

Αοι



Arcangelo Distante  
Cosimo Distante

**Visione computazionale  
Dall'immagine ai simboli**

Volume II

*Prefazione di*  
Virginio Cantoni





Aracne editrice

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

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

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

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

ISBN 978-88-255-1533-6

*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 Operazioni locali: Edging</b>	<b>25</b>
1.1 Definizioni . . . . .	28
1.2 Filtro del Gradiente . . . . .	29
1.3 Approssimazione del filtro Gradiente . . . . .	32
1.4 Operatore di Roberts . . . . .	37
1.5 Manipolazione dell'immagine gradiente . . . . .	38
1.6 Operatore di Sobel . . . . .	39
1.7 Operatore di Prewitt . . . . .	41
1.8 Operatore di Frei&Chen . . . . .	41
1.9 Operatori LED a confronto . . . . .	41
1.10 Operatore gradiente direzionale . . . . .	47
1.11 Operatore derivata della Gaussiana (DroG) . . . . .	49
1.12 Operatore Laplaciano . . . . .	51
1.13 Laplaciano della Gaussiana <i>LoG</i> . . . . .	55
1.14 Differenza di Gaussiane <i>DoG</i> . . . . .	63
1.15 Operatore della derivata seconda direzionale . . . . .	64
1.16 Operatore di Canny . . . . .	64
1.16.1 Algoritmo di Canny . . . . .	70
1.17 Estrazione di punti . . . . .	73
1.18 Estrazione di linee . . . . .	74
1.19 Filtraggio Passa-Alto . . . . .	75
1.19.1 Filtro Ideale Passa-Alto (IHPF) . . . . .	76
1.19.2 Filtro Passa-Alto di Butterworth (BHPF) . . . . .	78
1.19.3 Filtro Gaussiano Passa-Alto (GHPF) . . . . .	78
1.20 Filtro Ideale Blocco-Banda . . . . .	79
1.20.1 Filtro Butterworth e Gaussiano Blocco-Banda . . . . .	80
1.21 Filtro Ideale Passa-Banda . . . . .	81
1.21.1 Filtro Passa-Banda di Butterworth e Gaussiano . . . . .	82

1.21.2	Filtro Passa-Banda da Differenza di Gaussiane . . . . .	82
1.21.3	Filtro Laplaciano nel dominio delle frequenze . . . . .	83
1.22	Filtri di Sharpening . . . . .	84
1.22.1	Filtri lineari di Sharpening . . . . .	84
1.22.2	Unsharp Masking . . . . .	86
1.22.3	Filtro di Sharpening High-Boost . . . . .	88
1.22.4	Filtraggio di Sharpening nel dominio delle frequenze . .	90
1.22.5	Filtro Omomorfo . . . . .	91

## **2 Trasformate Lineari Fondamentali** **95**

2.1	Introduzione . . . . .	95
2.2	Trasformazione lineare discreta monodimensionale . . . . .	96
2.2.1	Trasformate Unitarie . . . . .	96
2.2.2	Trasformate Ortogonali . . . . .	96
2.2.3	Trasformate Ortonormali . . . . .	97
2.2.4	Esempio di trasformata unitaria monodimensionale . . .	98
2.3	Trasformazione lineare discreta bidimensionale . . . . .	99
2.3.1	Esempio di trasformata unitaria bidimensionale . . . . .	100
2.4	Considerazioni sulle trasformazioni unitarie . . . . .	100
2.4.0.1	Proprietà delle trasformate unitarie . . . . .	102
2.5	Trasformate sinusoidali . . . . .	103
2.6	La trasformata coseno discreta (DCT) . . . . .	104
2.7	La trasformata seno discreta (DST) . . . . .	108
2.8	La trasformata di Hartley discreta . . . . .	108
2.9	Trasformate con funzioni rettangolari . . . . .	110
2.9.1	Trasformata discreta di Hadamard - DHaT . . . . .	110
2.9.2	Trasformata discreta di Walsh (DWHT) . . . . .	115
2.9.3	Trasformata di Slant . . . . .	117
2.9.4	Trasformata di Haar . . . . .	119
2.10	Trasformate basate sugli autovettori . . . . .	122
2.10.1	Analisi delle componenti principali (PCA) . . . . .	125
2.10.1.1	Riepilogo caratteristiche della PCA . . . . .	128
2.10.1.2	Passi essenziali per il calcolo delle componenti PCA . . . . .	129
2.10.2	PCA/KLT per la compressione dati . . . . .	131
2.10.3	Calcolo degli assi principali di un oggetto bidimensionale	137
2.10.4	Riduzione della dimensionalità . . . . .	138
2.10.5	Calcolo delle componenti significative nelle immagini mul- tispettrali . . . . .	139
2.10.6	Eigenface - Riconoscimento dei volti . . . . .	143
2.11	La trasformata SVD - Singular Value Decomposition . . . . .	148
2.12	Trasformata Wavelets . . . . .	152
2.12.1	Trasformata Wavelet Continua - CWT . . . . .	154
2.12.2	Trasformata Wavelet 2D Continua - CWT . . . . .	156
2.12.3	Trasformata Wavelet come filtraggio passa-banda . . . . .	156



