



POLITECNICO DI MILANO
Dipartimento di Elettronica e Informazione
DOTTORATO DI RICERCA IN INGEGNERIA INFORMATICA E AUTOMATICA

Una Metodologia Gerarchica per Analisi e Sintesi di Immagini Tridimensionali

Tesi di dottorato di:
Stefano Ferrari

Relatore:

Prof. Vincenzo Piuri

Correlatore:

Ing. Alberto Borghese

Tutore:

Prof. Mariagiovanna Sami

Coordinatore del programma di dottorato:

Prof. Carlo Ghezzi

XIII ciclo

POLITECNICO DI MILANO
Dipartimento di Elettronica e Informazione
Piazza Leonardo da Vinci 32 I 20133 — Milano



POLITECNICO DI MILANO
Dipartimento di Elettronica e Informazione
DOTTORATO DI RICERCA IN INGEGNERIA INFORMATICA E AUTOMATICA

Una Metodologia Gerarchica per Analisi e Sintesi di Immagini Tridimensionali

Ph.D. Dissertation of:
Stefano Ferrari

Advisor:

Prof. Vincenzo Piuri

Tutor:

Prof. Mariagiovanna Sami

Supervisor of the Ph.D. Program:

Prof. Carlo Ghezzi

XIII edition

Indice

1. Introduzione	1
1.1. La digitalizzazione di oggetti reali	1
1.2. La costruzione di un modello tridimensionale	4
1.3. Obiettivo della tesi	5
1.4. Organizzazione della tesi	8
2. Metodi e tecniche per la digitalizzazione	11
2.1. La digitalizzazione vista ad alto livello	12
2.1.1. Acquisizione	12
2.1.2. Ricostruzione	13
2.1.3. Ottimizzazione e trasformazione	16
2.2. Rappresentazione di oggetti reali	20
2.2.1. Rappresentazione della geometria	20
2.2.2. Parametrizzazione di una superficie	28
2.2.3. Rappresentazione delle caratteristiche cromatiche	28
2.3. Acquisizione	33
2.3.1. Dispositivi di acquisizione	34
2.3.2. Tecniche di misura	39
2.4. Ricostruzione	40
2.4.1. Ricostruzione da punti	42
2.4.2. Multirisoluzione	54
2.4.3. Parametrizzazione	57
2.4.4. Integrazione	59
2.4.5. Applicazione del colore	64
2.5. Elaborazioni post-ricostruzione	65
2.5.1. Surface simplification	67
2.5.2. Modeling/Editing	74
2.5.3. Animation	74
2.5.4. Morphing	75
2.6. Conclusioni	75

3. Una metodologia gerarchica per la digitalizzazione	77
3.1. La digitalizzazione	78
3.1.1. Acquisizione	78
3.1.2. Ricostruzione	78
3.1.3. Ottimizzazione	80
3.2. La ricostruzione	80
3.2.1. Formalizzazione della metodologia	83
3.2.2. Utilizzo dei paradigmi gerarchici	90
3.2.3. Alcuni problemi nell'acquisizione dei dati	91
3.3. Contributo della tesi	92
3.4. Conclusioni	95
4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D	101
4.1. Le Wavelet	102
4.1.1. La Trasformata Wavelet Continua	102
4.1.2. La Trasformata Wavelet Discreta	104
4.1.3. Applicazione alle immagini digitali	114
4.1.4. Estensioni	117
4.2. Reti HRBF	121
4.2.1. Introduzione	121
4.2.2. Rete RBF gaussiana regolare e filtro gaussiano	123
4.2.3. HRBF: la multirisoluzione	127
4.2.4. Estensione al caso bidimensionale	132
4.2.5. Approssimazioni Multiscala	133
4.2.6. Riformulazione HRBF in chiave digitale	135
4.3. Confronto tra wavelet e HRBF	137
4.3.1. Confronto teorico	138
4.3.2. Confronto sperimentale	142
4.4. Conclusioni	152
5. Un esempio applicativo: Autoscan	155
5.1. Fase di acquisizione	156
5.1.1. Acquisizione	156
5.1.2. Accuratezza della acquisizione	158
5.2. Fase di ricostruzione	160
5.2.1. Ricostruzione della superficie	160
5.2.2. Conversione in superficie poligonale	165
5.3. Conclusioni	165
6. Conclusioni	169
6.1. Risultati ottenuti	171
6.2. Spunti per ricerche future	173

Glossario

177

Indice

Elenco delle figure

1.1. Schema ad alto livello dei sistemi di acquisizione 3D	6
2.1. Registrazione di due curve.	18
2.2. Rappresentazione a diversi livelli di dettaglio dello stesso oggetto.	18
2.3. Modellazione multirisoluzione [GSS99].	19
2.4. Morphing [LDSS99].	19
2.5. Mesh poligonale [Hop94].	23
2.6. Gouraud shading.	23
2.7. <i>View-dependent rendering</i> [Hop97].	23
2.8. Modellazione con una curva NURBS [PT97]. La spezzata formata dai punti di controllo P_0, \dots, P_5 fornisce la for- ma generale della curva. Aumentando un peso la curva diventa più aderente al punto di controllo corrispondente (P_3 nell'esempio).	25
2.9. Procedura iterativa per ottenere una suddivisione cubica [SS96].	25
2.10. Subdivision surface [ZSS96].	27
2.11. Un reticolo di <i>patch</i> B-spline [EH96].	27
2.12. <i>Per vertex color</i> (a sinistra) e <i>texture mapping</i> (a destra) [TH].	32
2.13. Texture procedurale	37
2.14. Il fenomeno dello <i>scattering</i> provoca un errore sistematico nella misura [LPC ⁺ 00].	37
2.15. Un modello di scanner Digibotics (sinistra) e la tecnica utilizzata per la misura (destra) [DIGb].	41
2.16. Lo scanner Faro [FAR].	41
2.17. Ricostruzione di una superficie effettuata con <i>marching cubes</i> a differenti risoluzioni [Sha]. La superficie originale è quella a sinistra.	47
2.18. α -shape di un insieme di punti al variare del parametro α	47
2.19. Comportamento di <i>Ball pivoting</i> su dati con differenti caratteristiche [BMR ⁺ 99].	47

Elenco delle figure

2.20. <i>Carving</i> [RW97]. Lo spazio occupato dai dati dell'acquisizione viene quantizzato e i voxel attraversati troppo spesso dalla linea di visione vengono eliminati (sinistra). La superficie viene ricostruita a partire dalle facce dei voxel rimasti visibili dall'infinito dalle sei direzioni parallele agli assi.	50
2.21. Superquadriche deformabili [TM91].	53
2.22. Una mappa di Kohonen approssima una distribuzione complessa [Koh95].	53
2.23. MAPS: il modello a bassa risoluzione è utilizzato come base per la parametrizzazione che poi viene propagata ai vertici del modello originale [LSS ⁺ 98].	61
2.24. La formazione di <i>clique</i> è una situazione che la registrazione deve evitare [Pul99].	61
2.25. Descrizione multirisoluzione di una mesh [Hop96]	68
2.26. Semplificazione tramite <i>quadric error metrics</i> [GH97].	73
3.1. Schema ad alto livello della metodologia di digitalizzazione	79
3.2. La ricostruzione può essere effettuata utilizzando i dati forniti dall'acquisizione solo per ricostruire le superfici corrispondenti alle viste.	82
3.3. Le elaborazioni vengono effettuate sulle ricostruzioni parziali e poi vengono riportate ai dati.	82
3.4. Acquisizione multi- <i>vista</i>	85
3.5. I dati raccolti dalla sessione raffigurata in figura 3.4 vengono generalizzati (il vero profilo dell'oggetto è riportato tratteggiato).	88
3.6. Estrazione delle <i>caratteristiche</i>	88
3.7. La registrazione viene effettuata allineando le <i>caratteristiche</i>	96
3.8. Le ricostruzioni dello stesso cluster, allineate (a), vengono integrate tra loro (b).	97
3.9. Le trasformazioni di integrazione vengono propagate ai dati.	98
3.10. La parte visibile dell'oggetto può nascondere altre regioni.	98
3.11. La ricostruzione in regioni dove i dati siano mancanti è affidata alle proprietà di generalizzazione del paradigma scelto per la ricostruzione.	99
3.12. Dati non contigui devono essere separati(segmentazione).	99
4.1. Annidamento degli spazi approssimazione (V_j) e dettaglio (W_j).	106

4.2.	Calcolo della trasformata e della antitrasformata con le differenti famiglie di MRA	109
4.3.	Schema per una implementazione digitale della FWT . . .	113
4.4.	Schema per una implementazione digitale della IFWT . .	113
4.5.	Funzioni base e filtri della MRA biortogonale 3.3	115
4.6.	Funzioni base e filtri della MRA biortogonale 3.7	116
4.7.	Filtro gaussiano	125
4.8.	Schema di ricostruzione della HRBF	131
4.9.	Schema per una implementazione digitale della procedura di analisi HRBF	136
4.10.	Schema per una implementazione digitale della procedura di sintesi HRBF	137
4.11.	Struttura <i>fine to coarse</i> della MRA e <i>coarse to fine</i> della HRBF.	140
4.12.	Rappresentazione della ricostruzione HRBF mediante analogia con vettori del piano \mathbb{R}^2 : spazi non paralleli (G_1 e G_2) consentono la ricostruzione del residuo (r_1).	140
4.13.	Dati 3D.	142
4.14.	Triangolazione di Delaunay (a) e interpolazione mediante spline cubica (b) dei dati in figura 4.13.	143
4.15.	Ricostruzione HRBF dei dati in figura 4.13	144
4.16.	Uscite dei singoli livelli HRBF	144
4.17.	Posizione delle gaussiane nei livelli HRBF	145
4.18.	Ricampionamento su griglia regolare	146
4.19.	Ricostruzione con MRA biortogonale 3.3	147
4.20.	Effetti delle notazioni in virgola fissa e mobile sulla rappresentazione di un segmento di linea retta	149
4.21.	Set-up per la prova di robustezza rispetto alla notazione usata	150
4.22.	La funzione utilizzata per le prove	151
4.23.	Errore di generalizzazione nel caso di perturbazioni dovute alla notazione in virgola fissa	151
4.24.	Errore di generalizzazione nel caso di perturbazioni dovute alla notazione in virgola mobile	152
5.1.	Schema ad alto livello dello scanner tridimensionale Autoscan.	157
5.2.	Sessione di acquisizione con Autoscan.	159
5.3.	Distribuzione dell'errore di misura.	159
5.4.	Dati geometrici (a) e cromatici (b) acquisiti con Autoscan. 162	
5.5.	Ricostruzione con HRBF a quattro livelli.	163
5.6.	Posizione delle unità utilizzate nei quattro livelli.	164

Elenco delle figure

5.7. Ricampionamento della superficie HRBF e assegnamento del colore tramite texture mapping.	166
--	-----

Elenco delle tabelle

4.1. Cifre di merito della ricostruzione mediante HRBF	145
4.2. Cifre di merito relative alla ricostruzione con MRA bior- tgonale 3.3	147
5.1. Risultati quantitativi della ricostruzione	162

Elenco delle tabelle

1. Introduzione

1.1. La digitalizzazione di oggetti reali

In *computer graphics*, un modello tridimensionale di un oggetto è una rappresentazione numerica delle caratteristiche “visive” (forma, colore) dell’oggetto stesso. A partire dal modello tridimensionale, è possibile riprodurre l’oggetto, animarlo e manipolarlo. La crescente disponibilità di dispositivi per la gestione hardware della grafica tridimensionale ed in generale l’abbassarsi del costo della potenza di calcolo ha recentemente reso comuni applicazioni basate sul trattamento di modelli tridimensionali. Ambiti in cui vengono utilizzate applicazioni di questo genere sono, solo per citarne alcuni, la realtà virtuale in genere, la medicina, la progettazione assistita dal calcolatore, l’intrattenimento e gli effetti speciali cinematografici.

Alla maggiore disponibilità di potenza di calcolo si accompagna la necessità di utilizzare modelli sempre più dettagliati per ottenere effetti di maggiore realismo. La digitalizzazione di oggetti reali è un procedimento che permette di ottenere in maniera semi-automatica i modelli tridimensionali. Essa è basata sulla misura delle caratteristiche geometriche di un oggetto e di quelle che contribuiscono alla formazione del suo colore. Rispetto alla modellazione manuale (eseguita da un utente con software di modellazione dedicati), la digitalizzazione è caratterizzata da una maggiore velocità del procedimento e da una accuratezza più elevata (o quantomeno quantificabile). Inoltre, almeno in via teorica, non richiede all’operatore alcuna capacità artistica.

Le applicazioni coprono uno spettro molto ampio. Ad esempio:

- Nel campo della moda, Cyberware ha sviluppato un sistema di misura, *DigiSize* [DIGa], che traduce le dimensioni corporee in taglie di vestiario. È usato dalla marina militare USA e, rispetto alle procedure tradizionali, ha diversi vantaggi. Innanzitutto rende più rapide le procedure per l’assegnamento dei capi di vestiario ed eventuali modifiche dei vestiti. L’archiviazione delle misure consente inoltre di pianificare meglio le scorte di magazzino. In

1. Introduzione

futuro, una volta effettuata la misura del proprio fisico, vi sarà la possibilità di provare i vestiti virtualmente.

- Il fax *3D* è una apparecchiatura in grado di realizzare a distanza una copia “fisica” di un oggetto tridimensionale. La stazione trasmittente è costituita da un dispositivo in grado di collezionare informazioni sulla struttura dell’oggetto, informazioni che il dispositivo ricevente utilizza per guidare la costruzione della copia, ad esempio tramite stereolitografia [CL96] [3DS].
- La realtà virtuale può essere utilizzata per addestrare, in situazione controllabile, all’esecuzione di compiti critici. Per esempio, in medicina e chirurgia è necessario acquisire una certa esperienza prima di poter operare correttamente. L’inesperienza, cioè il trovarsi a gestire una situazione mai conosciuta prima, può in questo campo comportare un danno grave al paziente. In [SSS00] viene presentato una applicazione che permette di simulare operazioni ortopediche. Il sistema è in grado di simulare svariate situazioni di frattura del femore e gli strumenti e gli impianti per la ricomposizione delle fratture.
- La progettazione è estremamente facilitata dalla possibilità di utilizzare un calcolatore. Alcuni progettisti, soprattutto nel campo architettonico e del design, preferisco modellare fisicamente i prototipi e solo in un secondo momento modellarli al calcolatore (per esempio, per simulazioni o presentazioni). La digitalizzazione dei modelli consente quindi una maggiore libertà espressiva. Inoltre, talvolta bisogna inglobare nella progettazione elementi già esistenti dei quali non si dispone di una documentazione elettronica (o esistente in un formato ormai abbandonato). Le procedure di *reverse engineering* possono essere velocizzate (ad esempio, [TOdSg⁺99]).
- La quantificazione di caratteristiche ha anche lo scopo di confrontare diversi oggetti appartenenti alla stessa classe. Questo concetto può essere applicato in diversi ambiti. Per esempio, in ambito industriale viene utilizzato per il controllo della qualità, mentre nell’ambito della sicurezza viene usato per l’identificazione di persone tramite misure antropometriche [jM96] [PMWP00].
- L’archeologia e l’arte sono caratterizzate dalle contrastanti necessità di proteggere e di rendere fruibili rari reperti o irripetibili opere d’arte [PST00] [LPC⁺00]. L’acquisizione di questi “oggetti” (in realtà possono essere interi complessi) consente da una parte una misura ed una catalogazione più accurata, e dall’altra permette l’accesso alle copie digitali per scopo di studio o per visite cultu-

rali ad un numero elevato di persone, senza rischio per i reperti originali.

- L'architettura virtuale (cioè la progettazione in ambiente virtuale di opere architettoniche) è una disciplina che si sta sviluppando. Essa non solo offre vantaggi dal punto di vista mediatico, ma consente una maggiore libertà di sperimentazione al progettista [Dob97] [FW00] [Sen97]. La possibilità di navigare in ambienti virtuali, giudicare accostamenti, valutare i materiali, può essere vantaggiosa sia in fase di progettazione che per la valutazione da parte di chi commissiona il progetto e che generalmente non ha le conoscenze tecniche per “leggere” la documentazione di progetto. Ad essa si possono affiancare opportuni strumenti automatici di valutazione mediante simulazione (per esempio, stabilità strutturale). Perché una applicazione del genere sia ragionevolmente proponibile, è necessario che il progettista possa disporre di un'ampia gamma di librerie di oggetti e materiali virtuali.
- A causa della larghezza di banda necessaria, la video conferenza ha ancora costi elevati. Disponendo di modelli digitali dei partecipanti, la quantità di banda necessaria può essere drasticamente ridotta. Una volta trasmessi i modelli, infatti, l'informazione trasmessa può essere limitata a quella relativa al movimento ed al parlato. Questo principio è stato considerato nella stesura dello standard MPEG4 [MPE], il quale contiene l'indicazione sui punti di reperi del volto umano.

Le differenti applicazioni hanno caratteristiche computazionali e ambientali differenti. Per esempio, l'archeologia virtuale richiede una grande accuratezza nella ricostruzione, una bassa invasività delle tecniche di acquisizione, ma generalmente non impone limitazioni di tempo per l'elaborazione. La videoconferenza, invece, necessita di una elaborazione in tempo reale, mentre la qualità della ricostruzione richiesta è relativamente bassa. Inoltre, poiché l'informazione da trasferire è generalmente legata alla mimica facciale, le tecniche di acquisizione e ricostruzione possono avvantaggiarsi di modelli del volto e della dinamica facciale. Un discorso simile vale per il controllo qualità industriale: esso deve essere rapido e poco costoso, in quanto deve essere applicato ad un gran numero di pezzi. La conoscenza a priori dell'oggetto da misurare consente di irrobustire e velocizzare le procedure utilizzando un modello. Il fax 3D presuppone che ad utilizzarlo sia un operatore senza particolari conoscenze tecnologiche. La facilità di utilizzo e il basso costo del dispositivo (nonché la sua compattezza) sono quindi le caratteristiche principalmente desiderate.

1.2. La costruzione di un modello tridimensionale

I sistemi utilizzati per la digitalizzazione di oggetti tridimensionali devono quindi rispondere a diverse (e spesso contrastanti) esigenze. Le tecniche scelte per la loro implementazione hanno quindi caratteristiche differenti sia per prestazioni che per principi basanti. Tuttavia, la digitalizzazione è descrivibile ad alto livello indipendentemente dai sistemi utilizzati per ottenerla.

La digitalizzazione di un oggetto è definibile come un processo gerarchico e modulare (cf. fig. 1.1). Il primo livello è deputato al campionamento delle proprietà geometriche e visive della superficie. I sensori, attraverso una prima elaborazione, raccolgono quindi un insieme finito e discreto di caratteristiche geometriche e visive della superficie quali punti, spigoli, colore e tessitura. In un secondo momento, le procedure di ricostruzione generalizzano le informazioni raccolte in modo da generare una rappresentazione della superficie che copra anche le regioni non campionate. Analogamente, le informazioni legate al colore sono elaborate per eliminare l'effetto delle condizioni di illuminazione al momento della misura, in modo da ottenere il colore reale della superficie (*deshading*).

L'acquisizione, come tutte le misure, è soggetta ad errori. La ricostruzione, quindi, deve essere tale da filtrare il rumore di misura. Inoltre, per poter ricostruire forme complesse, il sistema deve disporre di più di un sensore. Alternativamente — e, in generale, equivalentemente — la ricostruzione può essere operata a partire da più acquisizioni. Le informazioni provenienti dai vari sensori devono essere integrate tra loro. Questo implica due ulteriori fasi di elaborazione: la registrazione dei sensori e la fusione delle informazioni.

La ricostruzione della superficie visibile a partire da punti ad essa appartenenti è un problema ben formalizzato [Ter88] [PST00], ma generalmente mal condizionato. In generale, la ricerca di un compromesso tra l'aderenza ai campioni ed il filtraggio del rumore implica un'ottimizzazione non lineare. Metodi iterativi e euristiche legate alle caratteristiche del sistema d'acquisizione e/o del tipo di oggetto acquisito possono essere utilizzati per contenere la complessità del problema.

Una volta ottenuto un modello affidabile dell'oggetto acquisito, la sua rappresentazione (finora descritta in un formato adatto alla procedura di ricostruzione) può essere trasformato in un formato più adatto alle applicazioni che ne faranno uso. Per esempio, può essere sottoposto ad una procedura di compressione per diminuirne la dimensione.

I sistemi di acquisizione possono quindi essere descritti ad alto livello utilizzando lo schema di figura 1.1:

- Un insieme di dispositivi di acquisizione è utilizzato per catturare dell'informazione riguardo alle caratteristiche geometriche e cromatiche della scena di interesse.
- I dati acquisiti sono utilizzati per ricostruire informazioni correlate alla posizione nello spazio $3D$ (per esempio, punti tridimensionali, attributi di luminanza o colore di regioni $3D$).
- Generalizzando questi dati è possibile ottenere una ricostruzione $3D$ anche nelle regioni nelle quali non è stato effettuato il campionamento e ridurre gli effetti degli errori di misura.
- Le informazioni ottenute dai singoli dispositivi possono essere integrate tra di loro.
- Il processo può essere iterato per ottenere una soluzione migliore.
- La rappresentazione ottenuta può essere ottimizzata o trasformata, esprimendo la superficie mediante un paradigma più adatto ai successivi utilizzi.

Ogni fase di elaborazione può essere resa più efficiente incorporando conoscenza a priori sugli oggetti che compongono la scena o sulle modalità di acquisizione. Le tecniche sviluppate per situazioni generali possono quindi essere migliorate personalizzandole sul caso particolare.

Va inoltre notato che lo schema può essere esteso per includere alcune applicazioni che fanno uso di *motion capture*, quali animazione di modelli $3D$ o *augmented reality*. Esse presuppongono un'analisi e/o una sintesi di oggetti che modificano (in modo relativamente lento, cioè rispetto alle capacità degli strumenti) le loro caratteristiche nel tempo. La trattazione approfondita dei problemi legati a questa disciplina è al di fuori dello scopo di questo lavoro.

1.3. Obiettivo della tesi

Dato l'ampio spettro di applicazioni e di requisiti, i sistemi e le tecniche da essi utilizzate vengono in generale sviluppate dando maggior peso ad alcune problematiche piuttosto che ad altre.

Una metodologia consente di elevare il livello di astrazione della trattazione, epurandola dai dettagli implementativi legati alle esigenze dell'applicazione specifica. L'uso di una metodologia consente la valutazione delle funzionalità e l'esplicitazione delle relazioni tra le varie fasi elaborative piuttosto che delle prestazioni delle singole tecniche.

1. Introduzione

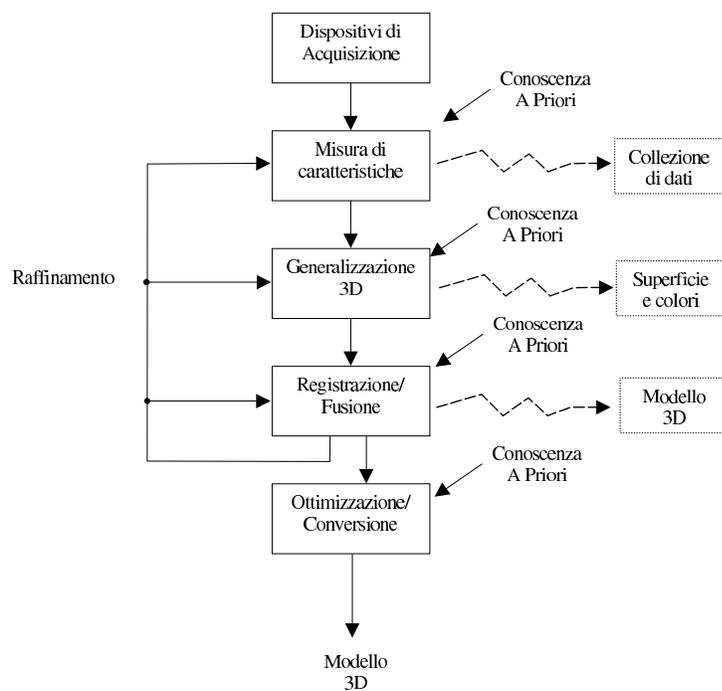


Figura 1.1.: Schema ad alto livello dei sistemi di acquisizione 3D

Lo scopo di questa tesi è pertanto lo studio di una metodologia di elaborazione e di sintesi di immagini tridimensionali.

Maggiore risalto verrà dato alle problematiche della ricostruzione, in quanto costituiscono il nocciolo del problema della digitalizzazione. In questo ambito, particolare considerazione verrà data alla ricostruzione di superficie a partire da punti campionati. I punti, infatti, sono le primitive geometriche minime e pertanto le tecniche di ricostruzione a partire da essi consentono un approccio più generale, meno legato a particolari caratteristiche del sistema di acquisizione o dell'oggetto acquisito.

La ricostruzione della superficie è una procedura onerosa in termini di costo computazionale e generalmente poco robusta: lo spazio di ricerca della soluzione ottimale è vasto e con molti minimi locali.

L'uso di tecniche multirisoluzione può migliorare sia l'efficienza computazionale che la sua robustezza. Il mondo osservato è generalmente composto da oggetti che possono essere caratterizzati a differenti livelli di dettaglio. L'utilizzo di paradigmi in grado di trattare le informazioni a diverse scale può semplificare gli algoritmi e renderli computazionalmente più leggeri.

Un approccio di questo tipo ben si coniuga alla struttura iterativa della procedura di digitalizzazione. In un primo momento viene prestata maggiore attenzione alle informazioni ad ampia scala. Il risultato è un modello poco dettagliato, ma che riesce a ricostruire le caratteristiche salienti dell'oggetto acquisito. Questo modello può essere utilizzato, ripercorrendo le varie fasi di elaborazione, per migliorare i parametri stimati inizialmente.

L'utilizzo di modelli poco particolareggiati implica in genere una diminuzione del costo di elaborazione. A seconda delle necessità, si può quindi ottenere un modello in un tempo minore o utilizzare il guadagno computazionale così ottenuto per migliorare l'accuratezza. D'altra parte, la minore richiesta di risorse computazionali può comportare un risparmio nel costo delle apparecchiature.

Un modello gerarchico è composto di diversi sottomodelli, comunemente chiamati *livelli* o *strati*. I livelli sono organizzati in modo gerarchico: l'effetto di un livello nella ricostruzione dipende dal livello superiore. I livelli sono a loro volta composti da diversi elementi i quali in genere dipendono da un parametro di scala. Il parametro di scala di un dato livello caratterizza il comportamento in spazio/frequenza del livello stesso. Questi modelli (detti anche "multirisoluzione") sono in grado di fornire il comportamento in frequenza del segnale in una determinata finestra temporale. Questa proprietà li rende uno strumento adatto allo studio ed all'elaborazione di segnali non stazionari.

1. Introduzione

L'approccio metodologico sviluppato in questa tesi si basa su modelli multirisoluzione al fine di ottenere una gerarchia di approssimazioni del segnale a scale diverse, a risoluzione via via crescente. Questa rappresentazione agevola l'individuazione del rumore ed il suo filtraggio. Anche le elaborazioni basate sul confronto di segnali — per esempio, allineamento, riconoscimento, estrazione di *feature* — sono facilitate. Esse possono essere effettuate per approssimazioni successive, a partire dalle ricostruzioni a bassa risoluzione (in modo robusto, per via dell'assenza dei dettagli), raffinando i risultati ottenuti alla scala precedente.

Inoltre, le elaborazioni seguenti alla ricostruzione possono avvantaggiarsi di una rappresentazione mediante un modello gerarchico. Per esempio, una modifica ad un dato livello si ripercuote anche ai livelli gerarchicamente inferiori [FvFH90]. Ciò rende l'editing di modelli complessi più agevole e intuitivo.

Il contributo innovativo della presente tesi può essere riassunto nei seguenti punti:

- definizione di una metodologia modulare e strutturata per la generazione di modelli tridimensionali;
- ideazione e valutazione di HRBF — un paradigma gerarchico basato sulla combinazione delle teorie del filtraggio e delle reti neurali — come paradigma multirisoluzione per la ricostruzione di superfici a partire da dati campionati affetti da rumore;
- comparazione di HRBF con MRA — paradigma gerarchico basato sulla teoria della trasformata wavelet — per esaminare la capacità di ricostruzione, espressa come accuratezza e robustezza, nell'ottica di una realizzazione su architetture digitali ad alte prestazioni;
- riformulazione degli algoritmi HRBF in chiave digitale;
- contributo rilevante alla creazione del modulo di ricostruzione di Autoscan — un sistema di digitalizzazione di oggetti tridimensionali —, nonché valutazione delle prestazioni di accuratezza nel campionamento di superfici.

1.4. Organizzazione della tesi

Nel capitolo 2 sarà trattato in modo più dettagliato e formale il problema della digitalizzazione. In particolare, si intende fornire al lettore, sia pure senza pretesa di esaustività, una panoramica sui dispositivi di

digitalizzazione attualmente disponibili, sulle problematiche della ricostruzione tridimensionale, e sulle elaborazioni alle quali solitamente un modello tridimensionale viene sottoposto per poterlo effettivamente utilizzare nelle applicazioni. L'intenzione è di descrivere sommariamente, ma in modo quanto più possibile completo, gli approcci più utilizzati per ottenere la digitalizzazione di oggetti reali e le caratteristiche che sono richieste ad un modello tridimensionale.

Successivamente, nel capitolo 3, sarà introdotta la metodologia di analisi e sintesi di immagini tridimensionali sviluppata in questa tesi. Essa è in grado di considerare in modo organico i passi necessari ad affrontare il problema della generalizzazione di informazioni provenienti da differenti sorgenti.

Nel capitolo 4 saranno trattati in modo più approfondito i paradigmi computazionali gerarchici MRA e HRBF, fornendo elementi delle teorie sulle quali si fondano. Ciò consentirà l'indagine sulle proprietà di HRBF — paradigma ideato e sviluppato nel contesto di questa tesi — come approssimatore di funzioni. Inoltre ciò ageverà la comparazione con il paradigma gerarchico più utilizzato in letteratura: MRA.

Nel capitolo 5 sarà descritto lo scanner tridimensionale Autoscan. Il suo flusso di elaborazione e la sua struttura modulare rispecchiano le fasi individuate nella metodologia proposta nel capitolo 3, della quale Autoscan è una istanza implementativa.

Infine, il capitolo 6 riassumerà le idee e i risultati esposti per tracciare le conclusioni e le possibili linee di ricerca future.

