetto visibile, pertanto con gli algoritmi di SfS si ha una ricostruzione non metrica ma qualitativa della superficie 3D. È descritta la teoria dello SfS basata sulla conoscenza della sorgente luminosa (direzione e distribuzione), il modello di riflettanza della scena, il punto di osservazione e la geometria della superficie visibile, che insieme, concorrono nel processo di formazione dell'immagine. Sono derivate le relazioni tra i valori d'intensità luminosa dell'immagine e la geometria della superficie visibile (in termini di orientazione della superficie punto per punto) sotto alcune condizioni di illuminazione e del modello di riflettanza. Sono inoltre descritti altri algoritmi di ricostruzione 3D della superficie basato sul paradigma Shape from xxx, dove con xxx può essere la tessitura, luce strutturata proiettata sulla superficie da ricostruire, oppure immagini 2D della superficie focalizzata o defocalizzata.

Nel cap.6 *Analisi del Movimento* sono riportati gli algoritmi di percezione della dinamica della scena, in analogia a quanto avviene, nei sistemi di visione di diversi esseri viventi. Con gli algoritmi di analisi del movimento è possibile derivare il moto 3D, quasi in tempo reale, dall'analisi di sequenze d'immagini 2D tempo-varianti. Paradigmi sull'analisi del movimento hanno dimostrato che la percezione del movimento deriva dalle informazioni degli oggetti valutando la presenza di occlusioni, tessitura, contorni, ecc. Sono descritti gli algoritmi per la percezione del movimento che avviene nella realtà fisica e non al movimento apparente. Sono analizzate diverse metodologie di analisi del movimento da quelli con carico computazionale limitato come quelli basati sulla differenza delle immagini tempo-varianti a quelli più complessi basati sul flusso ottico considerando contesti applicativi con diversi livelli di entità di movimento e scene-ambiente con diversa complessità. Nel contesto di corpi rigidi, dall'analisi del moto, derivato da una sequenza di immagini tempo-variante, sono descritti gli algoritmi che, oltre al movimento (di traslazione e rotazione), stimano la ricostruzione della struttura 3D della scena e la distanza di tale struttura dall'osservatore. Informazioni utili, nel caso di osservatore mobile (robot o veicolo), per stimare anche il tempo di collisione. Sono infatti descritti i metodi per risolvere il problema della ricostruzione 3D della scena acquisendo una sequenza di immagini con una singola telecamera i cui parametri intrinseci rimangono costanti anche se non noti (telecamera non calibrata) insieme alla non conoscenza del moto. I metodi proposti rientrano nella problematica di risolvere un problema inverso. Sono descritti algoritmi per ricostruire la struttura 3D della scena (ed il moto), ovvero per calcolare le coordinate di punti 3D della scena di cui è nota la loro proiezione 2D in ciascuna immagine della sequenza tempo-variante.

Infine, nel cap.7 *Calibrazione* sono riportati gli algoritmi per calibrare il sistema di acquisizione immagini (normalmente una telecamera e stereovisione) fondamentali per rilevare informazioni metriche (rilevare dimensioni di un oggetto oppure determinare misure accurate di distanza oggetto-osservatore) della scena dall'immagine. Sono descritti i vari metodi di calibrazione della telecamera che determinano i relativi parametri intrinseci (lunghezza focale, la dimensione orizzontale e verticale del singolo fotorecettore del sensore oppure il rapporto d'aspetto, la dimensione della matrice del sensore, i coefficienti del modello di distorsione radiale, le coordinate del punto principale ovvero il centro ottico) ed i parametri estrinseci che definiscono la trasformazione geometrica per passare dal sistema di riferimento del mondo a quello della telecamera. La geometria epipolare introdotta nel cap.19 è descritta in questo capitolo per risolvere il problema della corrispondenza di punti omologhi in un sistema di visione stereo con le due telecamere calibrate e non. Con la geometria epipolare è semplificata la ricerca dei punti omologhi tra le immagini stereo introducendo la matrice Essenziale e la matrice Fondamentale. Sono inoltre descritti gli algoritmi per la stima di tali matrici, noti a priori i punti corrispondenti di una piattaforma di calibrazione. Con la geometria epipolare il problema della ricerca dei punti omologhi si riduce a mappare un punto di una immagine sulla corrispondente linea epipolare nell'altra immagine. È possibile semplificare il problema della corrispondenza attraverso una ricerca monodimensionale punto a punto tra le immagine stereo. Ciò è realizzato con la procedura di allineamento delle immagini, noto come rettifica delle immagini stereo. I diversi algoritmi sono stati descritti alcuni basati sui vincoli della geometria epipolare (telecamere non calibrate dove la matrice fondamentale include i parametri intrinseci) e sulla conoscenza o meno dei parametri intrinseci ed estrinseci di telecamere calibrate. Il capitolo 7 termina con il paragrafo della ricostruzione 3D della scena in relazione della conoscenza a disposizione del sistema stereo di acquisizione. Sono descritte le procedure di triangolazione per la ricostruzione 3D della geometria della scena senza ambiguità, date le proiezioni 2D dei punti omologhi delle immagini stereo, noti i parametri di calibrazione del sistema stereo. Se sono noti solo i parametri intrinseci la geometria 3D della scena è ricostruita stimando i parametri estrinseci del sistema a meno di un fattore di scala non determinabile. Se i parametri di calibrazione del sistema stereo non sono disponibili ma sono note solo le corrispondenze tra le immagini stereo la struttura della scena è recuperata attraverso una trasformazione omografica incognita.