2.12.4	Trasformata Wavelet Discreta - DWT . . . . .	158
2.12.5	Trasformata Wavelet Discreta Veloce - FWT . . . . .	160
2.12.6	Trasformata Wavelet Discreta 2D - DWT2 . . . . .	163
2.12.7	Trasformata Wavelet Biortogonale . . . . .	168
2.12.8	Applicazioni della Trasformata Wavelet Discreta . . . . .	170
2.12.8.1	Compressione immagini . . . . .	170
2.12.8.2	Attenuazione del rumore . . . . .	171
2.12.8.3	Estrazione dei bordi . . . . .	173
2.12.8.4	Applicazioni varie della DWT . . . . .	173
2.13	Riepilogo . . . . .	175
<b>3</b>	<b>Trasformazioni geometriche</b> . . . . .	<b>177</b>
3.1	Geometria della visione . . . . .	178
3.1.1	Punti 2D . . . . .	178
3.2	Operatore geometrico . . . . .	178
3.2.1	Traslazione . . . . .	181
3.2.2	Ingrandimento o riduzione . . . . .	182
3.2.3	Rotazione . . . . .	182
3.2.4	Skew o Shear . . . . .	183
3.2.5	Speculare . . . . .	183
3.2.6	Trasposta . . . . .	184
3.2.7	Sistemi di Coordinate e Coordinate omogenee . . . . .	184
3.2.8	Trasformazioni geometriche omogenee elementari . . . . .	185
3.2.8.1	Traslazione in coordinate omogenee . . . . .	185
3.2.8.2	Rotazione in coordinate omogenee . . . . .	186
3.2.8.3	Traslazione e Rotazione combinate in coordina- te omogenee . . . . .	186
3.3	Trasformazioni geometriche affini . . . . .	188
3.3.0.1	Trasformazione affine di similarità . . . . .	188
3.3.0.2	Trasformazione affine generalizzata . . . . .	189
3.3.1	Trasformazioni affini elementari . . . . .	190
3.3.1.1	Trasformazione affine di cambiamento di scala . . . . .	191
3.3.1.2	Trasformazione affine di Shear . . . . .	191
3.3.1.3	Trasformazione affine composta . . . . .	191
3.4	Separabilità delle trasformazioni . . . . .	192
3.5	Trasformazione Omografica . . . . .	194
3.5.1	Applicazioni della trasformazione omografica . . . . .	200
3.6	Trasformazione Prospettica . . . . .	201
3.7	Trasformazioni geometriche per la registrazione di immagini . . . . .	205
3.8	Trasformazioni geometriche non lineari . . . . .	206
3.9	Trasformazione geometrica e ricampionamento . . . . .	210
3.9.1	Interpolazione ideale . . . . .	212
3.9.2	Interpolazione di ordine zero (Nearest-Neighbor) . . . . .	218
3.9.3	Interpolazione lineare del primo ordine . . . . .	219
3.9.4	Interpolazione Biquadratica . . . . .	221