1. *Introduzione*

2. Metodi e tecniche per la digitalizzazione

Nel capitolo precedente è stata tracciata a grandi linee la procedura per ottenere una copia digitale di un oggetto reale, evidenziando le fasi in cui tale procedura può essere scomposta. In questo capitolo si approfondirà la descrizione delle singole fasi, discutendo gli approcci utilizzati in letteratura.

La digitalizzazione può essere scomposta in tre macro-fasi: acquisizione, ricostruzione ed ottimizzazione. Durante l'*acquisizione* l'oggetto viene analizzato mediante opportuni dispositivi che ne misurano alcune caratteristiche. La *ricostruzione* utilizza queste informazioni per generare un modello tridimensionale dell'oggetto analizzato. Il modello può essere quindi sottoposto alla fase di *ottimizzazione* dove la sua rappresentazione viene trasformata in una semanticamente equivalente, ma aderente alle specifiche dell'applicazione che farà uso della copia digitale.

Particolare attenzione verrà posta alla fase di ricostruzione, argomento centrale di questa tesi. È però importante comprendere, anche se non in dettaglio, quali siano le peculiarità delle singole fasi e come interagiscano tra di loro. Le diverse fasi sono infatti descrivibili come problemi a se stanti, ma interdipendenti.

Inizialmente, §2.1, verrà data una descrizione ad alto livello delle attività che permettono di realizzare la procedura di digitalizzazione. Successivamente, §2.2, verranno introdotti i formalismi utilizzabili per descrivere le caratteristiche degli oggetti reali. In particolare, per gli scopi di questa tesi, l'interesse è rivolto verso la modellizzazione di caratteristiche geometriche e cromatiche. Nel paragrafo 2.3 verranno illustrate le tecniche utilizzate per misurare le caratteristiche degli oggetti. La collezione di misure ottenute dalla fase di acquisizione sono in genere affette da rumore e non costituiscono un modello continuo dell'oggetto. Per ottenere un modello realistico è necessario operare una generalizzazione delle misure acquisite. Le tecniche per operare questa operazione verranno presentate nel paragrafo 2.4. Le applicazioni che usano i modelli tridimensionali coprono un vasto spettro. Il formalismo utilizzato

2. Metodi e tecniche per la digitalizzazione

per la ricostruzione non sempre si presta ad un utilizzo diretto da parte di queste applicazioni. Il modello 3D deve quindi essere sottoposto ad una fase di ottimizzazione finalizzata ad una particolare applicazione obbiettivo. Alcune delle tecniche utilizzate in questo campo verranno illustrate in §2.5.

2.1. La digitalizzazione vista ad alto livello

Come già accennato nel paragrafo 1.2, la digitalizzazione di un oggetto è definibile come un processo gerarchico e modulare (cf. fig. 1.1). In questo paragrafo approfondiremo la descrizione allora strutturandola maggiormente. In particolare, le attività necessarie per ottenere a digitalizzazione possono essere raggruppate in tre fasi: acquisizione, ricostruzione e ottimizzazione.

2.1.1. Acquisizione

Il primo livello è deputato al campionamento delle proprietà geometriche e cromatiche della superficie ed è denominato fase di *acquisizione*. I sensori, attraverso una prima elaborazione, raccolgono un insieme finito e discreto di caratteristiche geometriche e visive della superficie quali punti, spigoli, colore e tessitura. Questa fase non è monolitica, ma si compone a sua volta di diversi passi: pianificazione delle acquisizioni, calibrazione dei sensori e misura.

Un oggetto reale difficilmente è osservabile mediante un solo sensore ed una sola acquisizione. Un sensore ha un campo di acquisizione (cioè la regione di spazio dalla quale può ricevere segnali) ed una risoluzione (cioè un livello di dettaglio) limitati. Inoltre, a causa di possibili ostacoli, non tutti i punti del campo di acquisizione sono rilevabili. In particolare, assumendo che gli oggetti da acquisire formino una superficie chiusa, una parte della superficie sarà coperta dalla parte visibile dell'oggetto stesso. È pertanto necessario studiare l'insieme di acquisizioni che permettono di coprire tutta la superficie visibile dell'oggetto. Questa operazione viene effettuata nella fase chiamata *view planning*. Essa può essere supportata da procedure automatiche studiate nell'ambito della visione artificiale e precisamente nell'area denominata *active vision*. Il problema è conosciuto in letteratura come *next best view* (NBV) [Pit99]: dato un oggetto parzialmente acquisito, individuare il posizionamento relativo sensore/oggetto che permette di rilevare la porzione di superficie non ancora acquisita. Sebbene necessario per la totale automatizzazione della digitalizzazione di oggetti, l'NBV non rientra tra le priorità di uno scanner generico. Per progetti di grandi dimensioni o con

costi per acquisizione elevati, tuttavia, una accurata pianificazione delle sessioni di acquisizione è quasi necessaria; la disponibilità di un tool di pianificazione automatico è quindi consigliabile [LPC⁺00]. Va notato che il problema NBV rientra naturalmente in uno schema iterativo del processo di digitalizzazione.

Per trasformare le informazioni rilevate dai sensori in un unico modello tridimensionale, è necessario determinare un modello matematico della trasformazione effettuata dai sensori. Questo modello contiene in generale due insiemi di parametri: i parametri interni (riguardanti cioè il singolo sensore) ed esterni (riguardanti le relazioni fra i sensori). La fase di determinazione di questi parametri viene chiamata *calibrazione* del sistema. La procedura di calibrazione varia da dispositivo a dispositivo, e non sempre è necessario effettuarla ad ogni acquisizione. Alcuni dispositivi di acquisizione, infatti, non hanno necessità di essere ricalibrati in quanto la calibrazione è effettuata al momento della fabbricazione. Generalmente, i dispositivi precalibrati esibiscono una flessibilità (cioè la capacità di essere utilizzati in condizioni ambientali differenti) minore dei dispositivi che devono essere calibrati ad ogni acquisizione. Per contro, un dispositivo precalibrato gode di una maggior facilità di utilizzo.

La fase di *misura* delle caratteristiche è la meno formalizzabile per via della molteplicità delle procedure utilizzate. Nel paragrafo 2.3.1 verranno descritti diversi sistemi di acquisizione. Le differenze consistono sia nei principi fisici sfruttati per la misura (per esempio, ottici, meccanici, acustici) sia nella tecnica di misura (per esempio, triangolazione, spostamento di fase). Lo scopo di questa fase è collezionare un insieme di misure di caratteristiche geometriche e cromatiche. Le informazioni raccolte da ogni sensore (o coppia di sensori, per alcuni dispositivi) vengono chiamate *viste* (*view*). Tipicamente, dopo questa fase vengono restituite la posizione nello spazio tridimensionale di primitive geometriche quali punti o spigoli, e le informazioni sulle caratteristiche cromatiche della superficie. Due gli approcci che vengono utilizzati per l'acquisizione delle caratteristiche. Il primo approccio, comunemente denominato *attivo*, utilizza un dispositivo per irradiare l'oggetto con un pattern facilmente riconoscibile dai sensori. Tipicamente, il pattern è costituito da punti o da linee. Il secondo approccio, chiamato *passivo*, è caratteristico della Computer Vision. Vengono in questo caso utilizzate solo immagini o sequenze di immagini con illuminazione ambientale.

2.1.2. Ricostruzione

Una volta ottenuta la collezione di primitive geometriche, le procedure di *ricostruzione* generalizzano le informazioni raccolte in modo da generare

2. Metodi e tecniche per la digitalizzazione

una rappresentazione continua della superficie che copra anche le regioni dove non sono definite primitive. Analogamente, le informazioni legate al colore sono elaborate per generare una rappresentazione continua del colore che viene chiamata *color field* (§6.2). Inoltre, per determinare il colore reale dell'oggetto è necessario operare una deconvoluzione delle sorgenti luminose. In generale, l'elaborazione dell'informazione di colore è seguente alla ricostruzione della geometria della superficie, in quanto il colore apparente dipende dalla posizione dell'osservatore, da quella della sorgente luminosa, ma anche dall'orientamento della superficie osservata (normale) [FvFH90].

La ricostruzione di una superficie a partire da punti campionati su di essa è un problema trattato in diverse discipline. In ambito neurale, viene visto come un problema di apprendimento di una funzione da esempi. In ambito statistico, i punti sono campioni di una distribuzione stocastica e la ricostruzione della superficie corrisponde alla stima ottimale dei parametri della distribuzione stessa. È una particolare istanza del problema di *learning* [Vap99]. In generale, il problema è formulabile come una approssimazione di funzione dove la funzione approssimante può essere scelta solo fra quelle appartenenti ad uno spazio determinato dal paradigma scelto per l'approssimazione. Tale spazio è definito dalle sue funzioni base, la forma delle quali è governata da parametri che vengono scelti in modo da raggiungere un compromesso tra l'aderenza ai punti campionati e una certa funzione costo che rende merito delle caratteristiche desiderate (ad esempio, dimensione del modello, *smoothness*). Questo tipo di rappresentazione è in forma *esplicita*. La superficie viene cioè rappresentata come una funzione definita su un dato dominio. Al problema della ricostruzione si somma quindi il problema della *topologia* della superficie. Tale problema non si pone se le superfici sono rappresentate in forma *implicita*. Ciò accade frequentemente nella grafica computerizzata, dove sono stati sviluppati soprattutto approcci costruttivi basati su tecniche volumetriche [AS96] o mesh triangolari [HDD⁺92]. Una revisione dei metodi più interessanti è riportata nel paragrafo 2.4.

La ricostruzione presenta principalmente due tipi di problemi: l'errore di misura e la fusione delle informazioni dei diversi sensori. Infatti, l'acquisizione, come tutte le misure, è soggetta ad errori. La ricostruzione, quindi, deve essere tale da filtrare il rumore di misura. Inoltre, per poter ricostruire forme complesse, il sistema deve disporre di più di un sensore o acquisire l'oggetto mediante più sessioni di acquisizione.

L'errore di misura viene parzialmente corretto sfruttando la correlazione tra misure spazialmente vicine utilizzando un opportuno paradigma rappresentativo che codifichi la conoscenza a priori sul problema.

Tale conoscenza può essere molto dettagliata (e in tal caso si userà un modello parametrico dell'oggetto, come, per esempio, in [BV99]) o solo superficiale (e quindi richiederà vincoli generalmente soddisfatti, per esempio, smoothness [PG90]).

Generalmente, per ogni sensore (o coppia di sensori) viene ricavato un insieme di informazioni utile per descrivere una superficie definita su un piano. Tali superfici, paragonabili ad un bassorilievo, vengono chiamate superfici $2\frac{1}{2}D^1$ per distinguerle da quelle completamente $3D$, che, continuando la metafora, possono essere paragonate a statue a tuffo. Le informazioni provenienti dai vari sensori devono essere integrate tra loro. Questo implica due ulteriori fasi di elaborazione: la *registrazione* dei dati prodotti dai sensori e la *fusione* di tali informazioni.

La *registrazione* consiste nell'individuare la posizione relativa dei dati acquisiti rispetto ad un unico sistema di riferimento (figura 2.1). Quest'operazione può essere parzialmente effettuata durante la calibrazione del sistema d'acquisizione (§2.1.1). L'uso di sensori precalibrati, o di bassa qualità, o, infine, di procedure di calibrazione poco accurate possono richiedere una ulteriore elaborazione. Durante le acquisizioni, infatti, la quantità di dati acquisiti è generalmente maggiore che durante la fase di calibrazione. Partendo da un allineamento iniziale abbastanza buono ottenuto durante la fase di calibrazione, i dati dell'acquisizione possono essere utilizzati per raffinarlo. Questa operazione viene effettuata individuando in *viste* differenti le regioni comuni e calcolando la trasformazione che ne massimizza la sovrapposizione.

La *fusione* consiste nell'unire le informazioni acquisite dalle diverse sorgenti in un'unica rappresentazione. Questo può implicare la risoluzione di ambiguità (ad esempio, [CL96]). Il caso più comune è rappresentato dalla mancata sovrapposizione della stessa caratteristica acquisita da sensori diversi. Regioni comuni a *viste* differenti, anche dopo la registrazione, non sono generalmente coincidenti: per ottenere una superficie senza brusche variazioni bisogna mediare sulle superfici ricostruite.

Due sono quindi gli approcci possibili:

- ricostruzione delle superfici $2\frac{1}{2}D$, seguita da registrazione e fusione delle ricostruzioni;
- registrazione e fusione delle informazioni provenienti dai sensori e poi ricostruzione della superficie $3D$.

È possibile utilizzare entrambi gli approcci. Una prima ricostruzione delle superfici $2\frac{1}{2}D$ viene utilizzata per allineare i dati, i quali, una volta

¹Talvolta vengono indicate anche come superfici $2.5D$

2. Metodi e tecniche per la digitalizzazione

fusi, vengono utilizzati per una ricostruzione $3D$. L'algoritmo più utilizzato per la soluzione di questo problema è quello detto *Iterative Closest Point* (ICP) [BM92], basato sulla ricerca delle somiglianze tra regioni di *viste* diverse. Data la stretta correlazione tra le due fasi, esse vengono talvolta considerate parte di una fase sola, chiamata *integrazione* (delle *view*).

In generale, la registrazione e la fusione devono essere applicabili a dati provenienti da sensori di diversa natura. Un caso molto frequente nella pratica — e di particolare interesse per il problema trattato in questa tesi — è la fusione delle informazioni geometriche e di quelle cromatiche. In generale, a meno di condizioni ambientali controllate, è necessario avere più di un sensore per il colore. Come già accennato, il colore apparente dipende dalla posizione relativa sorgente luminosa/osservatore e dalla direzione della normale della superficie. La disponibilità di più di un sensore colore permette di comporre le informazioni di colore pesandole con una misura di affidabilità [DTM96] [PARD⁺98].

La ricostruzione avviene generalmente per approssimazioni successive — tramite procedure iterative [TL94]. L'utilizzo di paradigmi multi-risoluzione per la rappresentazione dei modelli tridimensionali si colloca quindi in modo naturale in questo schema operativo. Esso consente la realizzazione di queste fasi in modo efficiente utilizzando modelli a bassa risoluzione per le prime iterazioni e incrementando via via la ricchezza di dettagli nelle fasi seguenti. L'efficienza deriva da diversi fattori:

- Un modello a bassa risoluzione ha pochi parametri. Ciò comporta un basso costo computazionale per il calcolo delle superfici $2\frac{1}{2}D$. Analogamente, i calcoli necessari per valutare l'allineamento delle superfici parziali sono contenuti.
- La ricerca delle trasformazioni ottimali per l'integrazione non rischia di rimanere intrappolata in minimi locali dovuti ai dettagli ed al rumore.

La ricostruzione è il tema centrale di questa tesi. Per questo motivo sarà dedicato ampio spazio (§2.4) alle tecniche utilizzate per risolvere il problema della ricostruzione.

2.1.3. Ottimizzazione e trasformazione

Una volta ottenuto un modello affidabile dell'oggetto acquisito, la sua rappresentazione (finora descritta in un formato adatto alla procedura di ricostruzione) può essere trasformato in un formato più adatto alle applicazioni che ne faranno uso.

2.1. La digitalizzazione vista ad alto livello

Uno degli scopi principali di questa fase è la traduzione in un formato che consenta di rappresentare la stessa superficie, ma con un ingombro (cioè utilizzando un numero di valori numerici) inferiore. La trasformazione di fatto equivalente ad una compressione *lossless*.

Succede spesso che un'applicazione abbia delle specifiche tali da non consentire l'utilizzo di modelli molto dettagliati. L'eliminazione di dettagli consente di ottenere modelli computazionalmente più leggeri e meno ingombranti. Questa trasformazione corrisponde ad una compressione *lossy*.

Come verrà descritto nel paragrafo 2.5.1, queste trasformazioni possono essere agevolate dall'uso di una rappresentazione della superficie mediante paradigmi gerarchici o multirisoluzione (figura 2.2).

Altre trasformazioni molto utilizzate nelle applicazioni che utilizzano i modelli tridimensionali sono:

Modifica: Spesso il modello tridimensionale è solo la base di partenza per un modello particolare. È il caso, per esempio, dei modelli utilizzati per i videogiochi: un personaggio viene modellato materialmente, digitalizzato, e poi gli vengono aggiunti accessori o dettagli. Il formato di rappresentazione della superficie può facilitare o ostacolare questa operazione. Per esempio, una rappresentazione tramite una paradigma gerarchico permette all'operatore di intervenire a più livelli di dettaglio, mantenendo intatti i rimanenti (*multiresolution editing*, [LSS⁺98]), oppure, come illustrato in figura 2.3, di modificare la forma generale di parte dell'oggetto intervenendo solo sul modello a bassa risoluzione.

Animazione: Le animazioni (sia per intrattenimento che per scopi didattici) spesso richiedono che parti del modello vengano modificate. Per esempio, il modello di un essere umano potrebbe dover flettere le braccia o correre. Una paradigma gerarchico consente di definire il moto a basso livello di risoluzione, ottenendo automaticamente il movimento consistente del dettaglio.

Morphing: Un effetto speciale molto utilizzato è la trasformazione (generalmente mediante una transizione smooth) di un modello in un altro, come illustrato in figura 2.4. Il morphing viene realizzato definendo una relazione suriettiva tra gli elementi del modello di partenza e gli elementi del modello finale. Una rappresentazione mediante una paradigma gerarchico facilita quest'operazione, perché la relazione suriettiva può essere limitata agli elementi (generalmente pochi) che compongono gli strati a bassa risoluzione [LDSS99].

2. Metodi e tecniche per la digitalizzazione

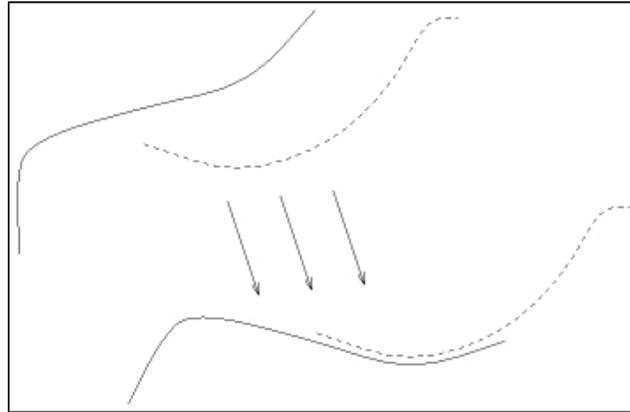


Figura 2.1.: Registrazione di due curve.

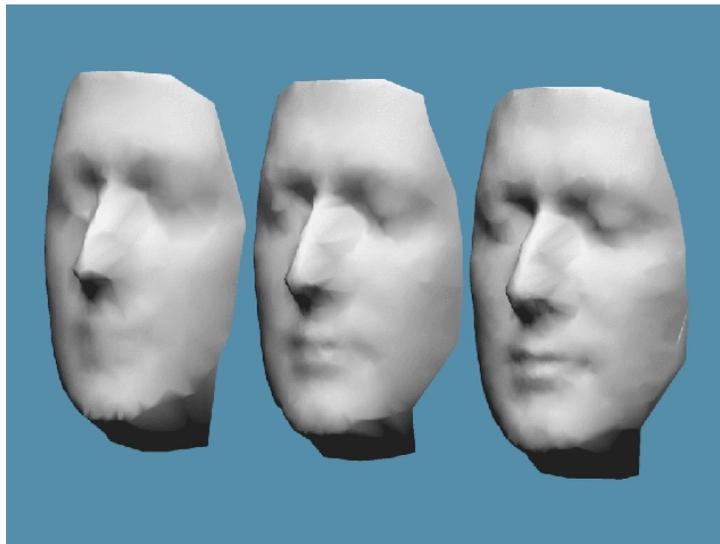


Figura 2.2.: Rappresentazione a diversi livelli di dettaglio dello stesso oggetto.

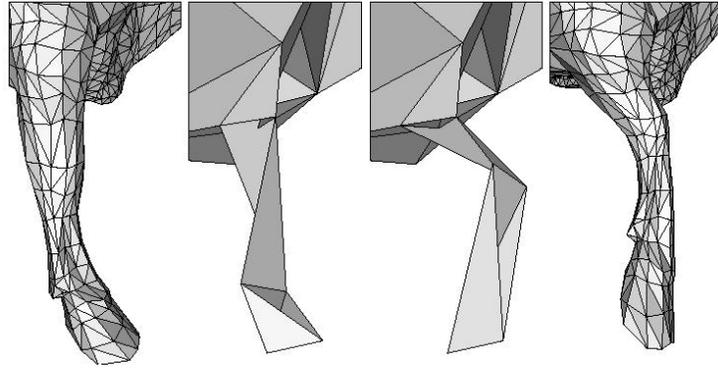


Figura 2.3.: Modellazione multirisoluzione [GSS99].



Figura 2.4.: Morphing [LDSS99].

2.2. Rappresentazione di oggetti reali

Definiamo in termini matematicamente più rigorosi le caratteristiche della superficie di un oggetto. Esse si possono dividere in due classi: le caratteristiche *geometriche* (quali la posizione dei punti della superficie o la direzione delle normali della superficie) e le caratteristiche *cromatiche* (ad esempio, colore, trasparenza, riflettanza).

Un modello tridimensionale di un oggetto reale deve quindi essere descritto tramite un formalismo capace di rappresentare entrambi i tipi di caratteristiche. In generale, vi sono formalismi in grado di descrivere la geometria e tecniche per descrivere le caratteristiche legate al colore, che sono sostanzialmente indipendenti dal formalismo utilizzato per la geometria.

2.2.1. Rappresentazione della geometria

Una superficie è definibile come un manifold (vedi glossario in §6.2) bidimensionale definito in uno spazio tridimensionale. I paradigmi utilizzati per la rappresentazione di superfici possono essere ricondotti a due classi [LSS⁺98]: funzioni implicite (spesso mesh poligonali) e funzioni esplicite (per le quali è necessario un dominio di definizione).

Le superfici rappresentate da funzioni implicite sono definite come l'insieme di punti che soddisfa una condizione del tipo:

$$f(\mathbf{x}, \theta) = 0, \quad \mathbf{x} \in \mathbb{R}^3$$

dove θ è un insieme di parametri (eventualmente vuoto) che governa alcune caratteristiche della funzione.

Una rappresentazione esplicita di una superficie necessita di un dominio di riferimento, chiamato spesso *spazio parametrico*. Una superficie del genere è quindi una funzione $f : \mathbb{U} \rightarrow \mathbb{R}^3$ del tipo:

$$\mathbf{x} = f(\mathbf{u}, \theta), \quad \mathbf{u} \in \mathbb{U}, \mathbf{x} \in \mathbb{R}^3$$

dove \mathbf{u} sono le coordinate nello spazio parametrico \mathbb{U} , mentre θ è l'insieme di parametri che governa le caratteristiche della superficie. Generalmente si ha che $\mathbb{U} \subseteq \mathbb{R}^2$. Tuttavia, lo spazio parametrico \mathbb{U} può coincidere con due coordinate dello spazio euclideo. In questo caso si ottiene una superficie $2\frac{1}{2}D$, $f : \mathbb{R}^2 \rightarrow \mathbb{R}$:

$$z = f(\mathbf{x}, \theta), \quad \mathbf{x} \in \mathbb{R}^2, z \in \mathbb{R}$$

Fra le rappresentazioni esplicite, i paradigmi gerarchici rappresentano un caso di particolare interesse per questa tesi. Essi assumono la

seguinte forma:

$$\mathbf{x} = a(\mathbf{u}_0, \theta_0) + \sum_{i=1}^L d_i(\mathbf{u}_i, \theta_i), \quad \mathbf{x} \in \mathbb{R}^3, \forall i \mathbf{u}_i \in \mathbb{U}_i$$

dove $a(\cdot, \cdot)$ è una funzione che approssima le caratteristiche a larga scala della superficie, mentre $d_i(\cdot, \cdot)$ sono un insieme di funzioni che descrivono i dettagli a risoluzioni via via crescenti. Le funzioni sono definite, in generale, su spazi parametrici differenti, $\{\mathbb{U}_0, \dots, \mathbb{U}_L\}$, tali che, per ogni j , $0 \leq j < L$, è definita una funzione suriettiva $\pi_j : \mathbb{U}_j \rightarrow \mathbb{U}_{j+1}$, e le coordinate nei differenti spazio parametrici rispettano questa catena di mappaggi: $\mathbf{u}_1 = \pi_0(\mathbf{u}_0)$, $\mathbf{u}_2 = \pi_1(\mathbf{u}_1)$, \dots , $\mathbf{u}_L = \pi_{L-1}(\mathbf{u}_{L-1})$.

Rappresentazioni implicite

La rappresentazione implicita più utilizzata è la mesh poligonale [FvFH90] (figura 2.5). In generale una mesh poligonale viene rappresentata attraverso un insieme di punti e di poligoni che connettono i punti. Una rappresentazione di questo tipo consente di diminuire l'ingombro in quanto su ogni vertice della mesh, generalmente, incide più di un poligono. Se l'insieme di punti viene rappresentato mediante un array e include i vertici dei poligoni che compongono la mesh, l'insieme dei poligoni può essere rappresentato utilizzando solo la collezione delle tuple di interi contenenti le posizioni dei vertici nell'array dei punti. Più formalmente, una mesh poligonale M può essere formalizzata come la tripla $\{V, E, P\}$, dove $V = \{v_i \in \mathbb{R}^3 \mid i = 1, \dots, z_v\}$ è l'insieme dei vertici, $E = \{e_i = (v_j, v_k) \mid v_j, v_k \in V, i = 1, \dots, z_e\}$ è l'insieme dei lati e $P = \{p_i \in 2^{|E|} \mid i = 1, \dots, z_p\}$, dove z_v , z_e e z_p sono rispettivamente il numero di vertici, di lati e di poligoni che compongono la mesh. Appare dunque evidente che anziché memorizzare l'insieme dei lati, E , e l'insieme dei poligoni, P , come gli insiemi dei vertici tridimensionali da essi toccati, è molto più economico rappresentarli come collezione di indici (puntatori) che si riferiscono alle posizioni dei vertici nell'insieme V . Oltre al minore ingombro, questa rappresentazione ha il vantaggio di esplicitare la condivisione di vertici. Ciò è molto utile per l'efficienza delle elaborazioni che modificano la forma della superficie modificando la posizione dei vertici. Una rappresentazione di questo genere può inoltre essere facilmente estesa. Alcuni formati grafici di questo tipo, infatti, considerano la possibilità di aggiungere informazioni quali:

- segmenti (per le applicazioni che richiedono l'esplicitazione degli spigoli)

2. Metodi e tecniche per la digitalizzazione

- normali (ai vertici e/o ai triangoli)

La rappresentazione come mesh poligonale (in generale triangoli) è molto utilizzata per la visualizzazione, principalmente a causa del fatto che molte schede acceleratrici grafiche ne permettono l'elaborazione hardware.

I modelli poligonali non possiedono infatti derivate continue. Questo può talvolta causare una visualizzazione poco realistica. Le superfici biologiche, per esempio, sono generalmente smooth. Opportune procedure di *visualizzazione* (ad esempio, lo *shading* di Gouraud) permettono, tuttavia, di conferire alle superfici poligonali un aspetto parzialmente liscio, ottenuto interpolando le normali ai vertici. L'osservatore, per effetto delle variazioni di tonalità del colore, percepisce la sensazione di smoothness della superficie, ma, come illustrato dalla figura 2.6, la geometria non cambia. Questo fatto è bene visibile lungo il profilo dell'oggetto, dove l'effetto dello shading non è realizzabile.

Superfici poligonali possono essere utilizzate per approssimare superfici di topologia arbitraria e con grado di smoothness visivamente accettabili [FvFH90]. Tuttavia, il numero di poligoni necessario allo scopo può superare le risorse del sistema di visualizzazione. Poiché, come precedentemente accennato, il numero di poligoni necessario varia a seconda della posizione dell'oggetto rispetto all'osservatore, sono state studiate delle procedure per adeguare run-time il numero di poligoni della rappresentazione nelle diverse regioni dell'oggetto visualizzato (figura 2.7). Questo tipo di elaborazione è nota come *view-dependent rendering* [Hop97]. Data la natura run-time di queste procedure, una rappresentazione efficiente dei modelli tridimensionali utilizzati è un fattore critico. Una rappresentazione gerarchica offre il vantaggio di poter selezionare rapidamente il grado di accuratezza necessario in ogni regione, data la località della rappresentazione.

Le mesh poligonali sono funzioni implicite lineari a tratti. In linea di massima una rappresentazione implicita è possibile anche per funzioni di grado superiore al primo. Superfici del genere non sono però facili da elaborare. Per questo motivo si preferisce in generale utilizzare rappresentazioni esplicite parametriche.

In alcune applicazioni vengono utilizzate le quadriche [FvFH90] per la vasta gamma di forme rappresentabili.

Rappresentazioni esplicite

Le superfici esplicite consentono di esprimere superfici aventi anche le derivate (fino ad un certo grado) continue. Nella maggior parte dei casi si tratta di superfici polinomiali a tratti (Bézier, B-spline e NURBS,

2.2. Rappresentazione di oggetti reali

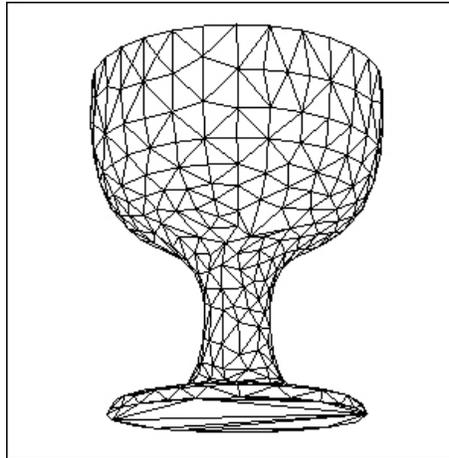


Figura 2.5.: Mesh poligonale [Hop94].

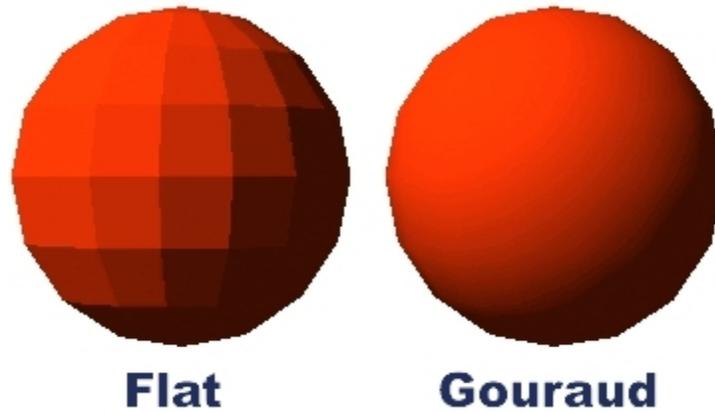


Figura 2.6.: Gouraud shading.

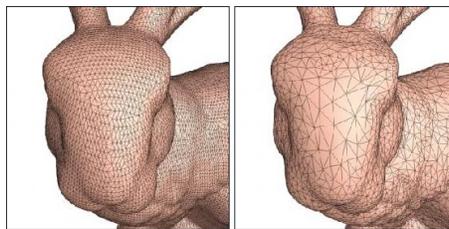


Figura 2.7.: *View-dependent rendering* [Hop97].

2. Metodi e tecniche per la digitalizzazione

per citare le più usate). Tra i vantaggi di una rappresentazione tramite polinomiali c'è anche la capacità di rappresentazione superfici complesse in modo compatto. Il numero di parametri è infatti minore del numero di vertici di una corrispondente rappresentazione poligonale.

In linea di massima, comunque, ogni superficie deve essere triangolarizzata per poter essere visualizzata. La Play Station 2 della Sony consente il rendering hardware anche di superfici di Bézier [PSX]. Il modello può quindi essere descritto usando superfici smooth e, a seconda delle risorse grafiche disponibili, viene calcolata in real-time la triangolazione più adeguata (*view-dependent rendering*).

Altri paradigmi di questo tipo fra i più utilizzati per rappresentare le superfici sono le *Non-Uniform Rational B-Splines* (NURBS) [PT97] e le *subdivision surfaces* [ZSS96] [Lev99], entrambe basate sull'interpolazione polinomiale. La superficie ricostruita da questi paradigmi è formalizzabile come una combinazione lineare di funzioni base:

$$\mathbf{x} = \sum_{i \in I} w_i b_i(\mathbf{u}), \quad \mathbf{u} \in \mathbb{U}, \mathbf{x} \in \mathbb{R}^3$$

dove $b_i : \mathbb{U} \rightarrow \mathbb{R}^3$ sono le funzioni base che caratterizzano il paradigma utilizzato, $w_i \in \mathbb{R}$ sono i coefficienti della combinazione lineare.

Le NURBS sono molto utilizzate nella modellazione di superfici per via della relativa facilità con la quale si possono associare i parametri che ne governano la forma con azioni di modellazione (figura 2.8). Inoltre è relativamente semplice definire zone di discontinuità delle derivate, descrivendo così curve e superfici di forma generale (*free-form surfaces*). Esse rappresentano un'evoluzione delle Bézier e delle curve parametriche polinomiali (ad esempio B-spline) delle quali sono un superinsieme. A parità di funzioni base impiegate, sono formalizzate mediante un numero maggiore di coefficienti, ma consentono di modellare forme più generali.

Le *subdivision surfaces*, invece definiscono una superficie come limite di una operazione di interpolazione polinomiale². La qualità della mesh risultante dipende dalle caratteristiche della mesh iniziale [ZSS96] [EDD⁺95] ed è legata alla connettività della suddivisione utilizzata [LDW97]. Le *subdivision surfaces* hanno una costruzione iterativa multirisoluzione, come evidenziato in figura 2.9, ma non una struttura gerarchica: è possibile ricostruire l'oggetto con diversi gradi di smoothness, ma l'informazione è tutta codificata nella mesh iniziale. Un esempio di applicazione della procedura di suddivisione di una mesh è mostrato in figura 2.10.

²È quindi possibile che le funzioni base non possano essere espresse in forma analitica, ma che sia possibile ottenerne una rappresentazione grafica con accuratezza arbitraria.

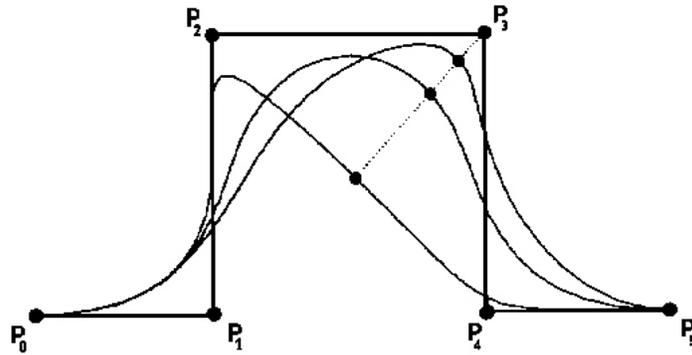


Figura 2.8.: Modellazione con una curva NURBS [PT97]. La spezzata formata dai punti di controllo P_0, \dots, P_5 fornisce la forma generale della curva. Aumentando un peso la curva diventa più aderente al punto di controllo corrispondente (P_3 nell'esempio).

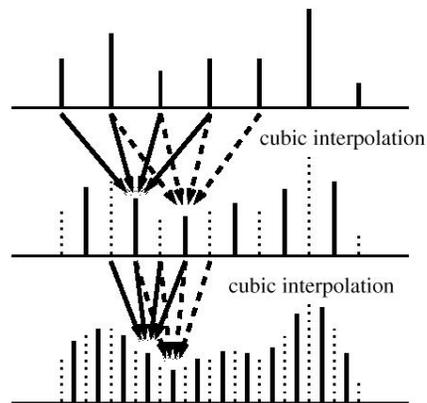


Figura 2.9.: Procedura iterativa per ottenere una suddivisione cubica [SS96].

2. Metodi e tecniche per la digitalizzazione

Di particolare interesse per il presente lavoro sono i paradigmi gerarchici. Fra di essi tratteremo, le Hierarchical Basis Functions Networks (HRBF) [BF98], le Hierarchical B-Spline [FB95], e le wavelet [Mal89] [Dau92]. Quest'ultime generalizzano anche le *subdivision surfaces*. Questi paradigmi gerarchici possono essere ricondotti ad una forma del tipo:

$$\mathbf{x} = a(\mathbf{u}) + \sum_{j \in J} d_j(\mathbf{u}), \quad \mathbf{u} \in \mathbb{U}, \mathbf{x} \in \mathbb{R}^3$$

dove $a(\cdot)$ e le $d_j(\cdot)$ sono funzioni $\mathbb{U} \rightarrow \mathbb{R}^3$ definite come combinazione lineare di opportune funzioni base. Le proprietà di queste ultime caratterizzano il paradigma utilizzato. La funzione $a(\cdot)$ fornisce un'approssimazione della superficie a scala elevata, e le funzioni $d_j(\cdot)$ codificano il dettaglio a scale via via decrescenti. La ricostruzione può essere limitata ad una scala adeguata alle applicazioni che ne usufruiranno. Bisogna notare che indipendentemente dal livello di dettaglio, lo spazio parametrico di riferimento è (formalmente) sempre lo stesso.

Data la necessità di rappresentare superfici $3D$, non è possibile usare una forma esplicita in cui la terza coordinata sia funzione delle altre due. In generale, poi, la topologia degli oggetti reali non è quella del piano. Per questo motivo, le superfici implicite vengono ricostruite a tratti. Lo spazio parametrico viene costruito accostando diverse *patch*. L'approccio è quindi quello di utilizzare una mesh poligonale (cioè un formalismo implicito) per definire il dominio di definizione (lo spazio parametrico) e poi utilizzare una curva parametrica (cioè un formalismo esplicito) per la rappresentazione dei punti della superficie.

Un oggetto viene quindi rappresentato da un reticolo di *patch* che determina la topologia e dall'insieme di coefficienti e parametri che determinano la forma delle funzione esplicita che descrive i punti della superficie dell'oggetto (approccio denominato anche *surface on surface*). Per garantire l'assenza di brusche variazioni nella superficie e nelle sue derivate (*seamless surfaces*) vengono imposte alcune condizioni di interpolazione e di continuità delle derivate al bordo delle patch adiacenti. Nella figura 2.11 è ben visibile il reticolo di *patch*.

Rispetto ad una rappresentazione implicita lineare, una rappresentazione esplicita mediante funzioni base con derivate continue ha il vantaggio di facilitare elaborazioni, come per esempio, l'estrazione di features [LM94] o il calcolo delle normali alla superficie. Rispetto ad una rappresentazione esplicita non gerarchica, i paradigmi gerarchici multi-risoluzione esibiscono generalmente una maggiore compattezza ed una maggiore facilità di elaborazione dovuta sia alla decorrelazione effettuata dalle funzioni base che alla località delle stesse.

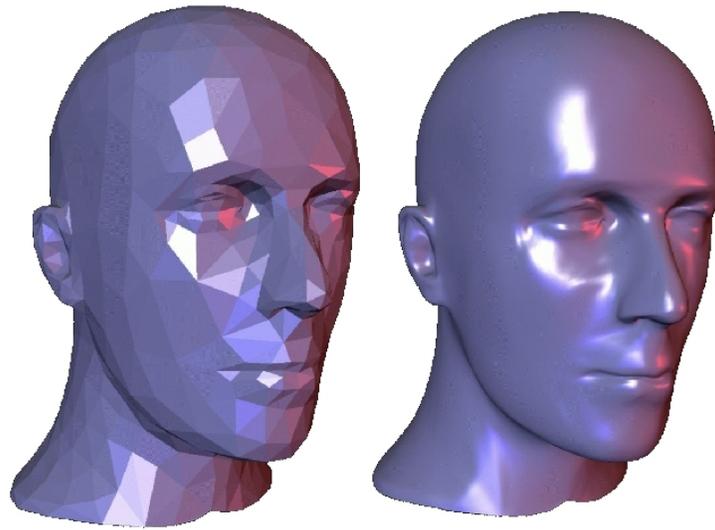


Figura 2.10.: Subdivision surface [ZSS96].



Figura 2.11.: Un reticolo di *patch* B-spline [EH96].

2. Metodi e tecniche per la digitalizzazione

Un caso particolare di paradigma gerarchico è rappresentato dalle superfici dette “normali” [GVSS00] [LMH00]. Un paradigma di questa classe è caratterizzato da una superficie base che ne determina il *genus* (vedi glossario, §6.2) e da un insieme di coefficienti che codificano l’altezza della superficie rappresentata rispetto alla superficie base. La definizione, come per le *subdivision surfaces*, può essere iterativa. Ad un primo strato di coefficienti possono aggiungersene altri. Il loro significato è sempre l’altezza rispetto alla superficie ricostruita allo strato precedente. A differenza dei paradigmi presentati precedentemente, lo spazio parametrico viene qui definito di livello in livello, in quanto è costituito dalla superficie ricostruita al livello precedente. Il vantaggio più evidente è l’efficienza della rappresentazione: pochi coefficienti per rappresentare la superficie base utilizzando una mesh poligonale (i cui vertici necessitano di tre coefficienti ognuno) e poi per ogni nuovo vertice è sufficiente un solo coefficiente.

La definizione della geometria e della topologia dello spazio parametrico non è limitata alle mesh poligonali. Per esempio, in [TM91] e [DM98] vengono utilizzate superquadriche ed ellissoidi.

2.2.2. Parametrizzazione di una superficie

Data una superficie di equazione $F(x, y, z) = 0$ nelle variabili x , y e z , si cerca una *parametrizzazione* della soluzione dell’equazione in termini di tre funzioni di due variabili, s e t :

$$(x, y, z) = (X(s, t), Y(s, t), Z(s, t)), \quad \forall x, y, z \in \mathbb{R} : F(x, y, z) = 0$$

La parametrizzazione della superficie serve per poter operare alcune elaborazioni. Innanzitutto, una superficie parametrizzata può essere utilizzata come dominio per costruire una superficie esplicita. Inoltre la parametrizzazione è necessaria per le tecniche di mapping (§2.2.3 e §2.4.5), ed è utile per elaborazioni quali il morphing (§2.5.4).

2.2.3. Rappresentazione delle caratteristiche cromatiche

Per una riproduzione realistica delle superfici bisogna codificare anche le informazioni che contribuiscono a formare il colore dell’oggetto. In generale, bisogna codificare quelli che vengono chiamati *appearance attribute*. Il colore è tuttavia, generalmente parlando, il più importante tra di essi e le tecniche utilizzate per la sua codifica possono essere applicate anche per codificare gli altri attributi (*bump*, *trasparenza*, *displacement*). Questo spiega perché semplificare la trattazione al colore non è limitante.

Modelli colore

Il colore percepito³ da un osservatore dipende da vari fattori, alcuni dei quali legati alle caratteristiche ottiche dell'oggetto (per esempio, opacità, rifrazione) ed altri alle condizioni ambientali (per esempio, illuminazione, caratteristiche ottiche del mezzo trasmissivo). Il colore apparente dell'oggetto dipende inoltre dalla posizione e dall'orientamento della superficie rispetto alle sorgenti luminose ed al punto di osservazione. Alcune caratteristiche dell'oggetto sono legate al materiale di cui esso è composto e possono essere assegnate in base a tabelle o mediante esami specifici (fotometrici o colorimetrici). Il colore apparente può essere acquisito mediante TV camere standard, possibilmente in situazioni di illuminazione ambientale controllate (cioè, assenza di ombre e riflessi). In alternativa, le immagini acquisite devono essere manipolate per eliminare gli effetti dovuti al tipo di illuminazione [PARD⁺98] [DTM96] [SI96]. Le caratteristiche ottiche vengono generalmente rappresentate mediante coefficienti numerici che rappresentano la risposta del materiale di cui l'oggetto è composto alle diverse componenti di illuminazione. La riproduzione è legata al tipo di algoritmo utilizzato per la sintesi dell'immagine. Per le caratteristiche di riflessione è largamente usato l'algoritmo di shading di Gouraud.

Il colore percepito è generato da uno spettro di onde elettromagnetiche nella banda di frequenza della luce visibile [FvFH90, §13.2]. In linea di massima, la sua rappresentazione richiederebbe quindi una quantità di informazione infinita. Poiché differenti composizioni spettrali possono dare la stessa sensazione percettiva, "colori" diversi possono essere rappresentati (per gli scopi della computer graphics) tramite alcune caratteristiche. La descrizione del colore può quindi avvenire in modo percettivamente univoco attraverso un numero limitato di basi. Questa modalità è supportata dal fatto che il nostro sistema visivo possiede tre tipi di recettori di colore (coni), ognuno dei quali sensibile ad un differente spettro luminoso. Gli spettri sono grosso modo centrati in corrispondenza delle lunghezze d'onda del blu, del verde e del rosso. Ciò equivale all'incirca ad affermare che ogni colore può essere espresso mediante una somma pesata con coefficienti positivi di rosso, verde e blu (detti perciò colori *primari*). Questa ipotesi è confermata, per la maggior parte dei colori. Permettendo ai coefficienti moltiplicativi di

³Sia ammessa qualche semplificazione. In realtà, infatti, il colore percepito dipende anche da altri fattori, per esempio il colore degli oggetti vicini. Tali effetti possono essere ignorati ai fini di questa tesi.

2. Metodi e tecniche per la digitalizzazione

assumere valori negativi⁴ si può coprire tutto lo spettro visibile.

Per ovviare alla necessità di pesi negativi per la definizione dei colori, sono stati definiti nel 1931 dalla *Commission Internationale de l'Éclairage* (CIE) tre primari standard, chiamati X , Y e Z . Per rispondere alle esigenze di uniformità percettiva⁵, nel 1976 è stato definito lo spazio colore uniforme CIE LUV. Tale spazio ha le stesse potenzialità del CIE XYZ, ma facilita le elaborazioni più comuni.

Per la grafica viene utilizzato un modello colore (RGB) che in generale non è in grado di coprire tutti i colori percepibili, ma che ha una diretta relazione con l'hardware più diffusamente utilizzato per la visualizzazione: il monitor CRT. In tale periferica, il colore viene formato eccitando fosfori di tre tipi, rosso, verde e blu. Il colore percepito è formato per addizione. Un colore può quindi essere rappresentato per mezzo di tre coefficienti (r , g , e b) che identificano la percentuale di luminosità del primario corrispondente. Una rappresentazione del genere non è adatta alle stampanti, le quali formano il colore in modalità sottrattiva ed utilizzano il modello CMY (ciano, magenta e giallo). Lo spazio CMY non coincide con lo spazio RGB. Particolare attenzione deve quindi essere posta per le elaborazioni che presuppongono una doppia modalità di output. Dispositivi di stampa più raffinati utilizzano un modello quadricromatico (il modello CMY con l'aggiunta del nero) o esacromatico (CMY, nero e due tonalità chiare di ciano e magenta).

Intuitivamente un colore può essere descritto in termini di lunghezza d'onda dominante (*hue*), di purezza (*saturation*) e di intensità (*value* o *brightness*). Esistono modelli colore basati su queste caratteristiche (HSV, HLS e HVC). Essi non descrivono uno spazio colore diverso dall'RGB, ma danno una codifica che rispecchia maggiormente le caratteristiche percettive del colore e che quindi può essere utile per talune elaborazioni.

Le regole per trasformare un colore espresso in uno spazio colore nelle coordinate di un'altro spazio colore vanno oltre gli interessi di questo lavoro. Una panoramica su questi temi può essere trovata in [FvFH90, §13] o [GW92].

Rappresentazione del colore nei modelli tridimensionali

Due sono le tecniche utilizzate per assegnare il colore alla superficie:

⁴Ciò corrisponde ad aggiungere al colore da replicare il primario con coefficiente negativo, e riprodurre il colore così ottenuto utilizzando gli altri due primari.

⁵In uno spazio colore percettivamente uniforme, variando due colori della stessa quantità, le variazioni devono essere percepite come uguali. Ciò non accade in generale nello spazio CIE XYZ.

- per vertex color;
- texture mapping.

La prima tecnica assegna ad ogni vertice della superficie un dato colore [LPC⁺00] [Rig00]. La seconda tecnica consiste nell'individuare una funzione iniettiva (*texture map*) fra lo spazio dei punti della superficie e lo spazio dei pixel di un'immagine. Questa tecnica (il *mapping*) ha una applicazione più generale: con essa si possono modificare le caratteristiche della superficie in modo economico (ad esempio, *bump-mapping*, *transparency-mapping*, *displacement-mapping*). Il bump-mapping perturba le normali della superficie, ma senza modificarne la geometria. Il displacement mapping, invece, ne modifica la geometria. Sebbene in teoria il displacement possa essere in ogni direzione (cioè, map tridimensionale), in generale viene usato solo nella direzione delle normali (depth-map) [LMH00].

È interessante notare che le rappresentazioni *per vertex* corrispondono ad una funzione espressa in forma implicita, mentre quelle legate al *mapping* corrispondono ad una funzione esplicita delle caratteristiche. L'applicazione di una tecnica del secondo tipo è subordinata alla possibilità di definire una parametrizzazione della superficie da utilizzare come dominio di definizione della funzione di mappatura (vedi [LSS⁺98]).

Una rappresentazione *per vertex* è utile per rappresentare il colore di sfondo dell'oggetto o di parte piuttosto estese di esso. Alcune superfici hanno dettagli che per essere rappresentati richiederebbero un elevato numero di poligoni (si pensi, per esempio, ad una lastra di marmo con venature metalliche). Una modellazione esplicita per mezzo di poligoni è spesso impraticabile. Il texture mapping è utile per codificare questo tipo di informazione. L'utilizzo del texture mapping richiede tuttavia una maggiore potenza di calcolo. Recentemente sono state introdotte schede grafiche a livello consumer che supportano il texture mapping hardware. Tuttavia, per sistemi che non dispongono di tali componenti la rappresentazione *per vertex* (anche a costo di un maggior numero di poligoni) è una soluzione più efficiente [RL00].

Il texture mapping è in generale effettuato utilizzando un'immagine. Se il colore del modello tridimensionale non è acquisito, ma è modellato, in alcune situazioni il texture mapping da un'immagine può essere poco realistico. È il caso degli oggetti scolpiti (per esempio in legno o in pietra). L'applicazione di un'immagine sulla loro superficie fornisce la sensazione di un oggetto dipinto. Ciò è causato dal fatto che il colore di un punto della superficie di un oggetto del genere è dipendente dalla posizione 3D del punto stesso in relazione con i punti vicini. In questi casi è meglio definire una texture *procedurale* o *volumetrica*

2. Metodi e tecniche per la digitalizzazione

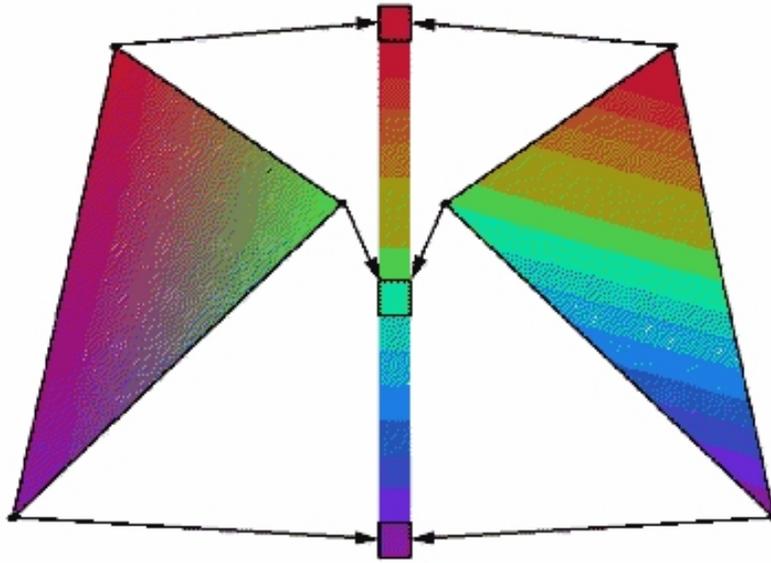


Figura 2.12.: *Per vertex color* (a sinistra) e *texture mapping* (a destra)
[TH].

[EMP⁺94]. Il meccanismo di mappatura del colore in questo caso è simile al caso dell'immagine, ma il colore viene calcolato in funzione della posizione nello spazio tridimensionale invece che nella posizione in uno spazio parametrico bidimensionale.

Le operazioni di ottimizzazione riguardano anche le caratteristiche cromatiche. Per esempio, in [Hop96] il colore (ed in generale gli *appearance attribute*) viene considerato nella procedura di compressione lossy. In questo ambito, una rappresentazione mediante un modello percettivamente uniforme è importante. Lo scopo è infatti l'eliminazione dei dettagli meno (percettivamente) significativi: la codifica delle caratteristiche deve rispecchiare gli effetti percettivi.

2.3. Acquisizione

Un sistema di acquisizione è composto da un insieme di dispositivi e di procedure in grado di catturare e codificare la forma e l'apparenza di un oggetto. Generalmente, il sistema di acquisizione campiona le caratteristiche della superficie dell'oggetto acquisito. Le informazioni collezionate durante il campionamento vengono poi utilizzate nella fase di ricostruzione per ottenere un modello dell'oggetto. I vari sistemi di acquisizione possono essere classificati in base alle loro prestazioni di:

Accuratezza L'accuratezza è la capacità del sistema di quantificare le caratteristiche geometriche ed il colore dell'oggetto con fedeltà. Diverse tecniche possono essere utilizzate per quantificare l'accuratezza. Esse si possono classificare in misure dirette (per esempio, i punti di un reticolo di passo conosciuto) e indirette (distanza media dei punti campionati su di un piano dal piano ottimale).

Risoluzione La risoluzione misura la densità di dettagli che il sistema riesce a riprodurre, o, analogamente, la distanza minima (in una metrica opportuna) che due caratteristiche devono assumere per essere distinte.

Rapidità La rapidità di un sistema di acquisizione viene valutata come il tempo impiegato per la misura di una caratteristica (per esempio, il modulo di acquisizione di Autoscan (descritto più approfonditamente in 5) può arrivare a 100 punti/secondo). È, ovviamente, da rapportare al tipo di caratteristiche acquisite: una linea contiene più informazione di un punto.

Flessibilità La flessibilità di un sistema esprime la capacità del sistema stesso ad acquisire una vasta gamma di forme e di dimensioni.

2. Metodi e tecniche per la digitalizzazione

Questa caratteristica è legata alla libertà lasciata all'utente nella disposizione dei sensori ed all'intervallo di dimensioni che il campo di acquisizione può assumere.

Invasività L'invasività esprime gli effetti che l'operazione di acquisizione può causare delle modifiche (più o meno irreversibili) all'oggetto misurato ed alla sua interferenza con la misura stessa. È chiaramente legata alla natura dell'oggetto da acquisire. In altri termini, l'invasività del sistema limita la gamma di oggetti che possono essere digitalizzati.

Robustezza La robustezza di un sistema descrive la sensibilità del sistema all'influenza delle condizioni ambientali.

Maneggevolezza La maneggevolezza di un sistema esprime le nozioni tecniche necessarie per il suo corretto utilizzo.

Costo Sistemi *consumer* necessitano di soluzioni a basso costo. Molti sistemi preferiscono appesantire la parte software per compensare l'utilizzo di hardware a basso costo.

Alcune di queste caratteristiche possono essere misurate, altre sono solo qualitative. Alcune dipendono direttamente da alcune componenti del sistema. Per esempio, l'invasività dipende fortemente dal principio fisico sfruttato per la misura.

Alcune caratteristiche possono essere in contrasto tra loro. Per esempio, la facilità di uso del sistema mal si accorda alla flessibilità o alla accuratezza. Queste due ultime qualità, infatti richiedono generalmente una maggiore complessità del sistema.

2.3.1. Dispositivi di acquisizione

Per la misura della geometria di un oggetto sono diversi i principi fisici sfruttabili. Gli scanner presenti in letteratura possono essere divisi nelle seguenti categorie:

Ultrasuoni : Gli scanner ad ultrasuoni sono basati sull'emissione di un'onda e sull'analisi delle caratteristiche dell'onda riflessa. Gli ultrasuoni possono attraversare alcuni materiali, in particolare i tessuti biologici. Il principio di funzionamento si basa sulla proprietà che hanno gli ultrasuoni di essere riflessi dalla superficie di confine fra due mezzi a impedenza acustica molto diversa, tipicamente un solido, o un liquido, e un gas. In queste condizioni si può misurare il tempo che impiega un impulso ultrasonoro, emesso

da una piccola sonda trasmittente-ricevente, per attraversare un certo spessore di materiale, essere riflesso dalla superficie opposta, o da un difetto, e tornare al punto di partenza; se la velocità di propagazione è nota, è possibile calcolare lo spazio percorso.

La velocità dipende dalla natura del materiale che l'onda deve attraversare. Una misura accurata richiede quindi la conoscenza della composizione dell'oggetto che si vuole analizzare. Lo strumento va infatti tarato con campioni di spessore noto e di composizione simile al materiale in esame.

Per questo motivo e per la loro bassa invasività, gli scanner ad ultrasuoni vengono utilizzati per indagini (qualitative) di oggetti complessi (ad esempio, ecografia) o per analisi accurate di oggetti di composizione nota (per esempio, individuazione di imperfezioni in oggetti metallici).

Meccanici Gli scanner meccanici (di cui un esempio, Faro [FAR], è riportato in figura 2.16) richiedono il contatto con l'oggetto della misura. In [TOdSg⁺99] viene riportato che dispositivi di contatto per la misura delle coordinate possono avere una precisione di 3 μm localmente e di 14 μm sul campo di acquisizione. A fronte di una elevata precisione, tuttavia, deve essere considerata una certa delicatezza del dispositivo e una bassa velocità di acquisizione. Il loro impiego è quindi limitato alle applicazioni per le quali sia richiesta una precisione elevata per un insieme ristretto di misure. I dispositivi basati su altri principi, per esempio, hanno una precisione inferiore alle capacità di alcune macchine utensili, e quindi non possono essere utilizzati per controllare la qualità dei pezzi lavorati.

I dispositivi di misura degli scanner meccanici possono essere utilizzati per facilitare la calibrazione di sistemi basati su altri principi. Per esempio, per la misura della posizione delle videocamere di un sistema ottico. In tal modo, è possibile registrare differenti acquisizioni rispetto ad un unico sistema di riferimento.

Ottici Gli scanner ottici si dividono in due categorie principali: passivi e attivi. Entrambi sono basati sulla analisi delle caratteristiche di una o più immagini digitali ritraenti la superficie da ricostruire. Gli scanner attivi sono dotati di un dispositivo di illuminazione per proiettare sulla scena un pattern luminoso conosciuto (per esempio, Digibotics [DIGb], riportato in figura 2.15). I pattern generalmente usati sono: punto, linea, multi-punto, multi-linea e griglia. Il sistema di acquisizione elabora l'immagine e stima la posizione

2. Metodi e tecniche per la digitalizzazione

dei punti appartenenti al pattern luminoso. Tipicamente, il pattern viene fatto scorrere sulla superficie, in modo da acquisire più punti possibile. La scansione può essere automatica o manuale. Il primo tipo di scansione assicura una copertura uniforme della superficie, mentre la scansione manuale permette di acquisire solo nelle regioni di interesse. Gli scanner passivi sono meno invasivi in quanto fanno uso solo della luce ambientale.

Gli scanner passivi sono generalmente basati sul calcolo della disparità di due o più immagini prese da diversi punti di vista. Sulle diverse immagini vengono identificati i punti omologhi. Poiché il matching a livello di pixel richiederebbe un pesante sforzo computazionale, molti algoritmi di matching sono basati sull'estrazione di primitive ad un livello di astrazione più elevato (come spigoli e vertici) [VZ98]. L'accuratezza della ricostruzione è quindi legata sia al tipo di caratteristiche che vengono usate per abbinare i punti, sia alla robustezza della procedura di estrazione di tali caratteristiche. Questa operazione è resa difficile dalla presenza di ombre e texture degli oggetti, ma è tuttavia meno sensibile al rumore di quanto lo siano le procedure di matching basate su regioni di pixel [CM92b]. Tipicamente vengono estratti gli spigoli e i vertici.

I dispositivi attualmente disponibili in commercio sono principalmente di tipo attivo [War00]. Essi consentono una maggiore precisione della misura [PTR⁺98]. Le caratteristiche di riflessione della superficie possono rendere difficoltosa l'operazione di acquisizione. Questo problema è più sentito negli scanner passivi, ma superfici particolarmente riflettenti (o particolarmente "assorbenti") possono creare problemi anche agli scanner attivi. Tale fenomeno viene chiamato *scattering* è causato dal materiale di cui è composto il soggetto dell'acquisizione (figura 2.14). Ciò può causare un errore sistematico nella misura [LPC⁺00].

Radar Il Radar Laser funziona analogamente ad un radar a microonde, rilevando la distanza di un oggetto misurando il tempo necessario per un segnale per essere riflesso dall'oggetto stesso. La misura del tempo è effettuata in due modi: misura diretta del tempo di riflessione e misura indiretta per esempio tramite spostamento di fase.