3.9.5	Interpolazione Bicubica . . . . .	223
3.9.6	Interpolazione B-Spline . . . . .	227
3.9.6.1	B-spline quadratica . . . . .	229
3.9.6.2	B-spline cubica . . . . .	230
3.9.7	Interpolazione per approssimazione ai minimi quadrati .	231
3.9.8	Interpolazione non polinomiale . . . . .	231
3.9.9	Metodi di Interpolazione a confronto . . . . .	232
3.9.9.1	Esempi di Riduzioni e Ingrandimenti . . . . .	235
3.9.9.2	Esempi di Rotazioni Multiple . . . . .	237
3.9.9.3	Esempi di Ingrandimenti con Rotazione . . . . .	237
<b>4</b>	<b>Ricostruzione dell'immagine degradata: Restoration</b>	<b>239</b>
4.1	Modello del rumore . . . . .	240
4.1.1	Rumore Additivo Gaussiano . . . . .	242
4.1.1.1	Processo stocastico stazionario normale . . . . .	242
4.1.1.2	Rumore bianco . . . . .	243
4.1.1.3	Processo stocastico ergotico . . . . .	244
4.1.2	Altri modelli statistici del rumore . . . . .	245
4.1.2.1	Rumore di Rayleigh . . . . .	245
4.1.2.2	Rumore di Erlang (Gamma) . . . . .	246
4.1.2.3	Rumore Esponenziale . . . . .	247
4.1.2.4	Rumore Uniforme . . . . .	247
4.1.3	Rumore Impulsivo Bipolare . . . . .	248
4.1.4	Rumore Periodico . . . . .	248
4.1.5	Stima dei Parametri del Rumore . . . . .	249
4.2	Filtraggio spaziale per la rimozione del rumore . . . . .	249
4.2.1	Filtro di Media Geometrica . . . . .	250
4.2.2	Filtro di Media Armonica . . . . .	251
4.2.3	Filtro di Media Contro-Armonica . . . . .	252
4.2.4	Filtri sulle statistiche d'ordine . . . . .	252
4.2.4.1	Filtro Minimo e Massimo . . . . .	252
4.2.4.2	Filtro del Punto Medio . . . . .	252
4.2.4.3	Filtro di Media Alpha-Trimmed . . . . .	253
4.2.5	Applicazione dei filtri spaziali di <i>Media</i> e delle <i>Statistiche d'ordine</i> per il restauro delle immagini . . . . .	254
4.3	Filtri Adattativi . . . . .	255
4.3.1	Filtro Mediano Adattativo . . . . .	257
4.4	Riduzione del Rumore Periodico con Filtraggio nel Dominio delle Frequenze . . . . .	260
4.4.1	Filtri di Notch . . . . .	261
4.4.2	Filtraggio ottimale con filtro Notch . . . . .	263
4.5	Determinazione della funzione di degradazione . . . . .	265
4.5.0.1	Derivazione di $H_D$ mediante osservazione dell'immagine degradata . . . . .	266
4.5.0.2	Determinazione di $H_D$ per via sperimentale . . . . .	266

4.5.0.3	Determinazione di $H_D$ basata su modello fisico-matematico: blurring da movimento . . . . .	266
4.5.0.4	Determinazione di $H_D$ basata su modello fisico-matematico: blurring da turbolenza atmosferica . . . . .	268
4.6	Filtro inverso - Deconvoluzione . . . . .	269
4.6.0.1	Applicazione del filtro inverso: esempio 1 . . . . .	272
4.6.0.2	Applicazione del filtro inverso: esempio 2 . . . . .	273
4.6.0.3	Applicazione del filtro inverso: esempio 3 . . . . .	274
4.7	Filtro ottimale . . . . .	274
4.7.1	Filtro di Wiener . . . . .	275
4.7.1.1	Filtro di Wiener con immagine originale $f_I$ e rumore $\eta$ non correlati . . . . .	279
4.7.2	Analisi del filtro di Wiener . . . . .	283
4.7.3	Applicazione del filtro di Wiener: caso monodimensionale . . . . .	285
4.7.4	Applicazione del filtro di Wiener: caso bidimensionale . . . . .	287
4.8	Filtro di equalizzazione dello spettro di potenza-PSE . . . . .	289
4.9	Filtro condizionato ai minimi quadrati - CMQ . . . . .	290
4.10	Filtri della media geometrica - MG . . . . .	291
4.11	Filtro di deconvoluzione iterativo non lineare . . . . .	293
4.12	Deconvoluzione cieca . . . . .	295
4.13	Filtro di Diffusione non lineare . . . . .	295
4.14	Filtro Bilaterale . . . . .	298
4.15	Dehazing . . . . .	299
<b>5</b>	<b>Segmentazione delle immagini</b> . . . . .	<b>301</b>
5.1	Introduzione . . . . .	301
5.2	Regioni e contorni . . . . .	302
5.3	Il processo di segmentazione . . . . .	302
5.3.1	Segmentazione mediante soglia globale . . . . .	303
5.4	Metodi di segmentazione mediante soglia . . . . .	305
5.4.1	Metodo basato sul rapporto dimensione oggetti/sfondo . . . . .	306
5.4.2	Metodo sull'analisi dell'istogramma . . . . .	306
5.4.3	Metodo basato sul gradiente e Laplaciano . . . . .	307
5.4.4	Metodo della soglia iterativa . . . . .	309
5.4.5	Metodo di massima varianza inter-classe - Otsu . . . . .	310
5.4.6	Metodo di segmentazione con soglia adattiva . . . . .	314
5.4.6.1	Metodo basato sulla scomposizione in sotto-immagini . . . . .	315
5.4.6.2	Metodo basato sulla normalizzazione dello sfondo . . . . .	316
5.4.6.3	Metodo basato sulle proprietà statistiche locali . . . . .	317
5.4.7	Metodo con soglia multi-livello per immagini a colori e multi-spettrali . . . . .	318
5.4.8	Segmentazione basata sull'estrazione dei contorni . . . . .	321
5.4.9	Inseguimento dei bordi (edge following) . . . . .	321
5.4.10	Collegamenti contorni interrotti . . . . .	323
5.4.11	Algoritmo di labeling di Componenti Connesse . . . . .	326