Nella misura diretta del tempo di riflessione (*time of flight* — TOF), data la velocità estremamente elevata dell'impulso luminoso, il componente critico del sistema è quello dedicato alla misura del tempo: un errore di 1 ps comporta un errore di 1 cm nella sti-

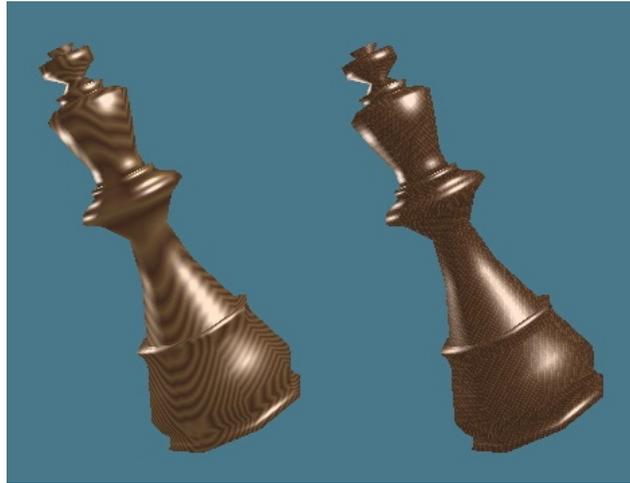


Figura 2.13.: Texture procedurale

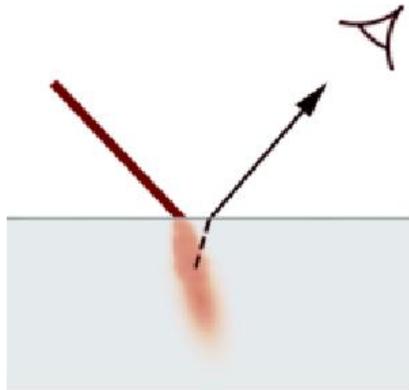


Figura 2.14.: Il fenomeno dello *scattering* provoca un errore sistematico nella misura [LPC⁺00].

2. Metodi e tecniche per la digitalizzazione

ma della distanza. Tuttavia per oggetti di grandi dimensioni (e a grandi distanze), questo può essere un errore accettabile. Inoltre, problemi di rumore nel dispositivo ricevente possono causare errori di misura consistenti. In questo caso, misure multiple possono migliorare l'accuratezza della misura.

La misura del tempo per spostamento di fase (*phase shift*) è basata sulla differenza di fase tra l'onda trasmessa e l'onda ricevuta.

Interferometria L'interferometria è una tecnica di misura basata sullo studio delle figure di interferenza create dalla sovrapposizione di due semifasci di luce provenienti dalla stessa sorgente. I due semifasci, originariamente in fase tra di loro, seguendo due percorsi differenti si vengono a trovare sfasati. La loro sovrapposizione provoca un rafforzamento o un'attenuazione dell'intensità della luce in alcune regioni. Conoscendo la lunghezza di uno dei due percorsi, e la figura di interferenza così composta, si può ricavare la lunghezza dell'altro percorso.

Nella moderna ingegneria, soprattutto meccanica, è necessario lavorare pezzi di grosse dimensioni con precisioni elevate; per esempio i componenti di un aereo hanno dimensioni superiori al metro e vengono lavorati con precisioni dell'ordine dei $10\ \mu\text{m}$. Accuratezze di questo tipo possono essere raggiunte effettuando le misure con metodi di interferometria laser, in virtù della coerenza della radiazione, raggiungendo precisioni dell'ordine di $\lambda/2$, dove λ è la lunghezza d'onda del raggio luminoso.

Scanner volumetrici Dati tridimensionali possono essere ottenuti anche utilizzando strumenti di tipo diverso da quelli finora descritti, per esempio un tomografo. In generale, però, si tratta di strumenti dedicati ad uno scopo ben preciso, e quindi con una scarsa flessibilità d'uso. Inoltre, richiedono per il loro utilizzo operatori specializzati. I dati acquisiti non sono relativi ad una singola superficie, ma sono volumetrici. Lo spazio acquisito è cioè uniformemente campionato nelle tre dimensioni. Il valore contenuto in ciascun campione non è la posizione tridimensionale (la quale è ricavabile dalla posizione del singolo campione nell'insieme di acquisizione), ma è una descrizione del materiale contenuto nel volume corrispondente. La superficie è ricavabile come frontiera tra regioni di diverso materiale. La ricostruzione deve quindi essere preceduta da un'operazione di segmentazione dell'insieme acquisito.

2.3.2. Tecniche di misura

Modalità di triangolazione

La tecnica più utilizzata per ottenere una misura di distanza è quella detta *triangolazione*. Essa si basa essenzialmente sulla proprietà dei triangoli di essere completamente specificati, cioè di avere definite le lunghezze dei tre lati e le ampiezze dei tre angoli che li compongono, dati tre di questi elementi. Sebbene nei casi reali la formalizzazione sia più complessa (per esempio, rette che possono essere sghembe, distorsioni dovute alle lenti dei dispositivi ottici, errori di quantizzazione), la spiegazione limitata al caso bidimensionale è più chiara ed essenziale.

I sistemi di acquisizione sono quindi progettati in modo formare un triangolo tra i dispositivi del sistema (sensori e puntatore, per esempio) ed il punto di cui misurare la posizione. Di questo triangolo almeno tre elementi (fra angoli e lati) devono essere conosciuti.

Una casistica estesa delle soluzioni proposte in questo ambito presenti in letteratura e sul mercato va oltre gli scopi di questa tesi. Verranno quindi analizzati solo tre approcci, a fini esemplificativi.

Digibotics [DIGb] produce tre modelli di scanner a differenti dimensioni. Essi hanno un laser e due sensori vincolati ad un piano. Inoltre, come illustrato in figura 2.15, i due sensori sono orientati in una direzione fissa di trenta gradi rispetto al laser. L'oggetto da acquisire viene posto di fronte al laser, il quale proietta un punto sull'oggetto stesso. I due sensori scorrono quindi su una guida orizzontale fino a trovare la macchia luminosa proiettata dal laser. Il triangolo per la misura ha quindi per vertici la posizione del laser, la posizione del sensore e la posizione della macchia luminosa, cioè del punto della superficie. Conoscendo la posizione del sensore rispetto al laser, l'orientamento del sensore (fisso) e la direzione del laser (fissa, perpendicolare al piano di scorrimento) è possibile calcolare la distanza laser-punto. Inoltre, gli scanner Digibotics sono composti da una piattaforma rotante e da guide laterali di scorrimento del blocco laser-sensori che permettono una maggiore flessibilità di acquisizione.

Il dispositivo di acquisizione di Autoscan [BFB⁺98] [BF00] è costituito da una coppia di TV-camere, da un comune puntatore laser da conferenza. Il sistema è estremamente flessibile perché le TV-camere possono essere posizionate ed orientate con la massima libertà ed il puntatore laser può essere, ovviamente, manovrato senza vincoli. In questo caso, la posizione e la direzione del laser non possono quindi essere conosciute. Ciò che è conosciuto è la posizione e la direzione dei sensori, le quali vengono rilevate durante la fase di calibrazione del sistema. In

2. *Metodi e tecniche per la digitalizzazione*

Autoscan, quindi la triangolazione viene basata sulla conoscenza della lunghezza di un lato (la distanza tra i sensori) ed i due angoli ad esso adiacenti (le direzioni lungo le quali il punto luminoso viene visto dalle TV-camere). Il sistema Autoscan rappresenta un esempio di applicazione della metodologia studiata in questa tesi. Esso sarà trattato più approfonditamente nel capitolo 5.

Il principio della triangolazione viene applicato iterativamente in Faro [FAR]. Faro è uno scanner meccanico costituito da un braccio a tre snodi (figura 2.16). Il braccio snodabile forma una successione di triangoli, dei quali la lunghezza di un lato (un segmento del braccio) è conosciuto (poiché è fisso) e gli angoli ad esso adiacenti vengono misurati dai dispositivi collegati agli snodi.

Acquisizione del colore

Se le tecniche per misurare la geometria di un oggetto sono varie, la misura delle caratteristiche cromatiche della superficie può avvenire solo per mezzo di una videocamera a colori.

Poiché il colore apparente di un oggetto dipende da vari fattori, le differenze maggiori tra le tecniche di acquisizione del colore differiscono tra loro per i metodi impiegati per epurare le misure dalle componenti causate dalle condizioni in cui la specifica misura è stata operata.

Per il Progetto Michelangelo [LPC⁺00] viene utilizzata una tecnica di scansione che acquisisce per ogni punto sia la geometria che il colore. Ogni acquisizione relativa al colore è composta di due immagini: una con la sola luce ambientale e una anche con la luce artificiale. Per determinare il colore reale della superficie viene utilizzata l'immagine data dalla differenza delle due immagini. In questo modo, le condizioni di illuminazione sono conosciute (solo illuminazione artificiale).

In [RB99] viene utilizzato un dispositivo dotato di cinque fonti di illuminazione. Ogni acquisizione, è quindi composta dalle cinque immagini rappresentati l'oggetto illuminato da una delle cinque luci. Per la stima del colore è necessario anche un modello (sia pure a bassa risoluzione) della superficie.

2.4. Ricostruzione

Nella fase di ricostruzione vien cercata la superficie che meglio approssima la collezione di dati fornita dalla fase di acquisizione. Due sono i fattori principali che condizionano questo problema: la conoscenza a priori sull'oggetto acquisito e il tipo di informazioni disponibili dalla acquisizione.



Figura 2.15.: Un modello di scanner Digibotics (sinistra)e la tecnica utilizzata per la misura (destra) [DIGb].

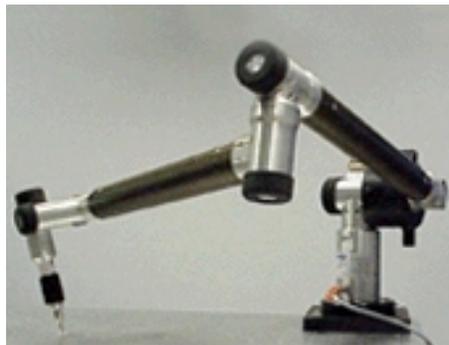


Figura 2.16.: Lo scanner Faro [FAR].

2. Metodi e tecniche per la digitalizzazione

Per risolvere il problema è necessario un paradigma di rappresentazione della superficie. Esso può essere usato per codificare la conoscenza a priori sul problema. Maggiore è la conoscenza a priori sull'oggetto acquisito di cui si dispone, più dettagliato, e per questo meno flessibile, è possibile scegliere il formalismo di rappresentazione. Se, per esempio, si è interessati all'acquisizione di oggetti derivanti da una classe ben precisa, è possibile esprimere una forma parametrica per la rappresentazione della classe ed eseguire la ricostruzione come ricerca dei valori dei parametri che meglio fanno adattare la forma ai dati disponibili. Tecniche di ottimizzazione *ad hoc* possono essere studiate per questi casi.

Se, d'altra parte, la conoscenza a priori è limitata, il paradigma di ricostruzione dovrà avere molti parametri per essere in grado di assumere molte forme e la tecnica di ricostruzione dovrà essere più complessa.

Il problema della ricostruzione può essere riformulato, quindi, come un problema di ottimizzazione. In generale, un approccio *model based* è più robusto [DTM96]. Tuttavia, in questa tesi l'interesse è rivolto alla metodologia e quindi si privilegerà la trattazione del caso più generale dove la conoscenza a priori è più limitata.

Un analogo discorso può essere fatto per i dati. Sebbene vi siano sistemi in grado di acquisire linee, spigoli o altre caratteristiche ad un livello di astrazione più elevato e dal maggiore contenuto informativo, la ricostruzione da punti è un problema più generale. Ciò giustifica l'attenzione che verrà prestata al problema della ricostruzione a partire da punti affetti da rumore.

Vale la pena di notare, però, che alcune informazioni ricavabili da alcuni dispositivi di acquisizione possono essere sfruttate per la ricostruzione. Per esempio, la direzione della normale alla superficie o la direzione nella quale il sensore era rivolto durante l'effettuazione della misura.

I problemi ai quali la fase di ricostruzione deve far fronte sono legati all'incertezza presente sui dati (rumore, dati mancanti o ambigui) e la mancanza di informazioni sulla topologia della superficie.

2.4.1. Ricostruzione da punti

La ricostruzione di una superficie a partire da una nuvola di punti tridimensionali comporta implicitamente la definizione della topologia della superficie stessa.

Per capire l'importanza della differenza tra geometria e topologia di un insieme di dati, si consideri come esempio l'acquisizione della superficie di una mano. Punti collocati sulle punte di due dita della stessa mano potranno risultare geometricamente vicini, ma sono topologica-

mente distanti. La distanza tra i due punti misurata come lunghezza del segmento di linea più corto aventi come estremi i punti stessi (cioè, distanza euclidea) sarà molto diversa dalla distanza misurata limitando l'insieme dei segmenti a quelli appartenenti alla superficie (cioè, distanza topologica). Le caratteristiche legate alla topologia non cambiano se la superficie viene sottoposta ad alcune trasformazioni.

La distanza topologica tende a coincidere con la distanza geometrica quanto più la regione di superficie contenente i due punti è approssimabile con un piano. Via via che ci si avvicina al palmo, i punti appartenenti alle due dita tendono ad avvicinarsi sia topologicamente che geometricamente. A causa del rumore, alla base delle dita i punti potrebbero confondersi. La conoscenza della topologia rende più semplice la ricostruzione. Per questo motivo, la conoscenza a priori sugli oggetti da ricostruire è una preziosa risorsa, qualora si possibile codificarla e includerla nella fase di ricostruzione.

Tuttavia, se le informazioni sulle quali la ricostruzione si basa sono limitate alla posizione dei punti acquisiti, la topologia deve essere stimata. Poiché la distanza geometrica e tra i punti dati è l'unica informazione a disposizione, l'ipotesi utilizzata è che punti geometricamente vicini lo siano anche topologicamente. Il metodo più usato per sopperire alla mancanza di informazioni sulla topologia consiste nel costruire un modello poligonale che rappresenti i dati a bassa risoluzione e utilizzarlo come riferimento topologico. Si cerca quindi di ottenere una approssimazione localmente planare per approssimare la topologia della superficie dell'oggetto. A partire da tale modello, si ricostruisce poi la superficie in dettaglio.

Verranno ora passate in rassegna le tecniche più utilizzate per la ricostruzione di superfici a partire da punti campionati. La scelta di non considerare altre informazioni (quali, ad esempio, la direzione dalla quale l'acquisizione è effettuata) è motivata dalla volontà di valutare le caratteristiche delle varie tecniche indipendentemente dal dispositivo di acquisizione utilizzato per collezionare i punti.

Il problema della ricostruzione di una superficie a partire da punti campionati su di essa è stato di grande interesse nel recente passato e lo è tutt'ora. Le tecniche per risolverlo possono suddividersi in due grandi categorie:

- *tecniche volumetriche*, che trattano i dati senza avere alcuna informazione riguardo alla topologia e che quindi cercano di individuare il volume occupato dall'oggetto;
- *approssimazione di funzione*, che cercano di individuare la superficie come funzione (espressa in forma implicita o esplicita) definita

2. Metodi e tecniche per la digitalizzazione

su un dominio dato.

È possibile utilizzare una combinazione di tecniche derivanti dalle due categorie per individuare prima un dominio avente le stesse caratteristiche topologiche della superficie dell'oggetto acquisito e utilizzare quindi tale dominio per costruire una approssimazione della superficie. Questo è un approccio frequentemente seguito nei lavori in letteratura.

Un'altra differenza importante tra i procedimenti di ricostruzione sono le caratteristiche della soluzione che si ottiene: interpolazione o approssimazione dei punti dati. Poiché, generalmente, i punti dati sono il risultato di una misura, una interpolazione non costituisce una ricostruzione fedele della superficie dell'oggetto. Tuttavia, una interpolazione può essere utilizzata come una prima approssimazione della superficie da sottoporre ad un successivo raffinamento.

Una classificazione più dettagliata può essere trovata in [MM98b].

Vi sono alcune definizioni e metodi che appartengono al vocabolario comunemente utilizzato nel settore della ricostruzione di superfici da punti: convex hull, triangolazione di Delaunay, diagramma di Voronoi, α -shapes, octree, marching cubes. Ad esse si farà riferimento nel seguito, ma, per non appesantire la trattazione, le loro definizioni sono state messe nel glossario (6.2).

Metodi volumetrici

Tipicamente, per ottenere la topologia dell'oggetto vengono utilizzati metodi volumetrici. Gli algoritmi di questa classe sono composti dai seguenti passi [MM98b]:

1. Scomposizione dello spazio occupato dai punti in celle.
2. Ricerca delle celle attraversate dalla superficie.
3. Calcolo della superficie a partire dalle celle selezionate.

In [AS96] il *bounding box* dei punti viene suddiviso in una griglia regolare di *voxel*. Una prima approssimazione della superficie viene ottenuta dalle facce esterne delle celle occupate da almeno un punto, le quali vengono poi suddivise diagonalmente per ottenere una mesh triangolare. Ogni vertice di tale mesh viene riposizionato tramite un'operazione di filtraggio. La posizione viene calcolata come media pesata della posizione originale e della posizione dei suoi vertici adiacenti.

In [HDD⁺92] viene utilizzato il concetto di funzione *distanza con segno* per approssimare la topologia della superficie. La funzione distanza con segno dalla superficie assume un segno positivo nei punti esterni alla

superficie e negativo all'interno della stessa. L'isosuperficie di distanza nulla è chiaramente la superficie da ricostruire. Nel lavoro citato, lo spazio occupato dai punti viene suddiviso in modo regolare ed ad ogni vertice viene assegnato il valore della distanza con segno dalla superficie. Vengono poi selezionati i soli cubi aventi vertici con segni discordi. Ciò equivale a scegliere i cubi attraversati dalla superficie. Infine, l'algoritmo *marching cubes* viene utilizzato per raffinare la superficie. Un approccio simile è stato utilizzato in [RW97]. La differenza principale è racchiusa nel metodo utilizzato per la stima della distanza con segno, il quale è un elemento critico del procedimento. La funzione distanza con segno, infatti, non è conosciuta a priori, altrimenti si conoscerebbe anche la superficie. Essa deve essere stimata a partire dai dati.

Nei metodi volumetrici con griglia regolare, la dimensione delle celle è una scelta critica. Alternativamente, la griglia può essere costruita adattivamente. In [BBX95] vengono utilizzate celle di forma tetragonale. All'inizio, vi è un'unica cella che include tutti i punti, la quale viene suddivisa in quattro aggiungendo all'insieme dei vertici il suo baricentro. Viene quindi calcolata la distanza con segno di tutti i vertici. I tetraedri con vertici di segno concorde vengono scartati. Per ognuno dei tetraedri rimasti viene calcolata una approssimazione della superficie che lo attraversa (una polinomiale di Bernstein-Bézier). Se l'errore di approssimazione di tale superficie rispetto ai punti contenuti nel tetraedro è maggiore di una soglia prefissata, la cella viene suddivisa e la procedura viene iterata. Per evitare la formazione di tetraedri troppo allungati, ad ogni suddivisione viene ricalcolata la tetraedrizazione secondo la condizione di Delaunay. In [PYL00] l'insieme di punti viene partizionato mediante un algoritmo di clustering (*k-means*). Ciascun cluster viene rappresentato mediante una approssimazione poligonale ed il modello risultante viene utilizzato come prima approssimazione della superficie da ricostruire. Per evitare di porre nello stesso cluster punti provenienti da facce opposte di oggetti sottili, nel clustering vengono utilizzate le normali ai punti. Esse sono dedotte dalle fasi di acquisizione.

Alcuni metodi lavorano per eliminazione di elementi. Tale procedimento viene anche chiamato *sculpturing* o *carving*. In [EM94] viene proposto un algoritmo di ricostruzione basato sugli α -shape. Inizialmente viene ricostruita la triangolazione di Delaunay e successivamente vengono eliminati gli elementi che non sono circoscrivibili da una sfera di raggio α : gli elementi rimasti costituiscono quindi l' α -shape dell'insieme di punti dati. Per determinare le facce che effettivamente compongono la superficie viene adottato il seguente criterio. Per ogni triangolo dell' α -shape vengono costruite le due sfere di raggio α che passano per i suoi vertici: il triangolo appartiene alla superficie se almeno una del-

2. Metodi e tecniche per la digitalizzazione

le due sfere non contiene alcun punto dell'insieme dato. A fronte della semplicità di utilizzo di un algoritmo del genere — un unico parametro globale di semplice interpretazione, il raggio α — va notato che la possibile variabilità della densità di campionamento può causare, se α è troppo elevato, una ricostruzione troppo smooth con errori nella ricostruzione della topologia o, se α è troppo piccolo, una ricostruzione con buchi non presenti nella superficie reale.

L'estrazione di superficie da α -shape è trattato anche da [GMW97]. In questo lavoro vengono affrontati i problemi della decisione della direzione della normale delle facce singole — triangoli che non appartengono ad alcun tetraedro — e del rilevamento automatico del bordo della superficie — rendendo così possibile la ricostruzione di superfici aperte. La procedura studiata è basata sulla costruzione di un grafo di adiacenza delle facce, dove ogni arco è pesato in modo proporzionale all'angolo formato dalle normali delle facce corrispondenti. Lo *spanning tree* minimo del grafo fornisce un ordinamento per l'orientamento delle normali.

Recentemente è stato proposto un algoritmo che ricostruisce una superficie con caratteristiche strettamente legate all' α -shape: il *Ball Pivoting Algorithm* [BMR⁺99]. Data una sfera di raggio ρ , partendo da un triangolo iniziale, la sfera viene fatta ruotare imperniata ad ogni lato esterno della mesh. Quando la sfera interseca un punto, un nuovo triangolo viene formato utilizzando gli estremi del lato perno e il nuovo punto stesso. Il nuovo triangolo viene aggiunto alla mesh e la procedura può continuare. Si dimostra che la mesh ottenuta è composta da un sottoinsieme delle facce appartenenti al ρ -shape dell'insieme di punti dato. L'utilizzo di sfere di raggio via via più grande permette la ricostruzione di superfici con differente densità di campionamento. La località delle operazioni richieste consente il trattamento di insiemi di punti di grandi dimensioni. In figura 2.19, viene mostrata l'applicazione di questo algoritmo su dati bidimensionali.

In [Wel96] viene espresso il concetto di stabilità di uno spigolo della triangolazione di Delaunay di un insieme di punti. Uno spigolo viene ritenuto stabile se effettuando la triangolazione dopo aver perturbato la posizione dei punti dell'insieme considerato lo spigolo permane. Spigoli topologicamente corretti tendono ad essere stabili.

Sempre sul tema della rimozione, [Boi84] propone un metodo che partendo dall'insieme di tetraedri ottenuti dalla triangolazione di Delaunay di un insieme di punti dati, elimina quelli che sono ritenuti esterni all'oggetto. I tetraedri selezionati per l'eliminazione sono scelti tramite un'euristica (quelli che presentano verso l'esterno due facce, cinque

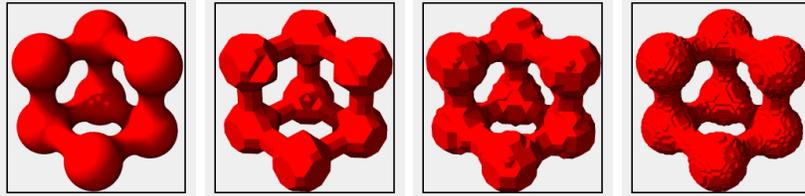


Figura 2.17.: Ricostruzione di una superficie effettuata con *marching cubes* a differenti risoluzioni [Sha]. La superficie originale è quella a sinistra.



Figura 2.18.: α -shape di un insieme di punti al variare del parametro α .



Figura 2.19.: Comportamento di *Ball pivoting* su dati con differenti caratteristiche [BMR⁺99].

2. Metodi e tecniche per la digitalizzazione

spigoli e quattro punti oppure una faccia, tre spigoli e tre punti)⁶. I tetraedri sono caratterizzati da un valore, il *decision value*, che ne individua la propensione all'eliminazione. Il *decision value* viene definito come la massima distanza di un punto sulla sfera che circoscrive un tetraedro dalle facce del tetraedro stesso. Tetraedri bassi e larghi, che generalmente nascondono dettagli della superficie, vengono quindi tendenzialmente eliminati. L'algoritmo termina se tutti i punti giacciono sulla superficie o se la rimozione del tetraedro con *decision value* più elevato non aumenta la somma dei *decision value* dei tetraedri rimanenti nella lista.

Questo algoritmo non consente la ricostruzione di superfici di *genus* (vedi glossario 6.2) maggiore di zero. In [IBS98] viene proposta una variante dell'algoritmo precedente in grado di ricostruire superfici di *genus* arbitrario. Ciò è effettuato aggiungendo alla lista dei tetraedri rimuovibili anche quelli con una faccia e quattro vertici o tutte e quattro le facce. Viene inoltre definita una *elimination function* — il massimo *decision value* dei tetraedri nella lista di eliminazione. Tale funzione viene utilizzata per decidere la terminazione dell'algoritmo. Viene notato, infatti, che l'andamento della *elimination function* presenta alcuni gradini. L'ipotesi è che uno di questi gradini identifichi la superficie reale.

Un altro miglioramento all'approccio in [Boi84] viene apportato in [Vel95]. In alternativa al *decision value* viene introdotto il γ -indicator, un numero reale associato ad ogni tetraedro con almeno una faccia all'esterno di una triangolazione di un insieme di punti. Il valore assoluto viene definito come $1 - \frac{r}{R}$, dove r è il raggio della sfera che circoscrive il triangolo esterno e R è la sfera che circoscrive il tetraedro. Il segno viene posto positivo se il centro della sfera circoscrivente è all'esterno del tetraedro e negativo se interno. Il vantaggio dell'utilizzo del γ -indicator per la scelta del tetraedro da rimuovere è la sua insensibilità alla dimensione del tetraedro stesso. Si adatta quindi alle variazioni di frequenza di campionamento nelle diverse regioni dell'oggetto.

Un approccio alternativo, basato sulla dualità tra la triangolazione di Delaunay e il diagramma di Voronoi, è proposto in [SB97]. Viene costruito lo *spanning tree* minimo del diagramma di Voronoi (dove i nodi corrispondono ai vertici del diagramma e la lunghezza di ogni arco è la lunghezza dello spigolo corrispondente) e vengono applicate alcune euristiche per sfoltirlo. L'eliminazione di un nodo dello *spanning tree* corrisponde all'eliminazione del tetraedro corrispondente nella triangolazione.

⁶L'eliminazione di un tetraedro con all'esterno tre facce, sei spigoli e quattro punti comporterebbe l'eliminazione di un punto dell'insieme dato.

In [BBCS99] viene utilizzato lo stesso principio di “scultura” di una triangolazione iniziale, qui ottenuta mediante α -shape. Il criterio di eliminazione è però differente ed è basato su una misura di smoothness locale. L’ α -shape viene usata per individuare la struttura generale della superficie, mentre la procedura di eliminazione dei tetraedri ha lo scopo di liberare i dettagli.

Come in [HDD⁺92], la superficie può essere espressa tramite una funzione distanza con segno. Tale funzione, come la superficie, è sconosciuta. Una sua approssimazione è però ricavabile ed utilizzabile per ottenere una migliore stima della superficie (per esempio attraverso l’algoritmo *marching cubes*). La funzione distanza stimata deve essere ragionevolmente affidabile solo in un intorno della superficie.

In [HDD⁺92] viene utilizzato il seguente procedimento. Inizialmente viene stimato il piano tangente alla superficie in ogni punto $p_i \in P$. Esso è ottenuto come il miglior piano (nel senso dei minimi quadrati) approssimante i k punti più vicini a p_i , per un certo k . Per ogni piano tangente viene definito un centroide, calcolato come il baricentro dei punti utilizzati per la sua stima. Al fine di assegnare il corretto orientamento delle normali, viene costruito un grafo dove i nodi corrispondono ai centroidi, gli archi congiungono i k centroidi più vicini e la lunghezza degli archi misura la differenza di orientamento dei piani. Il calcolo dello *spanning tree* minimo fornisce un ordinamento per la propagazione delle normali. I piani tangenti e l’orientamento consentono di calcolare la distanza con segno nelle vicinanze della superficie. La distanza con segno dalla superficie in un punto dello spazio viene calcolata come distanza (con segno) dal piano tangente corrispondente al centroide più vicino.

In [RW97] viene proposto un metodo semplice ed efficiente per stimare la funzione distanza con segno, illustrato nella figura 2.20. Lo spazio contenente i punti viene suddiviso in modo regolare in una griglia di voxel. La direzione della normale ad un punto viene calcolata individuando i due punti dati più vicini ad esso fra quelli appartenenti allo stesso voxel. per stabilire il verso della normale viene adottato un criterio di visibilità: se un voxel è visibile da un punto posto a (più o meno) infinito lungo uno dei tre assi, il segno della componente della sua normale corrispondente all’asse considerato è fissato. Il segno per i voxel non direttamente visibili viene definito per propagazione dai voxel vicini. Questa euristica permette solo la ricostruzione di superfici chiuse e senza buchi. Le normali vengono sottoposte ad un filtraggio passabasso per evitare brusche variazioni delle direzioni dovute al rumore di misura. Il valore della funzione distanza con segno viene quindi calcolato in ogni vertice della griglia come media pesata della distanza con segno di ogni punto appartenente ai voxel adiacenti. La distanza con segno

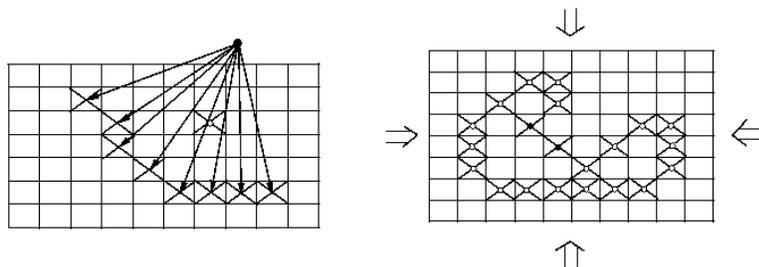


Figura 2.20.: *Carving* [RW97]. Lo spazio occupato dai dati dell'acquisizione viene quantizzato e i voxel attraversati troppo spesso dalla linea di visione vengono eliminati (sinistra). La superficie viene ricostruita a partire dalle facce dei voxel rimasti visibili dall'infinito dalle sei direzioni parallele agli assi.

di un punto ad un punto con normale viene definita come la distanza euclidea da quel punto, con segno positivo se l'angolo tra la normale e la congiungente è superiore a $\frac{\pi}{2}$.

L'approccio suggerito in [BTG95] utilizza una stima della distanza dalla superficie differente dalla distanza con segno: l'asse mediale (*medial axis*). Esso è definito come l'insieme dei centri delle sfere massimali, le quali sono le sfere interne alla superficie che non sono incluse in nessuna altra sfera interna. L'asse mediale vien calcolato da un partizionamento mediante griglia regolare del *bounding box* di un insieme di punti dati, P . I voxel contenenti punti di P sono ritenuti il bordo del volume racchiuso dalla superficie da ricostruire. Partendo dai voxel più esterni, quelli che non contengono alcun punto di P vengono rimossi. A quelli contenenti i punti viene assegnato una valore di distanza pari a zero. La distanza viene quindi propagata verso l'interno del volume incrementandola man mano che si procede. I voxel a distanza massima sono inclusi nell'asse mediale. A ogni voxel dell'asse mediale viene assegnata una sfera con raggio pari alla distanza. Ad ogni sfera è associato una funzione di campo che permette di calcolare un valore scalare di campo in un punto arbitrario dello spazio. La funzione campo delle sfere è additiva. Il campo scalare consente di definire una funzione implicita per l'approssimazione della superficie. Nel lavoro citato viene proposto una procedura incrementale per minimizzare il numero di elementi dell'asse mediale necessari alla ricostruzione della superficie e per ottimizzare il valore dei parametri che definiscono il campo di ogni singola sfera.

I metodi basati sulla triangolazione di Delaunay ipotizzano che i punti dati siano esenti da errore. Solitamente nei casi reali la posizione dei punti è il risultato di una misura e pertanto è affetta da rumore. Una qualche operazione di filtraggio deve quindi essere effettuata dopo la ricostruzione.

Approssimazione di funzioni

In [MBL⁺91] una mesh iniziale già abbastanza simile alla superficie da ricostruire viene adattata ai punti dati. Il modello viene sottoposto a deformazioni condizionate da una serie di vincoli. Ad ogni vertice viene associata una funzione che esprime il costo delle deformazioni locali, tenendo in considerazione il rumore e le caratteristiche della ricostruzione, nonché la semplicità della mesh.

Un approccio differente è seguito in [ST92] dove l'adattamento è ottenuto mediante particelle orientate (*oriented particles*). Ogni particella possiede alcuni parametri che vengono modificati durante l'adattamento. La modellazione delle interazioni fra particelle plasma la forma della superficie mediante forze attrattive e repulsive.

L'utilizzo della metafora fisica è stato utilizzato anche da altri autori. In [AS96] un modello poligonale vicino alla superficie viene trasformato in un modello fisico per sottoporlo ad un processo di adattamento ai dati. I vertici sono interpretati come punti dotati di massa, mentre gli spigoli vengono sostituiti da molle. Inoltre, ad ogni vertice viene assegnata una molla che lo collega al punto dato più vicino. Il modello può essere espresso in termini di un sistema di equazioni differenziali lineari, la soluzione del quale può essere calcolata iterativamente. Una volta calcolata la situazione di equilibrio, il modello fisico viene ritrasformato in un modello geometrico. La soluzione individua quindi un compromesso fra la smoothness della superficie e l'aderenza ai dati. Il grado di compromesso è determinato dai parametri fisici impostati.

La modellazione di un sistema fisico è stato utilizzato anche in [TM91], con l'intento di ottenere un approccio di modellizzazione delle forme che contemporaneamente soddisfacesse i requisiti di ricostruzione e riconoscimento. Il paradigma proposto, chiamato Superquadriche Deformabili (*Deformable Superquadrics*), consiste in una classe di modelli costituita da forme geometriche parametrizzate (ellissoidi superquadrici) usati come dominio per costruire una funzione (descritta mediante spline). Sono quindi consentite sia deformazioni locali che vincoli globali. Inoltre gli oggetti hanno un comportamento strutturale regolato da leggi di dinamiche rigide o non rigide, espresse da un insieme di equazioni lagrangiane del moto. La simulazione numerica delle equazioni del moto

2. Metodi e tecniche per la digitalizzazione

determina l'evoluzione dei gradi di libertà del modello sotto l'azione di forze e vincoli. I vantaggi di una dinamica guidata da forze sono legati al modo intuitivo con cui si possono correlare i gradi di libertà del modello con le caratteristiche dei dati. In risposta alle forze generate dai dati il modello si sposta, si deforma e si orienta in modo conforme ai dati.

Le strutture di adattamento iterativo sono state studiate anche nell'ambito dell'intelligenza artificiale, come metodi di soluzione del problema dell'apprendimento da esempi (generalizzazione).

Le *Self-Organizing Map* (SOM) [Koh95], anche conosciute come *Kohonen features map*, sono costituite da un insieme di unità organizzate in una struttura reticolare, $\{u_j\}$, e caratterizzate da un valore nello spazio delle caratteristiche, w_j , (senza perdere di generalità, può essere \mathbb{R}^D). La SOM viene sottoposta ad una fase di adattamento con lo scopo di approssimare la distribuzione di un insieme dato, $P \subset \mathbb{R}^D$. Ad ogni passo di adattamento, un elemento dell'insieme dato viene estratto, p_i , e confrontato con w_j . L'unità più vicina, u_k , chiamata *winning unit*, e le unità ad essa connesse vengono adattate secondo una regola del tipo:

$$w_j(i+1) = w_j(i) + \eta h_j(k) \cdot (p - w_j(i))$$

dove η è il tasso di apprendimento — parametro che regola la velocità di adattamento — e $h_j(k)$ è una funzione che esprime il grado di vicinanza dell'unità j all'unità k . Adattando oltre che la *winning unit* anche le unità ad essa vicine, al termine dell'adattamento le unità vicine nel reticolo avranno anche vettori delle caratteristiche simili. Ciò significa, utilizzando come dati punti tridimensionali, che la vicinanza topologica rispecchierà la vicinanza geometrica. Il reticolo delle unità costituirà quindi una superficie poligonale che approssimerà in modo smooth i punti dati.

Le SOM, come molti paradigmi di questo tipo, presentano due problemi. Sebbene la configurazione iniziale del vettore delle caratteristiche delle unità possa essere totalmente casuale, una situazione iniziale ragionevolmente vicina alla superficie risparmia risorse di calcolo che possono essere utilizzate per ottenere una soluzione più raffinata. Una pre-elaborazione per ottenere un modello iniziale è quindi indicata. In secondo luogo, alcune unità (chiamate *dead unit*) possono rimanere escluse dall'adattamento. La posizione finale delle *dead unit* sarà quindi vicina a quella di partenza e la superficie risultante poco smooth. Questo può capitare quando queste unità sono troppo lontane dai punti dati per risultare *winning unit* o quando il reticolo è troppo denso e le unità in questione non cadono mai nell'insieme delle unità da adattare. Per far fronte a questa situazione, in [BH94] viene proposta una procedura di

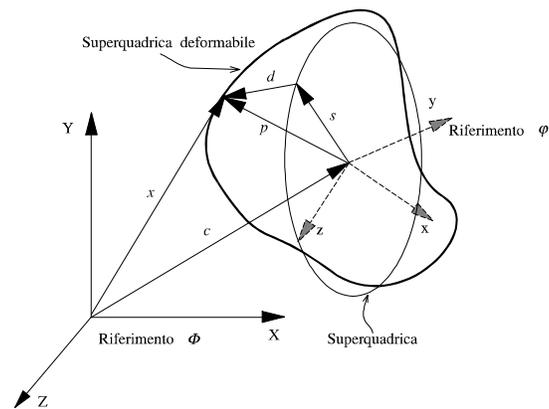


Figura 2.21.: Superquadriche deformabili [TM91].

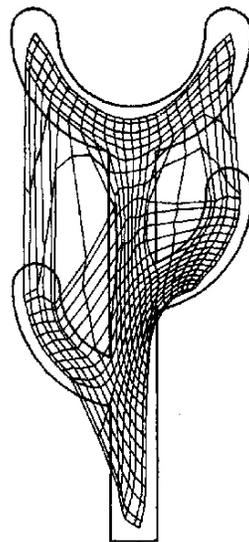


Figura 2.22.: Una mappa di Kohonen approssima una distribuzione complessa [Koh95].

2. Metodi e tecniche per la digitalizzazione

addestramento inverso, dove, per ogni unità, viene individuato l'elemento dell'insieme di punti dati più vicino e utilizzato per adattare l'unità stessa. La combinazione di adattamento diretto ed inverso permette di ottenere una superficie aderente ai punti dati.

Alcuni algoritmi ostruiscono la superficie in modo iterativo. Ad ogni iterazione, la superficie viene estesa per descrivere un insieme più ampio di dati. Questi metodi cercano di sfruttare la correlazione tra punti geometricamente vicini per stimare il comportamento della superficie in una regione circoscritta dello spazio tridimensionale. Essi sono quindi caratterizzati dalla località delle operazioni.

In [Boi84] viene proposto un algoritmo che costruisce la superficie in modo incrementale, aggiungendo un triangolo al bordo della mesh fino ad allora ricostruita. Ad ogni passo viene aggiunto un triangolo, selezionando un lato appartenente al bordo e individuando l'insieme di punti ad esso vicini. Questi punti vengono utilizzati per stimare il piano tangente passante per il lato. Il punto la cui proiezione massimizzi l'angolo sotteso tra la proiezione stessa ed il lato viene scelto per costruire il nuovo triangolo. La procedura termina quando non vi sono più lati disponibili.

In [MM98a] viene introdotto il concetto di *surface description graph* (SGD). Esso è definito come estensione del minimo *spanning tree* dell'insieme dei punti. Ogni nodo del SDG viene utilizzato per estrarre le caratteristiche della superficie nell'intorno attraverso una fase di *features extraction*. Lo scopo è di individuare le regioni con le stesse caratteristiche morfologiche e collegarle, ottenendo così una ricostruzione della superficie.

2.4.2. Multirisoluzione

Molti fenomeni del mondo reale hanno un contenuto informativo che si manifesta a differenti scale spaziali. In particolare, le caratteristiche di molti oggetti possono essere riferite come appartenenti alla loro forma globale o al dettaglio. Le caratteristiche a scala elevata generalmente sono legate al tipo di oggetto, mentre quelle a bassa scala costituiscono i particolari che differenziano gli elementi dello stesso tipo. Inoltre, le caratteristiche ad ampia scala spaziale hanno generalmente un basso contenuto in frequenza (cioè una bassa variabilità), mentre le caratteristiche a bassa scala hanno un contenuto in frequenza più elevato. Questo significa che organizzando le informazioni in base alla scala alla quale si manifestano, poche risorse computazionali sono generalmente sufficienti per descrivere il comportamento a scala elevata, mentre la configura-

zione delle risorse dedicate alle caratteristiche a bassa scala può essere effettuata limitando l'analisi ai dati in una regione circoscritta.

Su questi concetti si basano i paradigmi multirisoluzione e, più in particolare i paradigmi gerarchici. Essi sono infatti strutturati per fornire una descrizione a diverse scale.

Esistono essenzialmente due modi per ottenere una descrizione multirisoluzione, i quali forniscono una chiave di classificazione per i paradigmi multirisoluzione:

- *coarse to fine* (figura 4.11), quando vengono ottenute prima le descrizioni a bassa risoluzione (ampia scala) e via via la risoluzione viene aumentata aggiungendo il dettaglio (informazioni a bassa scala);
- *fine to coarse* (figura 4.11), quando viene prima ricavata la descrizione a massima risoluzione e poi, man mano che il dettaglio viene eliminato, vengono definite le descrizioni a bassa risoluzione.

Un paradigma multirisoluzione *coarse to fine* viene utilizzato, per esempio, in [ZHH98], dove una nuvola di punti viene ricorsivamente ricostruita per mezzo di superquadriche. Inizialmente viene stimata la superquadrica che meglio aderisce ai dati. Viene quindi calcolato l'errore e individuato il piano che meglio segmenta l'insieme dato. Per ogni sottoinsieme viene calcolata la superquadrica che rappresenta i dati in modo ottimale. Se il procedimento viene iterato, per esempio segmentando il sottoinsieme di punti peggio approssimati, per la superficie viene individuata una successione di modelli a complessità crescente che possono essere associati a diverse risoluzioni.

Non si può però parlare di gerarchia, in quanto la rappresentazione ad una data risoluzione non fa uso del modello alla risoluzione precedente (se non in modo parziale).

I paradigmi gerarchici costituiscono un particolare, ma interessante, sottoinsieme dei paradigmi multirisoluzione. Come è stato più formalmente descritto in §2.2.1, essi hanno una struttura a livelli (o strati), dove ogni livello è caratterizzato da una data scala (generalmente codificata da un opportuno parametro). Ogni livello è responsabile della codifica delle informazioni presenti alla scala corrispondente. La ricostruzione ad una data risoluzione viene ottenuta componendo i livelli corrispondenti ad una scala maggiore o uguale a quella data. La rappresentazione ha quindi una struttura gerarchica.

Una, sia pur debole, struttura gerarchica è ravvisabile nel paradigma delle superquadriche deformabili [TM91], descritto in un paragrafo precedente. I parametri che governano la forma degli ellissoidi superquadri-

2. Metodi e tecniche per la digitalizzazione

ci utilizzati come domino codificano le caratteristiche a larga scala della superficie, mentre le spline utilizzate per deformarne la forma codificano il dettaglio, cioè le caratteristiche a bassa scala.

La struttura gerarchica è maggiormente evidente nei paradigmi che fanno uso di elementi “omogenei” per la rappresentazione dei diversi livelli di dettaglio. Paradigmi di questo tipo possono essere derivati dalla approssimazione di funzioni, dove l’approssimazione viene scelta tra gli elementi di uno spazio vettoriale di funzioni generato dalle unità costitutive del paradigma.

Le *Hierarchical Radial Basis Functions* (HRBF) [BF98] [BF00], sviluppate durante questa tesi, sono costituite da più livelli formati come combinazione lineare di gaussiane. Ogni livello è costituito da gaussiane equispaziate, aventi la stessa deviazione standard — che ne determina la regione di azione e quindi la scala. La configurazione dei pesi della combinazione lineare è basata sull’analisi dei dati che cadono nella regione della corrispondente gaussiana. Un primo strato di gaussiane ad ampia scala viene impiegato per approssimare la forma generale della superficie. Il dettaglio viene codificato come differenza tra i dati originali e l’approssimazione. Il parametro di scala viene diminuito ed un nuovo strato viene generato per approssimare il dettaglio. Il procedimento può essere iterato fino a raggiungere la massima risoluzione consentita dalla densità dei dati. Il procedimento utilizzato per configurare il modello, non iterativo e basato su elementi della teoria del segnale, fa quindi delle HRBF un paradigma *coarse to fine*.

In modo simile, le *Multilevel B-spline*, proposte in [LWS97], operano la ricostruzione sovrapponendo diversi livelli di B-spline. Partendo da un reticolo di punti di controllo poco denso, i coefficienti del primo strato vengono ottenuti con una ottimizzazione ai minimi quadrati, considerando solo i punti che cadono nella regione di influenza della base corrispondente. Il dettaglio viene calcolato come differenza tra i dati originali e l’approssimazione del primo strato. Un nuovo livello di B-spline, il cui reticolo di punti di controllo è ottenuto infittendo il reticolo del livello precedente, viene quindi utilizzato per ricostruire il dettaglio. Un procedimento analogo è utilizzato per la ricostruzione con *Hierarchical Spline* in [FB88] [FB95], dove però la possibilità di ricostruzione della superficie è limitata ai dati disposti in modo equispaziato. Per esse, tuttavia viene proposto anche un algoritmo di configurazione *fine to coarse* in [FW98], dove la rappresentazione multirisoluzione viene derivata da una rappresentazione B-spline iniziale alla massima risoluzione.

Tradizionalmente, anche le Analisi Multirisoluzione (*Multiresolution Analysis* — MRA) basate sulla trasformata wavelet [Mal89] [Cha95] operano sui dati equispaziati. Una MRA è caratterizzata da una cop-

pia di funzioni che generano spazi tra loro complementari: la *scaling function* e la *wavelet*. La *scaling function* viene utilizzata per rappresentare l'approssimazione, mentre la *wavelet* viene usata per il dettaglio. La ricostruzione viene effettuata in modalità *fine to coarse* calcolando i coefficienti approssimazione e dettaglio come risultato della convoluzione dei dati con una coppia filtro (passa-basso per l'approssimazione e passa-alto per il dettaglio). L'operazione viene poi iterata sui coefficienti approssimazione, realizzando così una ricostruzione multirisoluzione.

Alcuni lavori presenti in letteratura si sono occupati del problema della generazione di una rappresentazione MRA di dati non equispaziati. In [PR99] i dati sono costituiti da una nuvola di punti tridimensionali. Si ipotizza che la superficie originale abbia la topologia ed il genus di una sfera. La mesh poligonale iniziale viene costruita adattando una sfera geodesica all'insieme di punti $3D$. Alla mesh risultante viene applicata la trasformata wavelet definita su un dominio sferico presentata in [SS95].

In [FE98] viene introdotta una tecnica *fine to coarse* per ottenere una ricostruzione wavelet a partire da dati campionati in modo non uniforme, limitata al caso monodimensionale. I coefficienti di scaling function e wavelet vengono ottenuti minimizzando l'errore di ricostruzione risolvendo un sistema lineare con il metodo dei minimi quadrati. Una delle ipotesi sulle quali il metodo si basa è che la topologia dell'insieme di dati sia quella della retta reale.

HRBF e MRA rivestono un ruolo di particolare interesse ai fini di questa tesi. Essi verranno descritti in modo più approfondito nel capitolo 4, dove le loro caratteristiche di approssimatori verranno messe a confronto.

2.4.3. Parametrizzazione

Data una superficie di equazione $F(x, y, z) = 0$ nelle variabili x , y e z , si cerca una *parametrizzazione* della soluzione dell'equazione in termini di tre funzioni di due variabili, s e t :

$$(x, y, z) = (X(s, t), Y(s, t), Z(s, t)), \quad \forall x, y, z \in \mathbb{R} : F(x, y, z) = 0$$

La parametrizzazione della superficie è utile per alcune elaborazioni. Innanzitutto, una superficie parametrizzata può essere utilizzata come dominio per costruire una superficie esplicita. Inoltre la parametrizzazione è necessaria per le tecniche di *mapping*.

La parametrizzazione di una superficie non è unica. Per applicazioni quali quelle menzionate, l'uniformità della parametrizzazione consente di ottenere buoni risultati. Pertanto, l'uniformità (cioè una corrispondenza tra distanza topologica e distanza nello spazio parametrico) è uno

2. Metodi e tecniche per la digitalizzazione

dei fattori utilizzati per la scelta della parametrizzazione. Una parametrizzazione veramente uniforme, se la curvatura della superficie non è costante, non è realizzabile. Quella che viene cercata è quindi una parametrizzazione *smooth* [EDD⁺95].

La smoothness della geometria di un modello non implica la smoothness della parametrizzazione. La smoothness geometrica descrive come varia la normale alla superficie, mentre la smoothness parametrica è una misura della variabilità della parametrizzazione. La smoothness parametrica è importante per la qualità dei risultati di alcuni algoritmi [GSS99]. In [GSS99] viene proposta una procedura di *remeshing* (procedura equivalente al ricampionamento di un segnale) che genera, per un dato modello poligonale, un modello avente la stessa smoothness geometrica, ma con i vertici disposti in modo più uniforme. Il nuovo modello può essere utilizzato come base per ottenere una parametrizzazione *smooth* del modello originale.

In [EDD⁺95] viene costruita la parametrizzazione di un modello poligonale su un dominio costituito da un modello base con poche facce triangolari avente la stessa topologia del modello originale. Sia per costruzione del modello base (ottenuto semplificando il modello originale) che per la costruzione della parametrizzazione, vengono utilizzate le mappe armoniche (*harmonic maps*). Una mappa armonica è una funzione che proietta i punti di una mesh avente la topologia di un disco in un sottoinsieme del piano in modo da mantenere il più possibile i lati dei singoli triangoli tra loro proporzionali e in modo da minimizzare le variazioni di lunghezza dei lati.

In [EH96] viene trattato il problema di ricostruire una superficie attraverso un reticolo di patch B-spline. Un modello topologicamente equivalente alla superficie data, ma con un numero di facce (triangolari) sensibilmente inferiore, viene costruito mediante la procedura in [EDD⁺95], ottenendo anche una funzione di proiezione della superficie originale sul modello base. Invertendo quest'ultima, si ottiene la parametrizzazione su patch triangolari. Poiché le B-spline sono costruite su un dominio rettangolare, le patch triangolari vengono abbinate due a due, costruendo così delle patch quadrangolari, avendo cura di minimizzare la distorsione parametrica. Il problema dell'abbinamento di triangoli è stato affrontato anche in [PYL00]. Poiché il problema è NP-completo, per la sua soluzione sono utilizzate alcune euristiche (basate sui grafi). La parametrizzazione in [PYL00] viene effettuata proiettando i punti sul modello a *patch* quadrangolari e calcolando le coordinate relative ai vertici della patch corrispondente.

In [LM98] viene presentato un metodo per il calcolo dello spazio parametrico di una mesh triangolare tale da minimizzare le distorsioni.

La procedura è modulare e si adatta alla elaborazione interattiva, dove l'utente può introdurre vincoli quali linee isoparametriche e regioni dove la distorsione è più critica.

In [LSS⁺98] viene proposto un algoritmo per la costruzione di una parametrizzazione smooth tra i diversi modelli poligonali di una rappresentazione multirisoluzione (MAPS). In particolare, viene definita la parametrizzazione del modello poligonale originale e la sua copia a bassa risoluzione (chiamata base) che costituisce il supporto della superficie multirisoluzione. Ogni punto della superficie originale, p , viene fatto corrispondere tramite una funzione proiezione, Π , al punto $p^0 = \Pi(p)$ della base. Il punto p_0 viene espresso come:

$$p^0 = \alpha p_i + \beta p_j + \gamma p_k$$

dove p_i , p_j e p_k sono i vertici del triangolo della base al quale p^0 appartiene e α , β e γ sono le coordinate baricentriche⁷: $\alpha + \beta + \gamma = 1$.

In [GVSS00] la parametrizzazione è un fattore critico per una buona prestazione del metodo di rappresentazione (*normal meshes*). La parametrizzazione inizialmente ottenuta mediante MAPS viene modificata sia cambiando i vertici che compongono il modello base, sia modificando la parametrizzazione dei vertici dei livelli risoluzione più elevata. In questo caso, lo scopo non è ottenere una parametrizzazione più uniforme, ma posizionare i punti a maggiore curvatura verso il centro dello spazio parametrico.

2.4.4. Integrazione

L'*integrazione* è necessaria quando si desidera acquisire oggetti ad un livello di dettaglio molto fine rispetto alle dimensioni dell'oggetto e/o quando la regione di interesse dell'oggetto non è osservabile da un solo sensore. In questi casi, per poter costruire una copia digitale dell'oggetto è necessario che i sensori ne acquisiscano le forme da più punti di vista. Le informazioni ricavate dalle diverse viste dell'oggetto devono poi essere integrate per fornire un'unica rappresentazione. Tale operazione può essere scomposta nelle seguenti fasi:

Registrazione le diverse sorgenti vengono riferite ad un'unico sistema di riferimento.

Fusione le informazioni delle diverse sorgenti vengono inserite in un unico modello.

⁷Fissando questa condizione, lo spazio parametrico è bidimensionale.

2. Metodi e tecniche per la digitalizzazione

Registrazione

La *registrazione* è un'operazione comune a molti settori, dove vi sia la necessità di unire diverse sorgenti di informazione, riferendole ad un comune modello. In particolare, nel caso della ricostruzione tridimensionale, le acquisizioni possono differenziarsi per:

sessione l'acquisizione viene operata in diversi momenti, per esempio a causa del ridotto numero di sensori disponibili;

senso diversi sensori possono essere impiegati per acquisire differenti caratteristiche dell'oggetto;

soggetto quando si intende comparare le caratteristiche di due oggetti dello stesso tipo o quando si vogliono attribuire le caratteristiche di un oggetto ad uno dello stesso tipo⁸.

La registrazione può essere facilitata dalla presenza di punti di repere comuni alle diverse acquisizioni o da informazioni di calibrazione del dispositivo di acquisizione. Un caso molto frequente è la registrazione dei dati raccolti in un intervallo di tempo. In questo caso, la *vista* è in realtà composta da una successione di (sotto)*viste* acquisite nei diversi istanti di tempo. La registrazione delle *viste* corrisponde quindi all'inversa della funzione che descrive il moto dell'oggetto rispetto al sensore. Questo caso può verificarsi, per esempio, quando l'oggetto da acquisire è posto su una tavola rotante (scansione cilindrica).

Se non sono disponibili punti di repere affidabili, è necessario individuare nelle diverse *viste* punti omologhi (e più in generale *caratteristiche*) per calcolare la trasformazione che li fa sovrapporre. Inoltre, poiché in generale non accade che tutte le *viste* abbiano una sovrapposizione con tutte le altre, bisogna evitare che si formi una partizione dell'insieme di *viste* tale che le *viste* appartenenti alla stessa classe siano ben registrate tra loro, ma che le diverse classi siano disallineate [Pul99]. Tale situazione è illustrata in figura 2.24.

I problemi da risolvere per ottenere la registrazione sono quindi due:

- ottenere una buona stima delle *caratteristiche*;
- realizzare una buona diffusione dell'errore di registrazione.

L'algoritmo *Iterative Closest Point* (ICP) [BM92] è uno dei più citati in letteratura. La sua validità è stata dimostrata provando un teorema

⁸Questa procedura è poco usata per ottenere una copia digitale di un oggetto reale, ma può essere usata, per esempio, per comparare i dati di un soggetto con quelli di un atlante medico.



Figura 2.23.: MAPS: il modello a bassa risoluzione è utilizzato come base per la parametrizzazione che poi viene propagata ai vertici del modello originale [LSS⁺98].

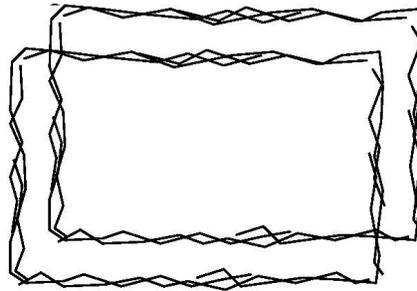


Figura 2.24.: La formazione di *clique* è una situazione che la registrazione deve evitare [Pul99].

2. Metodi e tecniche per la digitalizzazione

di convergenza formulato per una “qualsiasi” modalità di rappresentazione [BM92]. L’ICP è basato sulla definizione di una funzione distanza tra punti delle entità da registrare. Per un dato insieme di punti, viene calcolata la rototraslazione (espressa per mezzo di quaternioni) che minimizza la distanza dai corrispondenti punti più vicini di una superficie di riferimento. Iterando il procedimento, si ottiene la registrazione delle sue entità. Questo algoritmo può rimanere intrappolato in minimi locali. Necessita quindi di una buona inizializzazione e di una ampia sovrapposizione tra le entità da registrare.

In [TL94] viene utilizzato un ICP modificato. Le *viste* sono rappresentate come *range image*. Per ognuna di esse viene costruito un insieme di superfici poligonali, in modo da evitare cautamente la possibilità di introdurre connettività non esistenti. I vertici vengono inoltre pesati con una misura di confidenza della bontà della misura. L’ICP viene applicato solo ai vertici che trovano un corrispondente sull’altra mesh ad una distanza inferiore ad una certa soglia. Per velocizzare l’allineamento, le operazioni vengono effettuate inizialmente su una rappresentazione a risoluzione ridotta (ottenuta semplicemente sottocampionando le *viste*) e incrementando via via la risoluzione. La registrazione di viste multiple viene realizzata scegliendo una *vista* di riferimento ed allineando le altre ad essa. L’ipotesi su cui si basa questa soluzione è la disponibilità di una scansione cilindrica, la quale possa avere regioni in comune con tutte le altre viste.

In [Pul99], dove viene riportata anche una buona revisione (sia pure sintetica) dei metodi *parwise-matching* iterativi, il problema della diffusione dell’errore viene risolto operando prima una registrazione a coppie e utilizzando le informazioni derivanti da questa fase come vincoli per la fase di registrazione globale.

La registrazione in [DWJM98] viene effettuata tramite una modifica dell’algoritmo in [CM92a] (anche [CM91]). L’algoritmo originale si basa sulla minimizzazione della distanza tra i punti di una superficie di riferimento e gli omologhi stimati sulla superficie registrata. La modifica consiste nell’utilizzo di una tecnica per stimare la coerenza delle corrispondenze trovate. Le corrispondenze non coerenti non vengono utilizzate per valutare la trasformazione.

In [YA98] la registrazione viene effettuata in due passi, a partire dalle superfici corrispondenti alle diverse viste. Nella prima fase, vengono individuati i punti con gradiente elevato. L’ipotesi è che questi punti siano peculiari e che quindi le curve poligonali che li uniscono siano *caratteristiche* robuste. L’allineamento viene quindi calcolato allineando le curve.

Fusione

A causa di errori di calibrazione e di ricostruzione, le superfici anche dopo la registrazione, non sono perfettamente sovrapposte. La semplice unione dei dati di due *viste* registrate o delle superfici da essi ricostruite produce, in generale, una rappresentazione poco fedele della superficie dell'oggetto, con increspature nella regione di sovrapposizione. La *fusione* ha quindi lo scopo di individuare una rappresentazione che sostituisca le *viste* registrate, mediando opportunamente le informazioni da esse derivate.

In generale, ogni dispositivo di acquisizione ha una distribuzione dell'errore che dipende dalle caratteristiche dell'acquisizione. Per esempio, l'incertezza sui dati che formano il bordo della regione acquisita è generalmente superiore che all'interno della regione. Queste informazioni, dipendenti dal dispositivo di acquisizione, possono essere impiegate nella procedura di fusione per pesare l'affidabilità della ricostruzione delle differenti viste.

In [DWJM98], dove ogni vista è costituita da una *range image*, l'algoritmo di fusione pesa ogni pixel della *vista* con la distanza (in metrica 4-vicinanza) dal pixel del bordo più vicino. La profondità di ogni pixel dell'immagine risultante dalla fusione viene poi calcolata utilizzando una media pesata. I pixel del bordo vengono quindi "piegati" in modo smooth. La procedura può essere modificata per operare la fusione di più *viste*. In tale caso, si ricorre ad una voxelizzazione dello spazio occupato dall'oggetto. Per ogni voxel occupato viene calcolato un punto della mesh risultante, la posizione del quale è la media pesata dei punti che cadono nel voxel considerato.