5.4.12	Algoritmo di Filling per Regioni Complesse . . . . .	329
5.5	Estrazione dei contorni mediante la trasformata di Hough . . . . .	330
5.5.1	Trasformata di Hough per l'estrazione delle linee . . . . .	330
5.5.2	Trasformata di Hough per l'estrazione dei cerchi . . . . .	335
5.5.3	Trasformata generalizzata di Hough . . . . .	337
5.5.3.1	Considerazioni finali sulla trasformata di Hough . . . . .	339
5.5.4	Trasformata Randomizzata di Hough - RHT . . . . .	340
5.6	Segmentazione basata sulle regioni . . . . .	341
5.6.1	Segmentazione mediante la crescita delle regioni - growing . . . . .	342
5.6.2	Segmentazione mediante suddivisione delle regioni-splitting . . . . .	344
5.6.3	Segmentazione combinando splitting e merging . . . . .	346
5.7	Segmentazione mediante la trasformata watershed . . . . .	348
5.7.0.1	Algoritmo Watershed basato sulla simulazione dell'inondazione . . . . .	349
5.7.0.2	Algoritmo Watershed utilizzando dei <i>marker</i> . . . . .	351
5.8	Segmentazione mediante algoritmi di clustering . . . . .	352
5.8.1	Segmentazione con algoritmo K-means . . . . .	355
5.8.2	Segmentazione con algoritmo mean-shift . . . . .	355
5.8.2.1	Applicazione del mean-shift per la segmentazione . . . . .	362
5.8.2.2	Applicazione del mean-shift per il filtraggio . . . . .	364
5.8.2.3	Applicazione del mean-shift per il tracking . . . . .	364
5.8.2.4	Conclusioni . . . . .	365
<b>6</b>	<b>Rilevatori e Descrittori di Punti di Interesse</b> . . . . .	<b>367</b>
6.1	Introduzione . . . . .	367
6.2	Rilevatore di Punti di Interesse - Operatore di Moravec . . . . .	369
6.2.1	Limiti dell'Operatore di Moravec . . . . .	373
6.2.1.1	Operatore con risposta anisotropa . . . . .	373
6.2.1.2	Operatore con risposta rumorosa . . . . .	373
6.2.1.3	Operatore molto sensibile ai falsi bordi . . . . .	374
6.3	Rilevatore di Punti di Interesse - Algoritmo di Harris-Stephens . . . . .	374
6.3.1	Limiti e proprietà dell'algoritmo di Harris . . . . .	380
6.3.1.1	Invarianza alla rotazione . . . . .	381
6.3.1.2	Invarianza parziale fotometrica . . . . .	381
6.3.1.3	Cambiamento di scala dell'immagine . . . . .	382
6.4	Varianti dell'algoritmo di Harris-Stephens . . . . .	382
6.5	Rilevatore di Punti di Interesse - Hessian . . . . .	383
6.6	Punti di interesse scala invariante . . . . .	384
6.6.0.1	Rappresentazione dello Scale-Space . . . . .	387
6.7	Rilevatori e Descrittori di Punti di Interesse Scala Invariante . . . . .	392
6.7.1	Rilevatore e Descrittore - SIFT . . . . .	392
6.7.2	Componente Rilevatore - SIFT . . . . .	392
6.7.2.1	Determinazione degli estremi locali nello <i>scale-space</i> . . . . .	393
6.7.2.2	Localizzazione dei punti di interesse . . . . .	394