La fusione in [TL94] viene realizzata in due passi. Inizialmente l'attenzione viene prestata alla ricostruzione di una mesh che abbia la stessa topologia dell'oggetto. Successivamente, l'accuratezza geometrica viene recuperata adattando la mesh precedentemente costruita ai dati registrati. Al primo passo vengono rimosse le parti delle mesh che si sovrapporrebbero e vengono "incollati" i bordi, triangolando i vertici che lì vi giacciono. La procedura può generare triangoli di piccole dimensioni in corrispondenza dei bordi. Una opportuna procedura di semplificazione provvede a rimuoverli. La mesh così ottenuta viene considerata attendibile per quanto riguarda la topologia dell'oggetto, ma non per la sua geometria (alcuni punti, ritenuti ridondanti, vengono infatti scartati). Per ricostruire anche i dettagli, la mesh è sottoposta ad una fase di raffinamento basato sulla modifica dei suoi vertici nella direzione della normale della superficie. Per ogni vertice della mesh finale viene calcolato l'insieme di vertici delle mesh originali che cadono in un suo intorno

2. Metodi e tecniche per la digitalizzazione

sferico. La normale alla superficie finale nel vertice considerato viene stimata come media delle normali nei vertici selezionati. Il vertice viene riposizionato nella media dei punti risultanti dall'intersezione della normale con mesh originali.

Altre tecniche possono essere utilizzate per la fusione. Per esempio, in [NBC⁺99] vengono presentate alcune tecniche per integrare l'informazione contenuta in immagini volumetriche utilizzando la trasformata wavelet. Le immagini originali vengono scomposte nelle loro componenti (alta e bassa frequenza), integrate in questi domini e poi ricomposte.

La fusione può riguardare anche *viste* provenienti da sensori di diversa natura. Per esempio, in [YKOL96] viene fusa l'informazione sulla posizione di spigoli (acquisita mediante stereo matching) con quella di superficie (laser stripe).

2.4.5. Applicazione del colore

La stima del colore della superficie è una particolare operazione di integrazione delle informazioni. Le informazioni fotometriche devono infatti essere allineate a quelle geometriche.

Il colore apparente dipende da diversi parametri, tra i quali la curvatura della superficie e la posizione relativa delle sorgenti luminose e dell'oggetto rispetto all'osservatore. Il colore recepito dal sensore può quindi essere molto diverso dal colore effettivo dell'oggetto.

Per una stima accurata del colore effettivo è quindi necessario conoscere la posizione delle fonti luminose e la forma dell'oggetto.

Diverse tecniche sono state sviluppate per la stima del colore in modo robusto. Il tema comune è l'utilizzo di molteplici viste e di molteplici luci per evitare i riflessi e le ombre. Inoltre, la conoscenza della posizione del sensore e della curvatura della superficie (ricavate dopo una opportuna registrazione del sensore-colore rispetto ai sensori per la geometria [RB99] [Rig00]) possono essere utilizzate per stimare l'affidabilità relativa di una immagine e quindi ottenere il colore effettivo come media pesata dei colori osservati nelle diverse *viste* [PARD⁺98] [DTM96] [SI96] [RCMS99]. In tal caso è necessario un modello geometrico della superficie.

In [LPC⁺00] per evitare i riflessi, vengono scartati i pixel acquisiti perpendicolarmente. Inoltre, poiché ogni punto viene acquisito più volte, è possibile utilizzare la affidabilità dell'acquisizione per stimare il colore mediante una media pesata. In [RB99], inoltre, i pixel troppo chiari o troppo scuri rispetto alla tonalità dei pixel vicini non vengono considerati per la stima.

In [RB99] viene ipotizzato un modello lambertiano di illuminazione e riflessione. La conoscenza di un modello geometrico approssimato consente una stima iniziale delle normali. Cinque immagini prese dallo stesso punto di vista, ma con differenti sorgenti di illuminazione di direzione conosciuta, consentono di calcolare i parametri fotometrici del modello.

Le informazioni fotometriche devono poi essere utilizzate per assegnare il colore al modello geometrico. I paradigmi maggiormente utilizzati per questo scopo sono due:

- *texture mapping*
- *per vertex color*

Il *texture mapping* [Cat74] [BN76] è una delle più utilizzate nella sintesi di immagini per dare realismo all'immagine prodotta [Hec86]. Questa tecnica può essere utilizzata per conferire altre caratteristiche oltre a quelle di colore, sia geometriche (normali alla superficie, *displacement*) che cromatiche (specularità, trasparenza, illuminazione). L'utilizzo del *texture mapping* implica la conoscenza (e quindi la preventiva costruzione) di una parametrizzazione della superficie.

Il *per vertex color* [Rig00] [LPC⁺00] è una tecnica che permette l'utilizzo di una procedura di visualizzazione computazionalmente leggera. Inoltre, anche la costruzione del modello risulta più economica, in quanto non necessita né di parametrizzazione né di costruzione della texture.

2.5. Elaborazioni post-ricostruzione

Un modello 3D può essere utilizzato in diverse applicazioni. A seconda delle caratteristiche dell'applicazione, può essere necessario sottoporre il modello ad una elaborazione post-ricostruzione. Per esempio, una applicazione di realtà virtuale richiederà un modello con meno poligoni possibile, mentre il suo utilizzo per effetti speciali di un film potrebbe necessitare di un ritocco (più o meno pesante) o della inclusione di una struttura scheletrica per la sua animazione. Le trasformazioni possono essere facilitate da un opportuno paradigma utilizzato per rappresentare la superficie.

Fra le elaborazioni più comuni che in genere si operano sul modello si possono menzionare:

- la *conversione*
- la *compressione*

2. Metodi e tecniche per la digitalizzazione

- il *watermarking*
- la *modifica*
- l'*animazione*
- il *morphing*

La conversione del modello viene applicata per ottenere una rappresentazione delle informazioni contenute nel modello in un formato adatto ad essere elaborato da altre applicazioni. Una comune operazione di conversione è la triangolazione della superficie di un modello [HG99].

Un aspetto particolare della conversione è la riparametrizzazione. In [LSS⁺98] viene presentato un metodo multirisoluzione per parametrizzare una mesh triangolare. La mesh originale viene via via semplificata (in $O(N)$ livelli, N numero di vertici) fino ad ottenere una mesh con pochi poligoni, la quale viene utilizzata come base per la parametrizzazione. I vertici rimossi vengono cioè associati ad uno dei triangoli della base e vengono espressi in coordinate relative ai vertici del loro triangolo base. La descrizione risultante permette di realizzare una ristrutturazione della mesh, per esempio per distribuire i triangoli in modo più equilibrato o con una connettività costante e predeterminata. Quest'ultima caratteristica permette di utilizzare facilmente gli strumenti legati alle wavelet.

La compressione viene generalmente operata in modo *lossy* (cioè con perdita d'informazione). Lo scopo è, dato una certa quantità di risorse a disposizione per la rappresentazione (per esempio, un numero massimo di triangoli), minimizzare la perdita di informazione percepibile. L'utilizzo di una codifica in grado di decorrelare i coefficienti usati per la rappresentazione consente in genere di ottenere una buona compressione.

Il *watermarking* [PHF99] è un problema che per alcuni aspetti è legato alla compressione. In questo caso si tratta modificare la rappresentazione dell'oggetto inserendovi dell'informazione non visivamente percepibile. Solo una opportuna elaborazione è in grado di estrarre l'informazione aggiunta, rendendo così possibile il riconoscimento del modello. Il *watermarking* rende possibile la distribuzione dei modelli prodotti, permettendo però la rintracciabilità del singolo modello. La caratteristica fondamentale di una procedura di *watermarking* è la sua robustezza. Esso deve essere resistente alle operazioni più comuni, quali per esempio, variazione di scala, ricampionamento e sezionamento.

La *modifica*, l'*animazione* ed il *morphing* sono elaborazioni che vengono effettuate spesso nel settore della *computer graphics*.

Sebbene non rientrino nelle tematiche studiate in questa tesi, le elaborazioni post-ricostruzione sono interessanti perché evidenziano le caratteristiche che un paradigma per la rappresentazione di oggetti tridimensionali deve possedere. Nei prossimi paragrafi, verranno trattati più approfonditamente i temi legati all'ottimizzazione della rappresentazione poligonale (*surface simplification*, §2.5.1), della modifica (*Modeling/Editing*, 2.5.2), della animazione (*Animation*, 2.5.3) e del *morphing* (§2.5.4).

2.5.1. Surface simplification

Nelle applicazioni di realtà virtuale e nei CAD tridimensionali l'interattività, intesa come risposta immediata del programma all'input dell'utente, è una condizione fondamentale per l'efficacia del programma: per esempio, nei CAD, un lungo tempo di risposta ad un comando di modellazione o una mancanza di fluidità nell'ispezione di un oggetto possono stancare l'utente o diminuire l'intuitività di una operazione. In alcune situazioni, la descrizione di una superficie è ridondante; per esempio, se un oggetto è situato molto distante dal punto di vista, molti dei suoi elementi costitutivi saranno visualizzati dallo stesso pixel, oppure se l'oggetto è in movimento rispetto all'osservatore la precisione della visualizzazione può essere sacrificata in favore di una maggiore fluidità del movimento stesso (cioè del *frame rate*).

In casi come questi, una descrizione della superficie troppo ricca di dettagli è inutile e rischia di essere dannosa per l'efficienza dell'applicazione. Sono quindi state studiate diverse tecniche per ottenere una descrizione della stessa superficie a diversi livelli di dettaglio. Il problema viene chiamato *surface simplification*: ottenere una descrizione più compatta della superficie (cioè utilizzando meno elementi), minimizzando la differenza tra la superficie originale e quella semplificata.

Poiché molto spesso le superfici sono rappresentate attraverso un insieme di triangoli (vedi §2.2.1), il problema viene spesso ricondotto alla semplificazione di mesh poligonali (*mesh simplification*), cioè alla riduzione del numero di triangoli necessario per rappresentare un oggetto [CMS98]. Le tecniche di *mesh simplification* si differenziano dalle tecniche citate nel paragrafo 2.4.2 (ricostruzione multirisoluzione da punti) per l'approccio alla costruzione della superficie multirisoluzione: nella *mesh simplification* è necessario avere una descrizione della superficie in formato poligonale, mentre gli approcci precedentemente illustrati sono in grado di operare una ricostruzione a partire da un insieme di punti campionati sulla superficie stessa. In generale, quindi, le tecniche di semplificazione vengono utilizzate dopo la ricostruzione. Il loro uso è in-

2. Metodi e tecniche per la digitalizzazione

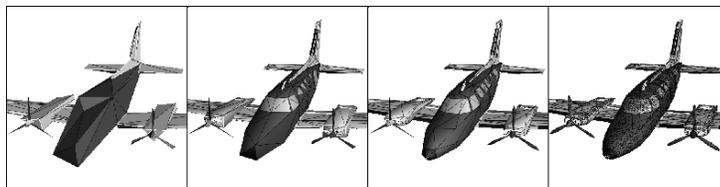


Figura 2.25.: Descrizione multirisoluzione di una mesh [Hop96]

fatti legato ai vantaggi che si possono ottenere nella sintesi di immagini tridimensionali.

L'obiettivo della semplificazione è di preservare, per quanto possibile, l'*aspetto* della superficie dell'oggetto, non la sua geometria o la sua topologia.

Una prima classificazione degli algoritmi di surface simplification può essere operata in base al fatto che la procedura sia studiata per un preprocessing dei modelli oppure che la semplificazione debba essere effettuata in real time. Il preprocessing viene effettuato per includere nella rappresentazione dell'oggetto informazioni riguardanti la forma dell'oggetto stesso a diversi gradi di precisione. Queste descrizioni possono essere essenzialmente di due tipi [CCMS97]: *Level of Detail* (LOD) e *Multirisoluzione*. Una descrizione LOD è composta da un insieme di modelli dell'oggetto in questione caratterizzati da diversi gradi di dettaglio. Una descrizione multirisoluzione contiene le informazioni necessarie per passare, con una certa continuità, attraverso operazioni "elementari", da un modello a descrizione minima (può essere anche un triangolo) al modello alla massima risoluzione.

Gli algoritmi di *mesh simplification* si possono inoltre differenziare a seconda dell'input e dell'output. La maggior parte degli algoritmi accetta in ingresso una rappresentazione della superficie descritta tramite un insieme di vertici ed un insieme di triangoli descritti attraverso una terna di vertici (vedi §2.2.1). Alcuni algoritmi sono in grado di gestire anche informazioni relative al colore o ad altre caratteristiche dei vertici o dei triangoli.

Vertex Clustering

Questa procedura di semplificazione è stata proposta da Rossignac e Borrel nel 1993 [RB93]. Richiede come input una struttura dati composta da una lista di vertici (coordinate $3D$) e da una lista di facce poligonali (non necessariamente triangolari) descritte tramite la sequenza di vertici che le delimitano. L'algoritmo è composto da sette fasi:

- *Grading*: Ad ogni vertice viene associato un peso che definisce l'importanza percettiva del vertice. Sono tenuti in considerazione i vertici che hanno una maggiore probabilità di essere sul profilo dell'oggetto e quelli che non contengono piccoli dettagli. Il primo fattore viene stimato mediante l'inverso del maggior angolo formato dalle coppie di lati che incidono sul vertice. Il secondo fattore viene stimato attraverso l'area delle facce alle quali il vertice appartiene.
- *Triangolazione*: Ogni faccia è decomposta in triangoli. La lista delle facce contiene così solo sequenze di tre vertici.
- *Clustering*: I vertici vengono quindi raggruppati in cluster, in base alla vicinanza geometrica. Il *bounding box* viene suddiviso uniformemente in celle (parallelepipedi). Il risultato di questa procedura è una tabella contenente per ogni vertice l'indice della cella alla quale appartiene.
- *Sintesi*: Viene calcolato il rappresentante per ogni cella. Una scelta comunemente utilizzata è il vertice più vicino alla media pesata dei vertici appartenenti alla cella. Ogni triangolo verrà ora ricostruito mediante i tre vertici rappresentanti le celle nelle quali si trovavano i vertici originali.
- *Eliminazione*: I triangoli originali possono essere degenerati in segmenti o in punti, a seconda che due dei vertici originali o tutti e tre siano stati mappati nello stesso vertice rappresentante. Se questi triangoli degeneri appartengono già ad un triangolo nella versione semplificata del modello possono essere rimossi. Inoltre più triangoli possono essere stati mappati nello stesso triangolo e quindi anche essi possono essere eliminati dalla descrizione.
- *Calcolo delle normali*: Le facce adiacenti e quasi coplanari vengono raggruppate e la normale al vertice comune viene calcolata come media delle normali alle facce in esame.

Mediante questa procedura, si possono calcolare una serie di modelli a differenti LOD e poi decidere run-time quale modello utilizzare. L'operazione di Grading è molto sensibile alla particolare tassellatura: con gli stessi vertici, una faccia poligonale può essere suddivisa in triangoli in diversi modi, influenzando così sia sul primo che sul secondo fattore del Grading. Se la potenza di calcolo a disposizione lo consente, la semplificazione può essere operata run-time. In tal caso, la regione da suddividere durante il clustering può essere il *bounding box* dell'oggetto

2. Metodi e tecniche per la digitalizzazione

o un unico parallelepipedo contenente la scena (per esempio se si vuole avere lo stesso livello di dettaglio in tutta la scena). Il numero di celle della suddivisione dipende dalla complessità dell'oggetto, dal livello di semplificazione desiderato e dall'importanza dell'oggetto relativamente alla scena. Questa scelta va fatta run-time anche tenendo conto delle caratteristiche grafiche del sistema. Questo approccio non permette di predeterminare in alcun modo il numero di triangoli finali o l'errore di rappresentazione. La scelta dei rappresentanti dei cluster come sottoinsieme dei vertici originali ha il vantaggio di mantenere compatta la descrizione dell'oggetto semplificato, in quanto non richiede spazio per la memorizzazione delle nuove coordinate, ma solo l'indice nella tabella dei vertici. Qualora la semplificazione sia realizzata su un modello già semplificato, al fine di contenere l'errore di approssimazione, è opportuno che la suddivisione al livello più fine sia propriamente contenuta nella suddivisione a scala più ampia.

Questo approccio è stato migliorato in [LT97], dove vengono suggerite le seguenti modifiche:

- *Grading*: utilizzo di $\cos(\frac{\alpha}{2})$ invece di $\frac{1}{\alpha}$, dove α è il massimo angolo di incidenza.
- *Clustering*: viene introdotta la strategia chiamata *floating cell clustering*, che consiste nel centrare ogni cella nel vertice con peso maggiore. Il partizionamento dello spazio non è più uniforme, ma viene eseguito iterativamente, mantenendo una lista dei vertici ancora da assegnare alle celle. In questo modo, il rappresentante è più vicino al vertice con peso maggiore. L'effetto visivo è che i triangoli stretti non tendono ad allargarsi.
- *Eliminazione*: triangoli allungati che collassano in lati vengono visualizzati mediante segmenti di spessore e colore proporzionale alle loro dimensioni e orientamento.

Progressive Mesh

L'approccio proposto da Hoppe in [Hop96] è di tipo multirisoluzione (figura 2.25). La rappresentazione multirisoluzione supporta *geomorphing*, trasmissione progressiva, compressione e raffinamento selettivo. A partire da una mesh triangolare (che può essere dotata anche di caratteristiche quali il colore) produce una rappresentazione multirisoluzione composta da un modello "minimo" e dalle informazioni per aggiungere i dettagli necessari a raggiungere la risoluzione del modello originario. L'algoritmo proposto si basa su una operazione di semplificazione (*edge collapsing*)

invertibile (tramite *vertex splitting*). La qualità delle rappresentazioni a risoluzione ridotta dipendono dall'ordine con il quale si scelgono i lati da collassare. Poiché la procedura viene eseguita off-line si può scegliere un algoritmo di selezione che assicuri un buon risultato percettivo. Allo scopo, viene definita una funzione energia E che misura la differenza tra una mesh M e la mesh originale M_o .

$$E(M) = E_{dist}(M) + E_{spring}(M) + E_{scalar}(M) + E_{disc}(M)$$

$E_{dist}(M)$ misura la distanza tra i punti di M_o dalla mesh M e la mesh originale, $E_{spring}(M)$ misura la lunghezza dei lati della mesh M (regolarizzatore), $E_{scalar}(M)$ misura l'accuratezza degli attributi dei vertici e $E_{disc}(M)$ misura l'accuratezza geometrica delle curve discontinue di M (le forti discontinuità nella curvatura sono fattori percettivamente importanti). All'inizio della semplificazione, viene calcolata la variazione di energia per ogni spigolo e viene così generata una lista di priorità. Ad ogni passo viene collassato il lato che comporta la minore variazione di energia e la lista viene aggiornata calcolando la variazione di energia relativa a ciascuno dei suoi vicini. Per calcolare il valore di E_{dist} relativo al collassamento di un dato lato, vengono presi i vertici di M_o più, opzionalmente, una serie di punti campionati dalle sue facce. Essi vengono proiettati sulla mesh M risultante dal collassamento e l'energia E_{dist} assume un valore proporzionale alla distanza media tra i punti prima e dopo il collassamento. Dopo il collassamento, una coppia di vertici assume una posizione coincidente. Tale punto viene scelto tra i due vertici che delimitano il lato ed il punto medio del lato stesso. La scelta ricade su quello che minimizza la variazione di energia. Il peso di E_{spring} (che equivale alla costante di allungamento elastico) viene fatto diminuire con il procedere della semplificazione. Per il calcolo di E_{scalar} bisogna determinare gli attributi del nuovo vertice in modo da minimizzare una funzione distanza definita sullo spazio degli attributi. E_{disc} rende conto sia delle discontinuità geometriche che di quelle degli attributi. Questa energia è calcolata campionando un insieme di punti dai lati a forte curvatura della mesh originale, e calcolando la distanza di questi punti dai lati a forte curvatura della mesh attuale.

La codifica *Progressive Mesh* di una mesh poligonale consente di raffinare selettivamente la rappresentazione dell'oggetto in funzione delle caratteristiche del dispositivo di visualizzazione [Hop97]. I criteri sui quali il raffinamento è basato prendono in considerazione la regione di spazio visualizzata nel frame considerato, l'orientamento della superficie e l'errore di visualizzazione dello schermo. La procedura di raffinamento o semplificazione viene eseguita in real-time. L'algoritmo utilizza geo-

2. Metodi e tecniche per la digitalizzazione

morphing per suddividere la trasformazione su diversi frame, evitando così l'apparizione o la sparizione istantanea di dettagli (*popping*).

Multiresolution Decimation

Un approccio simile è utilizzato in [CCMS97]. A differenza dei metodi precedenti, l'insieme dei vertici utilizzati nelle mesh alle varie risoluzioni è sempre un sottoinsieme dei vertici di partenza. Questo fatto consente di mantenere contenuta la dimensione della rappresentazione multirisoluzione. Ad ogni passo viene eliminato un vertice e la superficie viene ritriangolata in modo da minimizzare una funzione costo che tiene conto della variazione dell'aspetto della superficie. Gli errori introdotti dalla eliminazione di un vertice e dalla successiva ritriangolazione vengono calcolati in modo accurato. Il modello viene memorizzato come insieme di vertici e di facce. Per ogni faccia, viene memorizzata l'errore al momento della sua apparizione e della sua scomparsa. Ciò consente di ottenere, per un dato errore, il rappresentante del modello originale generato come composizione delle facce il cui intervallo di errore include il valore dato.

Quadric Error Metrics

Un approccio simile a quello in [Hop96] viene proposto in [GH97]. In esso, l'operazione di *edge collapsing* viene sostituita con quella di *pair contraction*. Entrambe le operazioni sostituiscono una coppia di vertici con un vertice solo, ma mentre la prima limita la selezione della coppia ai vertici topologicamente adiacenti, la seconda non pone questo limite. Questa particolarità permette la fusione di parti separate, ma vicine, di uno stesso modello. In questo modo viene mantenuto l'aspetto globale dell'oggetto rappresentato — o, più in generale, degli oggetti che appartengono alla scena. Ad ogni vertice è associata una quadrica (attraverso una matrice) rappresentante la distanza del vertice dai piani delle facce originali. Il vertice risultante dalla contrazione eredita le quadriche come somma delle matrici dei genitori. La posizione del nuovo vertice è calcolata, minimizzando una funzione errore quadratica, come soluzione di un sistema lineare. Ad ogni passo viene operata la contrazione che minimizza l'errore di rappresentazione. Data la località della procedura, il calcolo della variazione dell'errore va aggiornato solo per i punti vicini alla coppia collassata. L'algoritmo è stato successivamente esteso [GH98] per includere nel calcolo dell'errore anche attributi dei vertici quali colore e texture. La motivazione teorica della abilità dei metodi

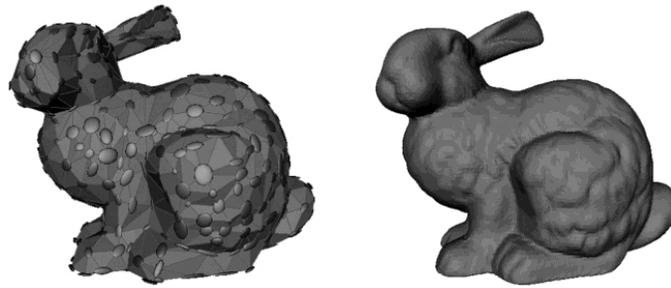


Figura 2.26.: Semplificazione tramite *quadric error metrics* [GH97].

basati sulla metrica di errore quadrica di misurare l'errore geometrico è stata discussa in [HG99].

Una metrica quadrica per misurare la differenza di rappresentazione degli attributi dei vertici (*appearance attributes*) è utilizzata anche in [Hop99].

Analisi Multirisoluzione

[LDW97] hanno proposto un'estensione dell'analisi multirisoluzione wavelet a superfici di topologia arbitraria. Tuttavia, il loro approccio richiede una mesh iniziale con connettività regolare (*subdivision connectivity*). In [EDD⁺95] viene proposto un algoritmo per calcolare una parametrizzazione smooth di una superficie poligonale ad alta risoluzione rispetto ad un dominio costituito da una copia a bassa risoluzione. Questa parametrizzazione può essere utilizzata per ricampionare la superficie ottenendo una mesh a connettività regolare da usare per una rappresentazione multirisoluzione.

Normal Meshes

In [GVSS00] viene proposta una rappresentazione gerarchica denominata *normal mesh*. In tale rappresentazione, una mesh a bassa risoluzione (ma con le stesse proprietà topologiche della superficie finale) viene utilizzata come livello gerarchico iniziale. I vertici degli strati seguenti vengono memorizzati utilizzando una sola coordinata (l'altezza). A partire da una mesh irregolare a bassa risoluzione, la risoluzione viene aumentata utilizzando una suddivisione regolare 1:4. I vertici aggiunti si trovano sulla normale al lato del triangolo originale, nel suo punto mediano. Particolare attenzione deve però essere posta alla costruzio-

2. Metodi e tecniche per la digitalizzazione

ne della trasformazione della mesh originale nella sua rappresentazione “normale”.

2.5.2. Modeling/Editing

Una rappresentazione gerarchica è utile per poter modificare le caratteristiche del modello a diversi livelli di dettaglio [FB95][ZSS97][EDD⁺95]-[LSS⁺98][GSS99][LJ94]. Nella realtà fisica, l'azione a diversi livelli di dettaglio viene operata utilizzando strumenti diversi (per esempio, una lima per sgrossare e un'altra per rifinire), ma la virtualità del modello permette di utilizzare anche la codifica dell'oggetto oltre che dello strumento. Ciò significa che un oggetto virtuale (per esempio, un segnale digitale) può essere elaborato utilizzando strumenti che agiscono a diverse scale (per esempio, filtri passabasso a differenti frequenze di taglio) oppure, se la codifica dell'oggetto lo permette, operando solo sul sottoinsieme di parametri dell'oggetto che codificano le caratteristiche ad una determinata scale (per esempio, i coefficienti wavelet). Indipendentemente dal tipo di multirisoluzione utilizzato per rappresentare la superficie, la modellazione può essere applicata ad una determinata scala agendo sulla codifica al livello di dettaglio corrispondente e lasciando invariati i rimanenti coefficienti della codifica. La località della rappresentazione è una caratteristica ricercata per questo tipo di trasformazioni, in quanto consente di agire su un limitato (e facilmente individuabile) numero di coefficienti.

2.5.3. Animation

L'animazione di un modello 3D si ottiene associando una legge (in funzione del tempo) di movimento dei punti della superficie. In pratica, per modelli minimamente complessi, non è possibile specificare la legge di moto di ogni punto della superficie. Una tecnica molto usata per l'animazione prevede l'utilizzo di una struttura scheletrica per la mappatura del movimento. Generalmente questa struttura è costituita da segmenti connessi fra loro in modo gerarchico e sottoposti a vincoli geometrici. Ogni punto della superficie viene vincolato ad uno o più segmenti e l'animazione viene ottenuta animando solo la struttura scheletrica.

Una rappresentazione gerarchica della superficie facilita l'animazione del modello. La costruzione dello scheletro e la sua successiva associazione al modello può essere effettuata utilizzando il modello a risoluzione più bassa [LSS⁺98] [LMH00].

Se il modello a bassa risoluzione corrisponde al dominio sul quale viene definito il dettaglio, la sua modifica si ripercuote automaticamente ai livelli superiori.

Alternativamente, viene usato lo *space warping* [LJ94]. Anche in questo caso, la modifica può essere effettuata, efficientemente, solo sul modello a bassa risoluzione.

2.5.4. Morphing

Il *morphing* è l'effetto visivo tramite il quale si realizza la trasformazione di un oggetto in un'altro. Il *morphing* viene realizzato mettendo in relazione i parametri dei due diversi modelli mediante una funzione (possibilmente smooth) del tempo. Inoltre, per l'ottenimento di un migliore effetto visivo, è importante che la relazione legghi tra loro i coefficienti dei diversi modelli che codificano le caratteristiche omologhe. Per esempio, per la realizzazione di un *morphing* tra due volti, il naso del volto originale deve trasformarsi nel naso del volto finale.

La presenza di una struttura gerarchica per la rappresentazione delle superfici facilita la realizzazione di tale effetto. Ciò è dovuto al fatto che le caratteristiche principali degli oggetti dovrebbero essere percepibili a scala elevata e quindi la costruzione della funzione di *morphing* può essere effettuata sui coefficienti dei livelli a bassa risoluzione.

Anche in questo caso, la località della rappresentazione può essere un vantaggio [LDSS99].

2.6. Conclusioni

La digitalizzazione di oggetti reali è un'operazione complessa, necessaria in una vasta gamma di settori. I casi reali sono quindi caratterizzati da una forte variabilità delle condizioni in cui la digitalizzazione deve essere effettuata. Nonostante l'ampio spettro di soluzioni proposte, la digitalizzazione, almeno ad un livello di astrazione sufficientemente elevato, può essere descritta come una successione di operazioni ben caratterizzate.

L'acquisizione viene effettuata con lo scopo di collezionare un insieme di dati che quantifichi le caratteristiche dell'oggetto acquisito. La ricostruzione integra fra loro i dati raccolti nella fase precedente, producendo una descrizione unica e accurata dell'oggetto. La fase di ottimizzazione converte il modello ricostruito in un formalismo adatto alle successive elaborazioni.

Lo scopo di questa tesi è di fornire una metodologia per la digitalizzazione. In particolar modo, l'attenzione sarà prestata alla fase di ricostruzione.

2. *Metodi e tecniche per la digitalizzazione*

Nel prossimo capitolo, verrà esposta la metodologia sviluppata nell'ambito di questa tesi. Essa si propone di organizzare in modo strutturato le varie attività che la compongono, esplicitandone le funzionalità e le relazioni che legano le varie fasi elaborative. Nel capitolo 4 verrà approfondita la descrizione di due paradigmi multirisoluzione, MRA ed HRBF, già introdotti nel paragrafo 2.4.2. HRBF è stato sviluppato nell'ambito di questa tesi. Le sue proprietà di approssimatore verranno comparate con quelle di MRA. Nel capitolo 5 verrà presentato un esempio applicativo della metodologia sviluppata, Autoscan.

3. Una metodologia gerarchica per la digitalizzazione

Il problema della generazione di copie digitali di oggetti reali scomponibile in diversi sottoproblemi, categorizzabili in acquisizione dei dati, ricostruzione del modello e ottimizzazione. Nel capitolo precedente sono stati analizzati i problemi e le peculiarità degli approcci proposti in letteratura per trattare con essi.

Una revisione critica degli approcci in letteratura è utile per fornire una visione del problema utilizzando diverse chiavi di lettura. Essa favorisce una categorizzazione dei problemi di maggiore interesse e delle tecniche utilizzate per risolverli.

Una metodologia è uno schema nel quale far rientrare una classe di problemi e le corrispondenti tecniche per la loro soluzione. Essa non fornisce quindi diretti spunti implementativi, ma si pone ad un livello di astrazione più elevato, epurando dai dettagli implementativi legati alle esigenze dell'applicazione specifica e suggerendo una strategia comune alla soluzione di una classe di problemi. L'uso di una metodologia consente la valutazione delle funzionalità e l'esplicitazione delle relazioni tra le varie fasi elaborative piuttosto che delle prestazioni delle singole tecniche.

Alla luce degli aspetti emersi nel capitolo 2, uno schema generale per la digitalizzazione sarà discusso in §3.1.

Il lavoro di questa tesi è stato focalizzato sul problema della ricostruzione. Un contributo innovativo di questa tesi è infatti costituito dalla formalizzazione di una metodologia per la generazione di modelli tridimensionali di oggetti reali a partire da dati campionati sulla superficie degli oggetti. Essa sarà presentata in §3.2. Inoltre, come verrà descritto in §3.3, la ricerca è stata rivolta alla ideazione ed allo sviluppo di un paradigma di rappresentazione del modello che rispondesse ai requisiti richiesti dalla metodologia proposta ed allo sviluppo di una realizzazione della metodologia, cioè di un sistema di digitalizzazione rientrante nello schema tracciato nella metodologia.

3.1. La digitalizzazione

La digitalizzazione si può scomporre in tre fasi: acquisizione, ricostruzione e ottimizzazione.

3.1.1. Acquisizione

Lo scopo della fase di acquisizione è fornire, a partire da un oggetto reale, una collezione di informazioni riguardanti la superficie dell'oggetto stesso. Essa si compone di tre attività: pianificazione, calibrazione e misura.

La *pianificazione* delle acquisizioni ha come scopo il posizionamento dei sensori del sistema in modo da coprire le regioni di interesse della superficie dell'oggetto. La *calibrazione* fornisce una descrizione della trasformazione operata dai sensori. La fase di *misura* utilizza le informazioni di calibrazione per trasformare i dati ricavati dai sensori nella collezione di dati da passare alla fase di ricostruzione.

Le variabili della fase di acquisizione sono legate essenzialmente al tipo di strumento utilizzato per il campionamento ed il tipo di oggetto: dimensioni dell'oggetto, condizioni ambientali, motilità dell'oggetto, sensibilità dell'oggetto a misure invasive, tipo di dispositivo di misura, numero di sensori, risoluzione dei sensori, campo di acquisizione dei sensori, livello di dettaglio desiderato. A seconda della conoscenza o del controllo che si possono esercitare su di esse è possibile determinare le condizioni operative ottimali. Fra le variabili si possono individuare alcuni gruppi tra loro dipendenti. Per esempio, numero di sensori, risoluzione dei sensori, campo di acquisizione dei sensori limitano il livello di dettaglio ottenibile in fase di ricostruzione.

La fase di acquisizione colloquia con la fase di ricostruzione. La prima passa alla seconda la collezione di dati acquisiti, eventualmente corredata da informazioni aggiuntive (per esempio, la posizione dei sensori rispetto ai dati). Da essa può ricevere la richiesta di nuove acquisizioni (dati mancanti o ambigui, maggiore dettaglio) o un modello approssimato per migliorare l'accuratezza delle misure effettuate (per esempio, migliorare la stima del colore effettivo data la curvatura della superficie).

3.1.2. Ricostruzione

La fase di ricostruzione deve fornire un modello digitale tridimensionale dell'oggetto a partire dalla collezione di dati forniti dalla fase di acquisizione. Essa si compone principalmente di tre attività: generalizzazione, registrazione e fusione. Questa fase è di particolare interesse

3.1. La digitalizzazione

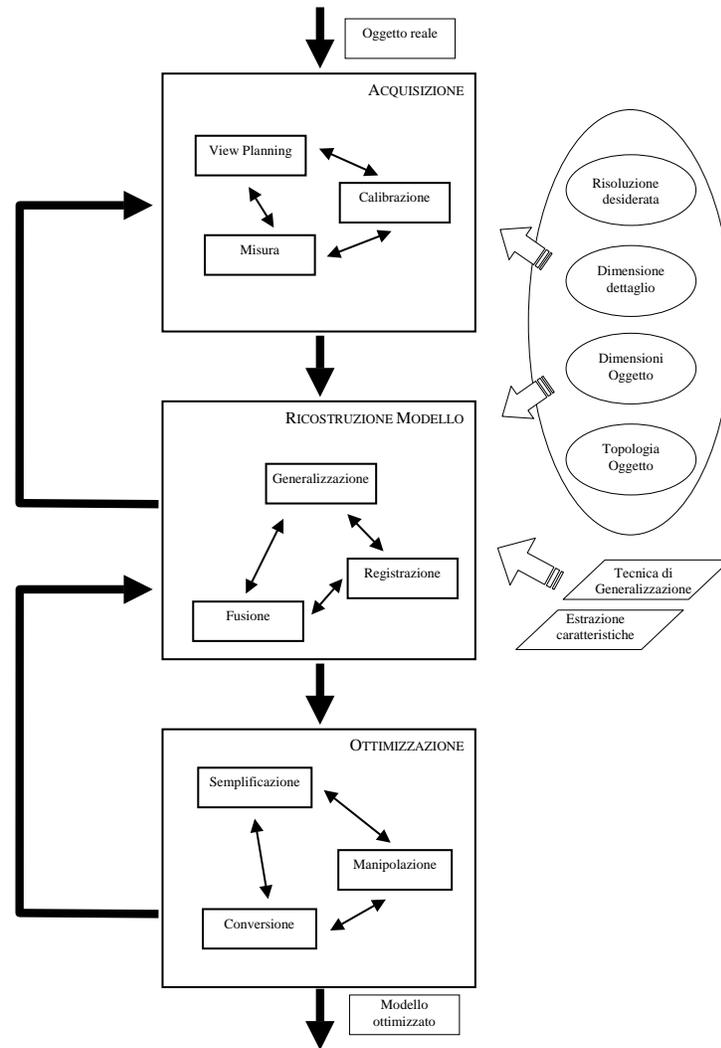


Figura 3.1.: Schema ad alto livello della metodologia di digitalizzazione

3. Una metodologia gerarchica per la digitalizzazione

per questa tesi. Essa sarà discussa in modo più formale ed approfondito nel paragrafo §3.2. Le variabili di questa fase sono: conoscenze a priori sulle caratteristiche dell'oggetto, appartenenza dell'oggetto ad una classe, scelta del paradigma di rappresentazione e dell'algoritmo di configurazione, risoluzione e accuratezza desiderata.

La fase di ricostruzione colloquia con la fase di acquisizione e con la fase di ottimizzazione. Le modalità e le informazioni scambiate sono descritte nei paragrafi dedicati a queste fasi (§3.1.1 e §3.1.3).

3.1.3. Ottimizzazione

La fase di ottimizzazione e conversione finalizza il modello generato dalla fase di ricostruzione per una determinata applicazione obbiettivo. Le attività svolte in questa fase sono quindi fortemente dipendenti dalle caratteristiche richieste dall'applicazione obbiettivo. Tipicamente, vi vengono operate una conversione del modello in un altro paradigma e la sua ottimizzazione. Le variabili sono: formalismo di rappresentazione utilizzato dall'applicazione obbiettivo, risorse computazionali a disposizione dell'applicazione, supporto di visualizzazione utilizzato, esigenze di animazione e/o modellazione.

Questa fase dialoga con la fase di ricostruzione, dalla quale riceve il modello tridimensionale dell'oggetto acquisito ed alla quale, in linea teorica, può richiedere un modello maggiormente dettagliato. Nella pratica, tuttavia, raramente vi è necessità di un ritorno di informazioni verso la fase di ricostruzione, in quanto la risoluzione di acquisizione è generalmente più elevata di quanto i dispositivi di visualizzazione riescano a supportare.

3.2. La ricostruzione

In questa tesi è stata dedicata maggiore attenzione alla fase di ricostruzione, in quanto essa rappresenta il nocciolo del problema della digitalizzazione. Essa sarà quindi analizzata più in dettaglio delle fasi di acquisizione ed ottimizzazione, strutturando e caratterizzando le attività di cui essa si compone in modo più particolareggiato.

La fase di acquisizione fornisce una collezione di *viste* (*view*) dalle quali vengono estratti (campionati) i dati per la ricostruzione. I dati vengono poi elaborati (generalmente mediante uno schema iterativo) per ricostruire un modello tridimensionale dell'oggetto acquisito.

Sia $S^0 = \{S_1^0, \dots, S_N^0\}$ l'insieme di dati campionati ricavati dalle N *viste*. L'indice in apice riporta il numero ordinale di iterazioni eseguite durante il processo di raffinamento — inizialmente 0.

Possono essere individuate essenzialmente due strategie per la ricostruzione di modelli digitali da informazioni raccolte mediante molteplici di *viste*:

- Per ogni insieme, S_i^0 , i dati vengono utilizzati per costruire, tramite generalizzazione, un modello M_i^0 . I modelli delle singole *viste* vengono suddivisi in gruppi di *viste* adiacenti e poi sottoposti ad una trasformazione al fine di riportarli in un unico sistema di riferimento per gruppo (registrazione). Infine, i modelli sono sottoposti ad una trasformazione che ne modifica la geometria nelle regioni di sovrapposizione, al fine di ottenere un unico modello tridimensionale per gruppo (fusione). Al termine di un ciclo di elaborazione appena descritto, si ha la collezione di modelli $M^1 = \{M_1^1, \dots, M_{N^1}^1\}$, dove N^1 è il numero di gruppi formati durante la registrazione dei modelli al passo 0. La procedura di registrazione e fusione viene iterata fino all'ottenimento di un unico modello tridimensionale. Queste operazioni sono schematizzabili come riportato in figura 3.2.
- Per ogni insieme, S_i^0 , i dati vengono utilizzati per costruire, tramite generalizzazione, un modello M_i^0 . I modelli delle singole *viste* vengono suddivisi in gruppi di ricostruzioni parzialmente sovrapposte e, all'interno di ogni gruppo, sottoposti ad una trasformazione al fine di riportarli ad un sistema di riferimento comune, (registrazione). Tale trasformazione viene applicata anche ai dati, generando così la collezione di insiemi $S^1 = \{S_1^1, \dots, S_{N^1}^1\}$. Tali insiemi vengono utilizzati per costruire un'altra collezione di modelli, $M^1 = \{M_1^1, \dots, M_{N^1}^1\}$. Agli elementi di M^1 può essere nuovamente applicata la procedura di registrazione, fino all'ottenimento di un unico insieme di dati dal quale ricavare il modello tridimensionale dell'oggetto acquisito. Questa strategia è stata schematizzata in figura 3.3.

La differenza tra i due approcci risiede essenzialmente nel ruolo della generalizzazione. Nel primo caso, tutta la potenza del metodo scelto per la ricostruzione della superficie è utilizzata nella ricostruzione delle superfici delle varie *viste*: i dati non saranno più utilizzati. La ricostruzione deve quindi estrarre tutta l'informazione che gli è possibile ricavare dai dati. Nel secondo caso, invece, la ricostruzione è utilizzata come supporto per la registrazione e la fusione. I dati vengono aggiornati ad ogni iterazione e un nuovo modello viene ricostruito.

La seconda strategia è quindi più costosa (si ricalcolano i modelli ad ogni iterazione), ma è anche la più scrupolosa: le ricostruzioni possono

3. Una metodologia gerarchica per la digitalizzazione

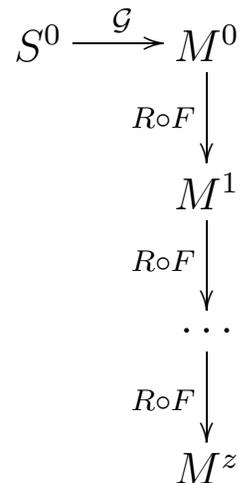


Figura 3.2.: La ricostruzione può essere effettuata utilizzando i dati forniti dall'acquisizione solo per ricostruire le superfici corrispondenti alle viste.

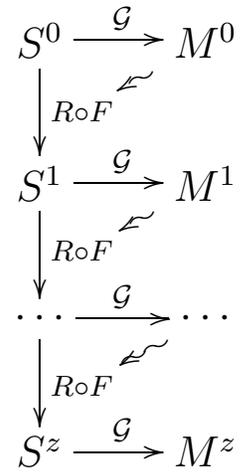


Figura 3.3.: Le elaborazioni vengono effettuate sulle ricostruzioni parziali e poi vengono riportate ai dati.

infatti giovare di insiemi di dati più estesi (effetto ai bordi, ridondanza). Inoltre, se l'operazione di fusione viene definita componendo la registrazione dei dati, la loro fusione e la successiva ricostruzione del modello, la prima strategia può essere vista come un caso particolare della seconda¹.

La multirisoluzione può essere impiegata per ridurre il costo computazionale delle procedure e per conferir loro maggiore robustezza. I paradigmi gerarchici, infatti, consentono di rappresentare la superficie ricostruita a diversi gradi di dettaglio. Questo fatto permette di procedere, nella ricerca delle trasformazioni ottimali, mediante approssimazioni successive. Prendendo in considerazione solo le caratteristiche essenziali delle superfici — le quali generalmente diventano evidenti anche a bassa risoluzione — è possibile esplorare in modo robusto (anche se non accurato) estese regioni dello spazio di ricerca. Partendo da una soluzione approssimata, il dettaglio — eventualmente a scale differenti — può essere preso in considerazione per trovare la soluzione ottimale. Al contrario, utilizzando superfici con molti dettagli a bassa scala, la soluzione trovata è costosa (più dettagli, più calcoli) e la procedura di ricerca è maggiormente soggetta a rimanere intrappolata nei minimi locali.

3.2.1. Formalizzazione della metodologia

Sia \mathcal{M} il paradigma utilizzato per la rappresentazione della superficie. Ai fini ed al livello di astrazione di questa trattazione, una istanza $M \in \mathcal{M}$ di questo paradigma può essere caratterizzata univocamente dall'insieme $\{\theta, K, P\}$, dove θ è un vettore di parametri che esprime le caratteristiche geometriche e fotometriche della superficie di M , K è una descrizione della topologia di M e P è la parametrizzazione della superficie. La procedura di generalizzazione di un insieme di dati S , formalizzabile come una funzione $\mathcal{G} : \mathcal{S} \rightarrow \mathcal{M}$, ha quindi lo scopo di scegliere, dato un paradigma \mathcal{M} l'elemento M_i più adatto alla rappresentazione della superficie dalla quale S è stato campionato. La scelta viene fatta, per una opportuna funzione \mathcal{D} che misura la distanza tra la superficie e i dati, individuando il modello ottimale:

$$\mathcal{G}(S) = \arg \min_{M_i \in \mathcal{M}} \mathcal{D}(M_i, S)$$

¹Questi approcci possono essere ulteriormente generalizzati prendendo in considerazione anche una retropropagazione alla fase di acquisizione delle informazioni risultanti dall'integrazione. La possibilità di utilizzare tali informazioni per diminuire l'incertezza presente sui dati è però fortemente dipendente dallo strumento di misura e dalle tecniche utilizzati per ottenere gli insiemi di dati relativi alle *viste*. L'analisi dettagliata di tale retropropagazione va oltre gli scopi della presente tesi.

3. Una metodologia gerarchica per la digitalizzazione

In teoria la ricerca dovrebbe estendersi su tutto l'insieme delle caratteristiche del modello, $\{\theta_i, K_i, P_i\}$. Ciò è però difficilmente realizzabile nella pratica, dove la topologia, K_i , e la parametrizzazione, P_i , vengono fornite a priori, limitando così la ricerca al vettore delle caratteristiche geometriche, θ_i . Solitamente, nella pratica, le informazioni di natura geometrica e quelle fotometriche vengono trattate in due momenti differenti. Per semplicità, le caratteristiche fotometriche vengono ricostruite dopo aver individuato le caratteristiche di superficie ottimali. Al livello di astrazione in cui si pone la presente trattazione, tuttavia, queste procedure possono essere come un'unica procedura di adattamento ai dati. La parametrizzazione, P_i , può poi essere adattata (*a posteriori*) sulla topologia e la geometria in base alle necessità della applicazione obiettivo. Inoltre, la ricerca è spesso vincolata dalla formalizzazione di alcune caratteristiche conosciute a priori (per esempio, grado di accuratezza, ingombro della rappresentazione). Una più realistica modellizzazione della procedura di generalizzazione è quindi una funzione $\mathcal{G} : \mathcal{S} \times \mathcal{K} \times \mathcal{P} \times \mathcal{V} \rightarrow \mathcal{M}$, dove \mathcal{K} e \mathcal{P} sono rispettivamente l'insieme delle topologie e delle parametrizzazioni² degli elementi di \mathcal{M} , e \mathcal{V} è la classe degli insiemi di vincoli che si possono imporre sui vettori delle caratteristiche geometriche θ_i .

La procedura di ricostruzione parte da una collezione di N^0 insiemi di dati $S^0 = \{S_1^0, \dots, S_{N^0}^0\}$ ed ottiene la sequenza di collezioni $\{S^j \mid j = 1, \dots, z\}$, composte ciascuna di N^j elementi, $N^{j+1} < N^j$, con $N^z = 1$. Al passo j -esimo della generazione del modello, vengono ricostruite le corrispondenti superfici $M^j = \{M_1^j, \dots, M_{N^j}^j\}$. Esse vengono utilizzate per la registrazione, la quale avviene partizionando l'insieme M^j in N^{j+1} gruppi di superfici che coprono regioni parzialmente sovrapposte dell'oggetto da ricostruire, e individuando una trasformazione che porti le loro rappresentazioni in un sistema di riferimento consistente con la realtà. Al termine della registrazione, quindi, ad ogni modello M_i^j è associata una trasformazione R_i^j che può essere applicata ai dati S_i^j . I dati appartenenti ad ogni partizione possono quindi essere fusi tra di loro, generando così l'insieme S^{j+1} . Per rendere più leggera la lettura, dove non vi siano possibili ambiguità, l'apice che indica il numero di passi di raffinamento eseguiti verrà tralasciato.

Per formalizzare in modo più chiaro le operazioni di integrazione, è necessario introdurre esplicitamente una notazione per descrivere la partizione. Essa verrà denotata come la collezione di indici che individuano

²La parametrizzazione è solitamente legata alla topologia, quindi la funzione \mathcal{G} può essere non suriettiva.

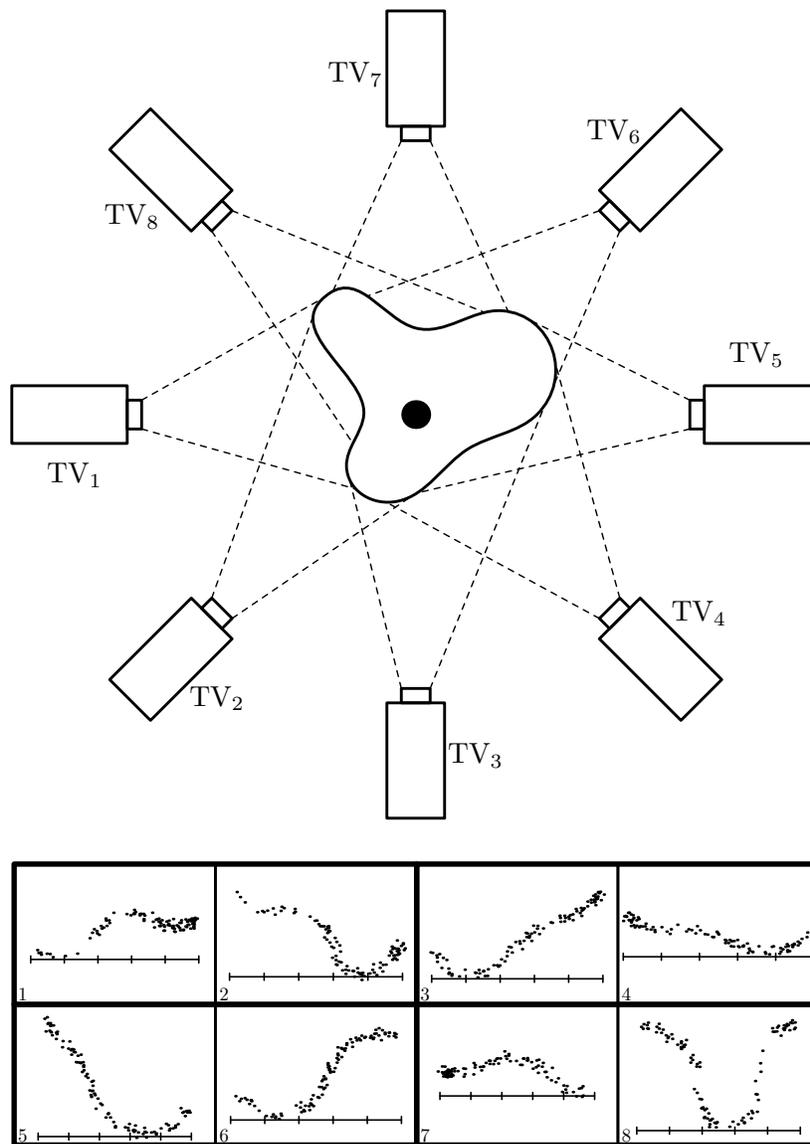


Figura 3.4.: Acquisizione multi-*vista*.

3. Una metodologia gerarchica per la digitalizzazione

i modelli dello stesso cluster:

$$\left\{ C_k^j \mid C_k^j \subseteq \{1, \dots, N^j\}, k = 1, \dots, N^{j+1} \right\}$$

tali per cui $\bigcup_k C_k = \{1, \dots, N\}$ e $C_k \cap C_h = \emptyset, \forall k, h \ k \neq h$. Le elaborazioni necessarie per l'integrazione dei dati vengono operate a livello di cluster. Dunque, al passo j , il k -esimo cluster è composto dai modelli $M_{C_k} = \{M_h \mid h \in C_k\}$ ed al termine del passo di integrazione, l'insieme di dati S_k^{j+i} sarà ottenuto come fusione della collezione di insiemi di dati $S_{C_k} = \{S_h^j \mid h \in C_k^j\}$. Per comodità di notazione definiamo, per ogni cluster C_k , una funzione enumerativa $c_k(\cdot)$, la quale enumera gli elementi dell'insieme di indici C_k . La collezione di modelli del k -esimo cluster è così esprimibile nel seguente modo:

$$M_{C_k} = \{M_{c_k(1)}, \dots, M_{c_k(|C_k|)}\}$$

dove $|C_k|$ indica la cardinalità dell'insieme C_k . Il partizionamento può essere effettuato utilizzando eventuali informazioni derivate dalla calibrazione (che potrebbe fornire la posizione dei sensori ed il campo di acquisizione rispetto ad un sistema di riferimento esterno) o dall'acquisizione (la posizione di marcatori speciali rispetto al sistema di riferimento dei sensori).

La registrazione consiste essenzialmente nella ricerca della trasformazioni affini (§6.2) da applicare ad ogni modello per far sì che le regioni comuni siano sovrapposte. Nei casi reali, i dati dai quali i modelli sono stati generati sono affetti da rumore e distorsioni che differiscono da *vista a vista*: anche nelle regioni comuni, quindi, i modelli differiscono tra loro. La sovrapposizione perfetta non si può perciò ottenere, e pertanto si cerca un insieme di trasformazioni tale da minimizzare la distanza delle superfici nelle regioni comuni. Questa ottimizzazione, generalmente, non viene effettuata utilizzando direttamente le superfici per via del costo computazionale che tale procedura richiederebbe e per la sua sensibilità ai minimi locali. Per questo motivo, la registrazione è generalmente composta di due passi:

- estrazione di *caratteristiche (features)*,
- allineamento.

La prima operazione ha per scopo la caratterizzazione delle regioni comuni alle superfici appartenenti allo stesso cluster. Le *caratteristiche* possono essere per esempio, punti o linee sulla superficie. L'importante

è che siano stimabili in modo robusto, perché da esse dipende l'affidabilità della registrazione. Esse vengono modellizzate come il risultato di una procedura:

$$f_i^j = \mathcal{E}(M_i^j).$$

Le *caratteristiche* sono definibili come proprietà dell'oggetto nelle regioni comuni del cluster, e il termine f_i identifica la loro rappresentazione nel modello M_i .

Le *caratteristiche* vengono utilizzate per calcolare le trasformazioni necessarie per allineare le superfici. La ricerca della trasformazione affine basata su poche *caratteristiche* affidabili è infatti molto più robusta ed economica. L'allineamento, per ogni cluster C_k , è quindi formalizzabile come:

$$[R_{c_k(1)} \cdots R_{c_k(|C_k|)}] = \arg \min_{[R_1 \cdots R_{|C_k|}]} Dist_f([R_1(f_{c_k(1)}) \cdots R_{|C_k|}(f_{c_k(|C_k|)})])$$

$[R_{c_k(1)} \cdots R_{c_k(|C_k|)}]$ identifica il vettore di trasformazioni di registrazione per i modelli appartenenti al cluster k -esimo. La funzione $Dist_f$ misura la differenza tra le trasformazioni subite dalle *caratteristiche* dei modelli dello stesso cluster.

Per rendere la soluzione univoca, in generale si può fissare una di queste trasformazioni come l'identità. Per esempio, senza perdere di generalità, si può fissare $R_{c_k(1)} = I$. Le altre trasformazioni, $\{R_{c_k(2)}, \dots, R_{c_k(|C_k|)}\}$, avranno quindi lo scopo di riferire i modelli ad esse corrispondenti al sistema di riferimento del primo modello del cluster (identificato da $c_k(1)$).

La minimizzazione della funzione $Dist$ non è facilmente ottenibile. Sebbene la trasformazione di registrazione sia essenzialmente una trasformazione lineare, il numero dei possibili abbinamenti delle *caratteristiche* rende improponibile una ricerca esaustiva dello spazio di ricerca. Per questo motivo si ricorre spesso a procedure iterative, come ad esempio ICP.

La registrazione viene quindi propagata agli insiemi di dati appartenenti al cluster, i quali vengono successivamente fusi tra loro per formare uno degli insiemi di dati del passo di integrazione successivo:

$$S_k^{j+1} = F(R_{C_k^j}(S_{C_k^j}^j))$$

Una definizione banale di F può essere l'operazione di unione insiemistica:

$$F(R_{C_k}(S_{C_k})) = \bigcup_{i \in C_k} R_{C_k}(S_i)$$

3. Una metodologia gerarchica per la digitalizzazione

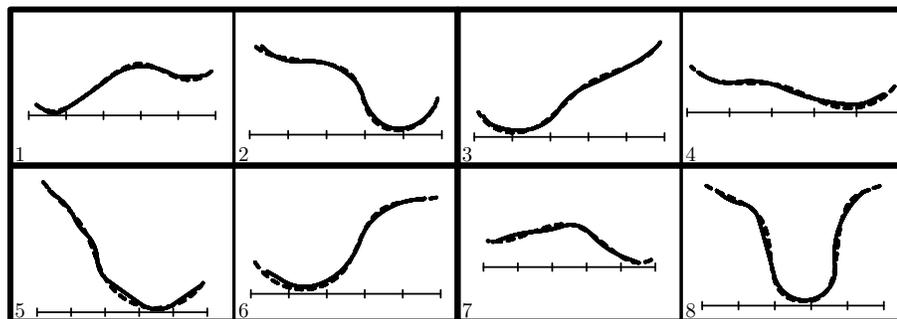


Figura 3.5.: I dati raccolti dalla sessione raffigurata in figura 3.4 vengono generalizzati (il vero profilo dell'oggetto è riportato tratteggiato).

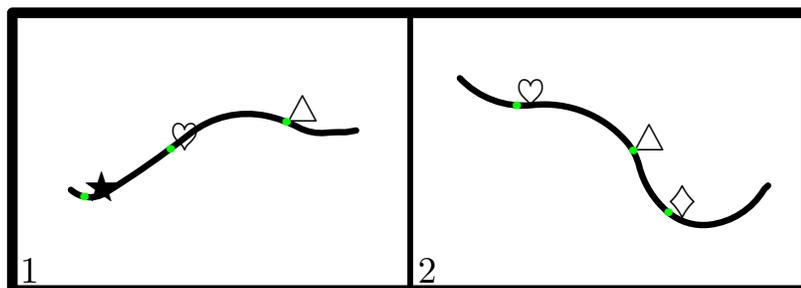


Figura 3.6.: Estrazione delle *caratteristiche*.

Ciò non tiene però conto delle caratteristiche dei sistemi di acquisizione: l'errore di misura non è uniforme, ma dipende dalle caratteristiche della superficie nel punto campionato. Per esempio, regioni quasi parallele alla direzione di acquisizione hanno una incertezza nella misura maggiore di quelle perpendicolari al sensore. La semplice unione degli insiemi di dati potrebbe quindi causare la presenza, nella stessa regione di spazio, di dati con differente grado di affidabilità. Questo non è un problema, se l'operatore di generalizzazione è in grado di trattare dati considerandone anche il grado di affidabilità.

Una soluzione più generale può includere nella fase di fusione l'omogeneizzazione dei dati delle regioni comuni, utilizzando i dati più affidabili per correggere quelli affetti da una maggiore incertezza. Un operatore di fusione con queste proprietà deve quindi essere costruito sulle caratteristiche dei dati appartenenti al cluster. Ciò può essere realizzato sfruttando la descrizione della superficie dell'oggetto fornita dal modello.

L'operazione di fusione può quindi essere prima effettuata sui modelli dello stesso cluster ($F_h(R_h(M_h)), \forall h \in C_k$) e poi retropropagata agli insiemi di dati corrispondenti ($F_h(R_h(S_h)), \forall h \in C_k$). Ciò equivale nei fatti alla costruzione di un nuovo modello. In particolare, è richiesta la fusione delle informazioni topologiche e di parametrizzazione. In generale, tuttavia, il modello risultante dalla fusione dei modelli di un cluster, $F_{C_k}^j(M_{C_k}^j)$ sarà diverso dal modello ottenuto generalizzando il risultato della fusione dei dati del cluster considerato, $M_k^{j+1} = \mathcal{G}(S_{C_k}^j)$.

Inoltre, come è stato detto in precedenza, solitamente la procedura di generalizzazione richiede una topologia sulla quale costruire il modello. Sebbene, come è stato illustrato in §2.4, esistano metodi per la ricostruzione della topologia a partire da informazioni geometriche, la conoscenza accumulata nel precedente passo di integrazione può essere sfruttata, migliorando così l'efficienza della metodologia. La fusione dei modelli al passo j richiede (o induce) la fusione della topologia dei modelli del cluster. Essa può essere utilizzata per la fase di generalizzazione del passo $j+1$. La topologia del passo 0 può essere omeomorfa al piano, così come la parametrizzazione. Dopo la fusione, si ottiene una nuova topologia e una nuova parametrizzazione, che possono essere utilizzate come topologia e parametrizzazione del passo 1. In generale, quindi:

$$K_k^{j+1} = F_k^j(K_{C_k}^j) \quad \text{e} \quad P_k^{j+1} = F_k^j(P_{C_k}^j).$$

3.2.2. Utilizzo dei paradigmi gerarchici

I paradigmi gerarchici (§2.4.2) si contraddistinguono per:

- multirisoluzione
- località spaziale

La multirisoluzione permette di descrivere le caratteristiche geometriche della superficie a diversi livelli di dettaglio. In un'ottica di teoria del segnale questa proprietà può essere vista come località nel dominio delle frequenze. La località spaziale consente di limitare l'utilizzo di risorse computazionali solo nelle regioni dove vi è effettivamente informazione.