6.7.2.3	Stima dell'accuratezza della locazione dei punti estremi . . . . .	395
6.7.2.4	Filtraggio dei punti estremi a basso contrasto . . . . .	396
6.7.2.5	Filtraggio dei punti estremi lungo i bordi . . . . .	396
6.7.2.6	Risultati del rilevatore di punti di interesse con l'algoritmo SIFT . . . . .	397
6.7.3	Componente Descrittore - SIFT . . . . .	398
6.7.3.1	Determinazione dell'orientazione dominante . . . . .	398
6.7.3.2	Generazione del Descrittore SIFT . . . . .	399
6.7.3.3	Invarianza al contrasto . . . . .	401
6.7.3.4	Motivazione biologica del descrittore SIFT . . . . .	401
6.7.3.5	Adeguatezza dimensionale e complessità computazionale dei descrittore SIFT . . . . .	402
6.7.4	Descrittore GLOH . . . . .	404
6.8	Rilevatore e Descrittore SURF-Speed-Up Robust Features . . . . .	405
6.8.1	Componente Rilevatore di SURF, Fast Hessian-FH . . . . .	405
6.8.1.1	Immagine Integrale . . . . .	405
6.8.1.2	Rilevatore Fast Hessian - FH . . . . .	406
6.8.1.3	Localizzazione dei punti di interesse . . . . .	411
6.8.1.4	Risultati del rilevatore Fast-Hessian . . . . .	411
6.8.2	Componente Descrittore di SURF . . . . .	412
6.8.2.1	Calcolo della Direzione Dominante . . . . .	412
6.8.2.2	Calcolo del Vettore Descrittore . . . . .	413
6.8.2.3	Indirizzamento veloce per il problema della corrispondenza . . . . .	416
6.8.2.4	Risultati del rilevatore-descrittore SURF . . . . .	416
6.8.3	Rilevatore di Harris-Laplace . . . . .	418
6.8.4	Rilevatore di Hessian-Laplace . . . . .	420
6.9	Punti di Interesse Invariante Affine . . . . .	423
6.9.1	Rilevatore di Harris-Affine . . . . .	424
6.9.1.1	Adattamento Affine . . . . .	424
6.9.1.2	Algoritmo Iterativo di Adattamento Affine . . . . .	427
6.9.1.3	Prestazioni del rilevatore Harris-Affine . . . . .	430
6.9.2	Rilevatore Hessian-Affine . . . . .	431
6.10	Rilevatori Veloci di Corner . . . . .	432
6.10.1	Rilevatore SUSAN - Smallest Univalued Segment Assimilating Nucleus . . . . .	432
6.10.2	Rilevatore Segment Test di Trajkovic-Hedley . . . . .	435
6.10.3	Rilevatore FAST - Features from Accelerated Segment Test . . . . .	436
6.10.3.1	Miglioramento dell'algoritmo mediante paradigmi di apprendimento automatico (Machine Learning) . . . . .	438
6.10.3.2	Soppressione dei non-massimi . . . . .	439
6.10.3.3	Risultati del rilevatore FAST . . . . .	440
6.11	Rilevatori e Descrittori di Regioni . . . . .	441

6.11.1	Rilevatore MSER-Maximally Stable Extremal Regions .	442
6.11.2	Rilevatore IBR-Intensity extrema-Based Regions . . . .	446
6.11.3	Rilevatore di Regioni Salienti Affine . . . . .	447
6.11.4	Rilevatore EBR-Edge-Based Region . . . . .	449
6.11.5	Rilevatore PCBR-Principal Curvature-Based Region . .	451
6.11.6	Rilevatore SISF-Scale Invariant Shape Features . . . . .	455
6.12	Riassunto e Conclusioni . . . . .	458
<b>Bibliografia</b>		<b>465</b>
<b>Ringraziamenti</b>		<b>475</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