Queste caratteristiche possono essere sfruttate in diverse fasi della metodologia introdotta nel precedente paragrafo.

L'uso più immediato è durante la fase di generalizzazione: la rappresentazione della geometria di una superficie è l'uso più naturale per questi paradigmi. La località spaziale consente di configurare il modello geometrico in modo economico in quanto coinvolge solo una parte dell'insieme di dati. La multirisoluzione consente di limitare l'ottimizzazione al grado di approssimazione necessario per le elaborazioni successive. Inoltre, in generale, l'utilizzo di una struttura gerarchica consente di limitare l'uso di risorse computazionali.

La loro capacità di decorrelare l'informazione a differenti livelli dettaglio fornisce uno strumento sia per limitare l'utilizzo delle risorse computazionali, che per disporre di informazioni già caratterizzate da una finestra spazio-frequenza. Questa particolarità può essere sfruttata per la registrazione, sia durante la fase di estrazione delle *caratteristiche* che durante l'allineamento. Le *caratteristiche* sono generalmente legate alle proprietà delle derivate della superficie (per esempio, spigoli) le quali vengono in genere facilmente individuate a partire dalla rappresentazione multirisoluzione. Appare, inoltre, evidente che prima di cercare una correlazione tra le caratteristiche a scala bassa dei modelli dello stesso cluster, è necessario cercarla fra le caratteristiche a larga scala.

La caratteristica di località spaziale può inoltre essere sfruttata al momento della fusione: solo le unità nelle regioni comuni saranno interessate dalle modifiche.

La rappresentazione multirisoluzione può essere sfruttata anche per un raffinamento della topologia. I dettagli sono ovviamente legati al particolare paradigma, \mathcal{M} , scelto per la rappresentazione dell'oggetto acquisito. Tuttavia, una rappresentazione a bassa risoluzione può essere utilizzata come descrizione della topologia dei modelli a risoluzione più elevata.

3.2.3. Alcuni problemi nell'acquisizione dei dati

La metodologia esposta poco sopra non considera esplicitamente alcuni problemi causati dalla forma che un generico oggetto reale può avere. La rilevazione e la soluzione di questi problemi può tuttavia essere inclusa nei passi di cui si compone la metodologia.

Può capitare che alcune parti dell'oggetto ne nascondano altre alla visione dei sensori (figura 3.10). Ciò può causare due situazioni:

- ambiguità,
- dati mancanti.

I dati si dicono *ambigui* quando portano a conclusioni discordanti. Nel contesto considerato, le ricostruzioni effettuate sulle singole viste sono geometricamente discordanti. All'interno della metodologia proposta, le ambiguità possono essere rilevate in fase di registrazione, dove può verificarsi che in alcune regioni vi siano corrispondenze, mentre in altre vi siano differenze sostanziali. Una delle ipotesi sulle quali si basano le ricostruzioni è che le *viste* siano composte da dati spazialmente contigui, il che permette alla generalizzazione di descrivere una superficie continua. Questa ipotesi viene chiaramente a cadere in caso di occlusioni in acquisizione.

I dati mancanti sono causati dall'impossibilità dei sensori di raccogliere informazioni in alcune regioni della superficie dell'oggetto. Essi possono essere il sintomo di un cattivo *view-planning*, ma possono essere anche dovuti alla effettiva inaccessibilità della superficie. In tal caso, il comportamento della metodologia è guidato dalla strategia di individuazione della topologia e dalla capacità della procedura di generalizzazione del paradigma prescelto di sopperire alla mancanza di informazioni in una regione più o meno estesa (figura 3.11).

Per correggere questa situazione, è necessario introdurre una fase di *segmentazione* dei dati delle *viste*. In genere, in questa fase possono essere utilizzate conoscenze e informazioni ausiliarie derivanti dal sistema di acquisizione utilizzato. Per esempio, analisi della pendenza [PST00] [HJBJ⁺96]. Se nella fase di registrazione si rileva un'ambiguità, bisogna quindi annullare alcune operazioni (*roll-back*) e modificare le condizioni di partenza. In questo caso, si annulla la ricostruzione (ed eventualmente il clustering) e si torna alle collezioni di dati delle *viste*. Su di esse si opera la segmentazione (utilizzando le informazioni derivate dalla registrazione fallita), suddividendo le viste che erano causa di ambiguità in due o più (sotto) *viste* (figura 3.12).

3.3. Contributo della tesi

Nell'ambito di questa tesi, la ricerca è stata svolta, sia in campo teorico che applicativo, con l'intento di sviluppare una metodologia che organicamente incorpora in modo innovativo i passi atti ad ottenere la riproduzione digitale di oggetti reali a partire da un campionamento della superficie. In particolare, lo sforzo è stato indirizzato nella soluzione del problema della ricostruzione delle caratteristiche di una superficie a partire da dati campionati su di essa.

In quest'ottica va considerata la metodologia esposta in §3.2. Essa è stata sviluppata ad alto livello di astrazione al fine di fornire un approccio strutturato, ma fortemente modulare e flessibile, capace di adattarsi efficacemente a numerose condizioni operative, includendo sinergicamente e organicamente non solo soluzioni innovative, ma anche la maggior parte delle tecniche presentate in letteratura per trattare i singoli aspetti del problema (capitolo 2). Il procedimento è stato suddiviso in diverse fasi, ognuna delle quali caratterizzata dall'elaborazione di un modello tridimensionale scelto all'interno di un paradigma di rappresentazione. Le fasi sono tra loro differenziate in base alle elaborazioni operate sui modelli. Il paradigma di rappresentazione non è specificato, ma deve essere in grado di descrivere le caratteristiche della superficie di un oggetto reale. Diversi paradigmi sono stati presentati nel capitolo 2. Sebbene quasi tutti abbiano le capacità descrittive richieste dalla metodologia individuata, appare evidente (§3.2.2) che i paradigmi gerarchici possano fornire un supporto più elegante alle diverse fasi elaborative.

Allo scopo è stato sviluppato il paradigma HRBF (accennato in §2.4.2, verrà illustrato nei dettagli nel capitolo 4), indagando sulle sue proprietà come approssimatore di funzioni in termini di accuratezza e robustezza della ricostruzione. Ad una dimostrazione teorica delle qualità citate, sono state fatte seguire prove sperimentali su insiemi di dati sia sintetici che reali. HRBF è stato utilizzato come paradigma ricostruttivo per il modulo di ricostruzione di uno scanner digitale sperimentale, Autoscan (§5). I vantaggi della multirisoluzione prodotta da HRBF sono stati illustrati all'interno del sistema Autoscan per la digitalizzazione di volti umani. Autoscan rappresenta una istanza realizzativa della metodologia illustrata precedentemente. La sua struttura modulare rispecchia la suddivisione operativa e la relazione tra i vari passi proposta nella metodologia.

HRBF è stato confrontato con il paradigma gerarchico più diffuso, MRA, basato sulla trasformata wavelet. Il confronto ha riguardato sia gli aspetti teorici, che quelli pratici. A tale scopo, sulla scia di quanto esistente per l'elaborazione mediante MRA, HRBF è stato riformulato

in chiave digitale. Ciò da un lato ha permesso un confronto più puntuale con MRA, dall'altro ha messo in evidenza la possibilità di una implementazione veloce e robusta eventualmente utilizzando componenti hardware standard.

Il contributo innovativo della presente tesi può quindi essere riassunto nei seguenti punti:

- definizione di una metodologia modulare e strutturata per la generazione di modelli tridimensionali;
- ideazione e valutazione di HRBF come paradigma multirisoluzione per la ricostruzione di superfici a partire da dati campionati affetti da rumore;
- comparazione con MRA;
- riformulazione degli algoritmi HRBF in chiave digitale;
- contributo rilevante alla creazione del modulo di ricostruzione di Autoscan nonché valutazione delle prestazioni di accuratezza nel campionamento di superfici.

Metodologia In questa tesi è stata definita una metodologia modulare e strutturata capace di inglobare organicamente i passi necessari ad affrontare specifiche condizioni operative delle applicazioni.

Valutazione di HRBF come paradigma multirisoluzione Una parte cospicua del lavoro di questa tesi è consistita nella comparazione tra le proprietà delle HRBF con quelle delle Analisi Multirisoluzione (MRA), il paradigma multirisoluzione più diffusamente utilizzato (un breve accenno in §2.4.2, una trattazione più estesa nel capitolo 4). Il loro utilizzo spazia, infatti, in diverse discipline e campi applicativi. In [BMF99] sono state esplorate le caratteristiche di approssimazione e di convergenza delle HRBF e confrontate con quelle delle MRA. È possibile dimostrare che, sebbene le HRBF non possano essere assimilate alle MRA, con esse condividono il fatto di realizzare una base di Riesz per lo spazio delle approssimazioni e possiedono le stesse proprietà di approssimazione asintotica. Inoltre, raggiungono più facilmente un desiderato grado di approssimazione.

Una delle ragioni della diffusione delle MRA è la possibilità di ricavare la codifica e di ricostruire un insieme di dati in modo rapido ed efficiente mediante un algoritmo chiamato *Fast Wavelet Transform* (FWT). Esso permette una implementazione mediante filtri FIR digitali

3. Una metodologia gerarchica per la digitalizzazione

che ne rendono possibile la realizzazione hardware utilizzando componenti poco complessi. Il campo di applicazione è limitato al caso di dati equispaziati, ma questo non rappresenta una limitazione nei casi reali.

Riformulazione HRBF in chiave digitale In [FBP00] è stata operata una riformulazione delle HRBF in chiave digitale, la quale è stata utilizzata per un confronto con la FWT.

Sebbene il filtro gaussiano non sia un filtro FIR, esso può esservi approssimato trascurando gli elementi inferiori ad una certa soglia. Poiché il valore della gaussiana decresce in modo esponenziale, l'approssimazione è generalmente accettabile. Allo scopo, le HRBF sono state riformulate per il caso di dati equispaziati.

Allo scopo di compararne le prestazioni in termini di robustezza, le grandezze coinvolte nella ricostruzione sono state perturbate simulandone la rappresentazione mediante notazioni a virgola fissa e mobile a diverse precisioni. La comparazione ha mostrato che i due paradigmi hanno compartimenti simili per precisioni elevate, ma che le HRBF resistono meglio a perturbazioni consistenti.

Valutazione delle procedure di acquisizione di Autoscan In [BF00] si è caratterizzato il rumore di misura di Autoscan. L'accuratezza della acquisizione è stata determinata sperimentalmente, utilizzando un metodo simile a quello riportato in [GR99]. L'esperimento è costituito da una serie di acquisizioni di un piano posizionato al centro del campo di visione delle TV-camere. Per ogni acquisizione è stata stimato il piano ottimale e calcolata la distanza dai punti acquisiti da esso. L'accuratezza è stata misurata mediante l'errore quadratico medio. Questa cifra di merito varia linearmente con le dimensioni del volume, ed è consistente con le misure di accuratezza riportate nel campo dell'analisi del movimento [BC00]. La distribuzione dell'errore di misura è risultata normale ed a media nulla.

Ricostruzione di superficie mediante HRBF In [BF00] si inoltre è mostrata la validità delle HRBF [BF98] nel campo della ricostruzione di superfici a partire da punti campionati. Esse sono state applicate al caso pratico della ricostruzione di punti campionati su un volto umano.

Dato il set-up utilizzato (due TV-camere) e considerato il tipo di superficie da ricostruire (un volto umano), la superficie da ricostruire può essere formalizzata come una funzione $\mathbb{R}^2 \rightarrow \mathbb{R}$ (§6).

Autoscan è stato spesso utilizzato per la ricostruzione di parti del corpo umano, volti in particolare. Quest'ultimo caso rappresenta un

compito particolarmente difficile, in quanto il contenuto in frequenza è altamente variabile nelle diverse regioni del viso.

Le HRBF ben si adattano a questo compito. Esse, infatti, realizzano una ricostruzione multirisoluzione mediante strati di unità gaussiane a differente ampiezza. Le gaussiane ad ampiezza maggiore forniscono una approssimazione delle regioni a basso contenuto di frequenza, mentre quelli con gaussiane più strette ricostruiscono il dettaglio più fine.

3.4. Conclusioni

In questo capitolo sono state rielaborate ad alto livello di astrazione le attività e le tecniche utilizzate in letteratura per operare la digitalizzazione di oggetti reali, le quali erano state esposte nel capitolo precedente. Esse sono state organizzate e strutturate in uno schema generale che riassume e formalizza la maggior parte degli approcci in letteratura. Particolare attenzione è stata prestata alla formalizzazione di una metodologia per la fase di ricostruzione.

Nei prossimi capitolo saranno illustrati gli altri contributi innovativi sviluppati durante questa tesi ed i risultati ottenuti.

Il capitolo 4 è dedicato ai paradigmi gerarchici utilizzati per la ricostruzione: HRBF — paradigma ideato e sviluppato nel contesto di questa tesi — ed MRA — il paradigma gerarchico più diffuso in letteratura —. Saranno illustrati i concetti ai quali sono ispirati e le proprietà di approssimazione di cui godono. L'inquadramento teorico fornirà una base per il confronto delle prestazioni dei due paradigmi impiegati nel ruolo di approssimatori di funzioni da dati sparsi.

Nel capitolo 5 sarà invece descritto lo scanner tridimensionale Autoscan. Il suo flusso di elaborazione e la sua struttura modulare rispecchiano le fasi individuate nella metodologia proposta in questo capitolo, della quale Autoscan è una istanza implementativa. Il modulo di ricostruzione di Autoscan utilizza HRBF come paradigma di rappresentazione del modello tridimensionale.

Infine, il capitolo 6 riassumerà le idee e i risultati esposti per tracciare le conclusioni e le possibili linee di ricerca future.

3. Una metodologia gerarchica per la digitalizzazione

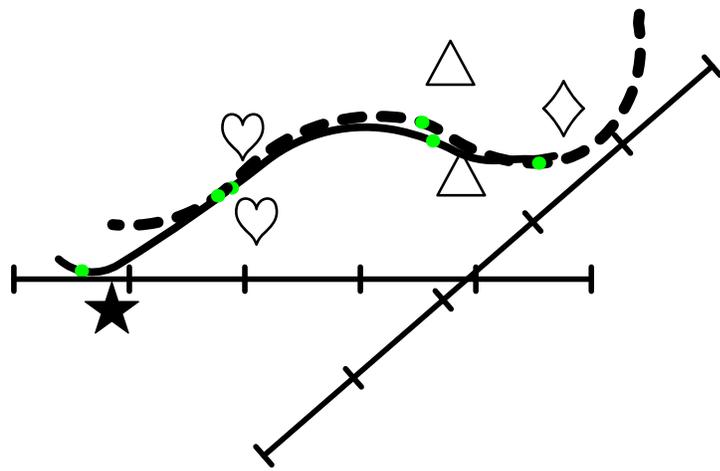


Figura 3.7.: La registrazione viene effettuata allineando le *caratteristiche*.

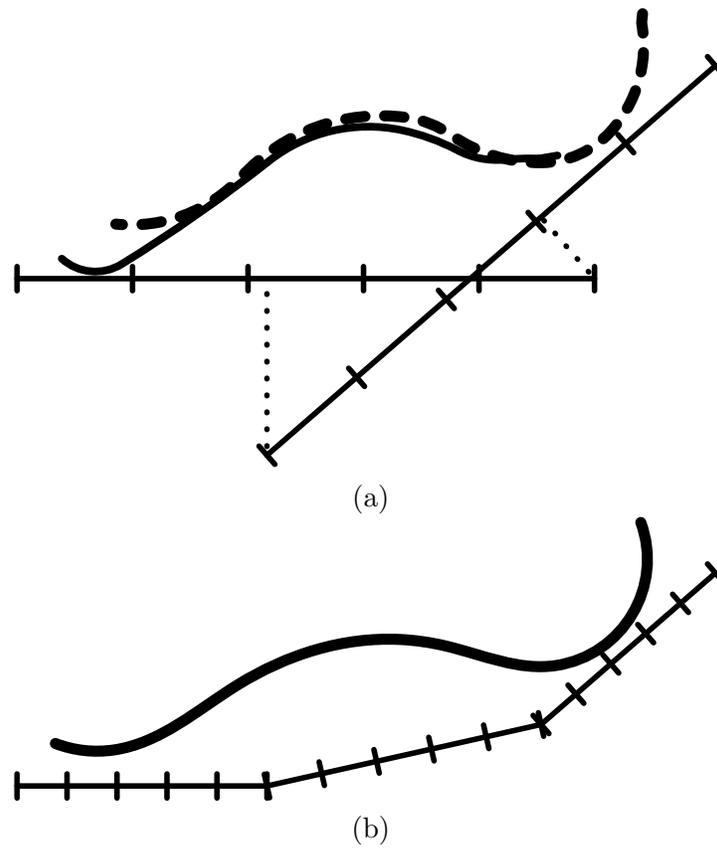


Figura 3.8.: Le ricostruzioni dello stesso cluster, allineate (a), vengono integrate tra loro (b).

3. Una metodologia gerarchica per la digitalizzazione

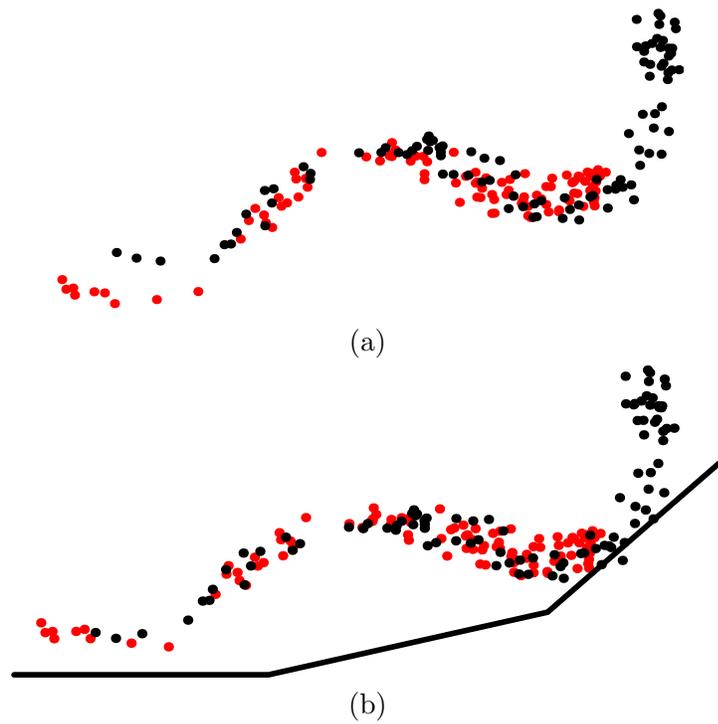


Figura 3.9.: Le trasformazioni di integrazione vengono propagate ai dati.

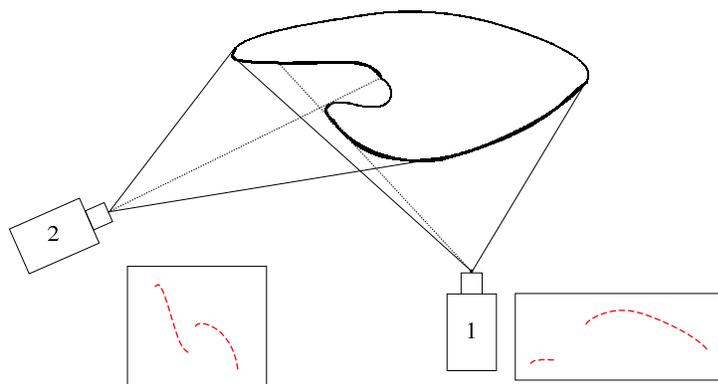


Figura 3.10.: La parte visibile dell'oggetto può nascondere altre regioni.

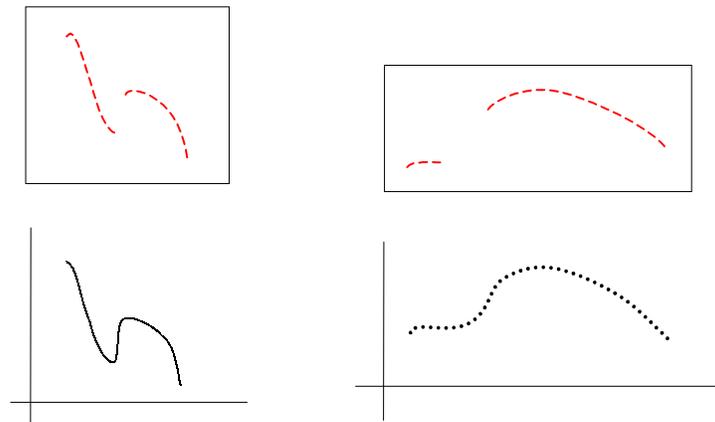


Figura 3.11.: La ricostruzione in regioni dove i dati siano mancanti è affidata alle proprietà di generalizzazione del paradigma scelto per la ricostruzione.

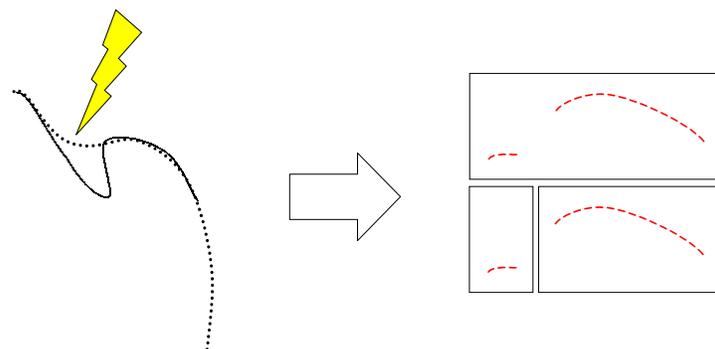


Figura 3.12.: Dati non contigui devono essere separati(segmentazione).

3. *Una metodologia gerarchica per la digitalizzazione*

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

La metodologia descritta nel capitolo 3 può avvantaggiarsi dell'utilizzo di un paradigma multirisoluzione di rappresentazione della ricostruzione.

I paradigmi gerarchici che permettono la ricostruzione e la rappresentazione multirisoluzione delle superfici sono raggruppabili in due classi, a seconda del loro modo di operare la multirisoluzione. Alcuni paradigmi operano prima sulle informazioni a larga scala (bassa frequenza) dei dati a disposizione, per poi raffinare la descrizione della superficie aggiungendo le descrizioni dell'andamento della superficie a scale via via più piccole. Questa modalità viene detta *coarse to fine*, e viene contrapposta a quella detta *fine to coarse* che caratterizza i paradigmi che operano estraendo prima il dettaglio e procedendo quindi a determinare le approssimazioni a scala via via più ampia.

Alla prima classe, *coarse to fine*, appartengono paradigmi quali *Hierarchical RBF* (HRBF) [BF98] o *Multilevel B-spline* [LWS97], mentre wavelet [Mal89] [EDD⁺95] o *Hierarchical Spline* [FW98] fanno generalmente uso di tecniche *fine to coarse*.

La possibilità di operare in modalità *coarse to fine* può rendere meno computazionalmente pesanti le operazioni nelle situazioni in cui l'informazione contenuta nel dettaglio possa essere trascurata.

Il paradigma HRBF, ideato e sviluppato nell'ambito di questa tesi, possiede queste caratteristiche. Le MRA sono state largamente studiate in letteratura e sono state utilizzate per molte applicazioni.

In questo capitolo le caratteristiche di questi due paradigmi gerarchici saranno analizzate. Nella loro formulazione originale il supporto utilizzato per l'espressione della funzione ricostruita (ciò che nel paragrafo §2.2.1 era stato chiamato spazio parametrico \mathbb{U}) è il piano XY. È possibile estendere la trattazione a domini più generici, ma ciò richiede ipotesi aggiuntive sulla definizione di tali domini.

Inoltre, si evidenzieranno i particolari teorici e applicativi che accomunano e differenziano le due tecniche di ricostruzione. Si discuteranno infine i risultati delle prove effettuate per saggiare le capacità di

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

ricostruzione in termini di accuratezza e robustezza in diverse condizioni.

Nel seguito i simboli in grassetto verranno utilizzati per indicare vettori e sequenze. Inoltre, con la notazione $\langle \cdot, \cdot \rangle$ si indicherà il prodotto scalare di due elementi di uno spazio vettoriale. La presenza di un solo termine tra le parentesi angolari, $\langle \cdot \rangle$, indicherà, invece, lo spazio generato dalla base indicata (*span*).

4.1. Le Wavelet

La teoria delle wavelet è stata derivata utilizzando conoscenze tipiche di differenti campi di ricerca: approssimazione di funzioni, computer vision, teoria del segnale, solo per citare i principali. I concetti che stanno alla base di questa teoria possono quindi essere trattati da diversi punti di vista, dando risalto alle diverse caratteristiche, e partendo da diversi presupposti. Una completa trattazione della teoria delle wavelet è al di fuori dello scopo di questo tesi. Per una descrizione più dettagliata si vedano [JS94] [Cha95] [Dau92] [Mal98] [SN96]. In [Dau96] la derivazione della teoria è trattata in modo meno formale. Una versione estesa della trattazione nelle sezioni seguenti si può trovare in [Fer99].

4.1.1. La Trasformata Wavelet Continua

La trasformata wavelet è una rappresentazione di una funzione $f(x) \in L^2(\mathbb{R})$ mediante una combinazione lineare di funzioni base. Le funzioni base, dette *wavelet*, sono generate traslando e scalando in ampiezza un'unica funzione detta *mother wavelet* $\psi(x) \in L^2(\mathbb{R})$. La forma delle wavelet $\psi_{a,b}(x)$ dipende da due parametri reali a e b . Essi rappresentano la scala e la posizione della wavelet $\psi_{a,b}(x)$ e possono variare in modo continuo. Formalmente, la generica wavelet alla scala a e nella posizione b , è espressa come:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) \quad \text{con } a, b \in \mathbb{R}, a \neq 0 \quad (4.1)$$

Si definisce trasformata wavelet continua (CWT) di una funzione $f(x) \in L^2(\mathbb{R})$:

$$W(a, b) = \langle f, \psi_{a,b} \rangle \quad (4.2)$$

L'ampiezza di $W(a, b)$ indica quanto la funzione $f(x)$ assomiglia alla wavelet $\psi_{a,b}(x)$. La trasformata di Fourier di $\psi_{a,b}(x)$ vale:

$$\hat{\psi}_{a,b}(\omega) = \frac{a}{\sqrt{|a|}} e^{-i\omega b} \hat{\psi}(a\omega) \quad (4.3)$$

La trasformata wavelet ha una alta risoluzione in frequenza ed una bassa risoluzione nel dominio dello spazio per alti valori del parametro di scala, a , e, viceversa, una bassa risoluzione in frequenza ed una elevata risoluzione nel tempo per bassi valori di a . Questa caratteristica è molto utile nella pratica, dove spesso gli eventi di breve durata sono caratterizzati da segnali con componenti ad alta frequenza, mentre gli eventi di lunga durata sono caratterizzati da segnali a bassa frequenza.

Se la wavelet $\psi(x)$ soddisfa la condizione di ammissibilità:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty \quad (4.4)$$

allora la CWT $W(a, b)$ è invertibile sul suo dominio, e la sua trasformata inversa è data dalla relazione:

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(a, b) \psi_{a,b}(x) \frac{da db}{2} \quad (4.5)$$

È quindi possibile ricostruire una funzione $f(x)$ a partire dall'insieme (infinito) dei coefficienti della trasformata wavelet, $W(a, b)$. Dalla condizione di ammissibilità si può dedurre che $\hat{\psi}(0) = 0$, e, quindi, $\psi(x)$ deve oscillare. L'attribuzione del nome wavelet (“ondina”) alla trasformata descritta, è da ascrivere a questa proprietà ed alla limitazione nel dominio dello spazio e della frequenza.

Campionamento diadico

La possibilità di ricostruire una funzione continua a partire da un insieme di campioni discreti è regolata dal teorema di Nyquist, secondo il quale, la frequenza di campionamento non può essere inferiore del doppio della massima frequenza contenuta nella funzione campionata. Il contenuto in frequenza della $W(a, b)$ diminuisce all'aumentare del parametro di scala, a . La frequenza di campionamento può quindi variare al variare della scala. La scelta più comune è il campionamento in corrispondenza della griglia diadica: il parametro a assume i valori delle potenze di due e la traslazione viene effettuata con un passo pari ad a . In altre parole, si pone $a = 2^{-j}$ e $b = ak$, con $j, k \in \mathbb{Z}$ (cf. [SN96]).

In generale, minore è il numero di valori di a e b che si vogliono utilizzare, maggiori sono le restrizioni che bisogna porre sulla wavelet perché quest'ultima possa generare un insieme di funzioni base che siano in grado di ricostruire una qualsiasi funzione appartenente ad $L^2(\mathbb{R})$.

4.1.2. La Trasformata Wavelet Discreta

La trasformazione che usa valori diadici è detta trasformazione wavelet discreta (DWT). La teoria della DWT è stata particolarmente sviluppata in questi ultimi anni, in quanto essa rende possibile il calcolo della trasformata wavelet senza necessità di ricorrere alle approssimazioni insite nella discretizzazione della CWT. Le funzioni base sono sempre continue (con eventualmente un insieme numerabile di discontinuità), ma la loro scala e traslazione variano in modo discreto.

Analisi Multirisoluzione

La teoria dell'Analisi Multirisoluzione [Mal89] consente di ricavare un algoritmo veloce per il calcolo della trasformata wavelet discreta, che in seguito sarà chiamata Fast Wavelet Transform (FWT). Allo scopo sarà introdotta nella teoria wavelet una famiglia di funzioni chiamate *scaling function*. Una analisi multirisoluzione definisce, per ogni funzione $f(x)$ appartenente ad $L^2(\mathbb{R})$, una successione di funzioni approssimanti $\{P_j[f(x)]\}_{j \in \mathbb{Z}}$ che converge alla funzione stessa. La funzione $P_j[f(x)]$ è una combinazione lineare di funzioni ottenute da un'unica funzione, detta *scaling function*, per traslazione. La forma della scaling function è controllata da un parametro di scala che caratterizza la risoluzione dell'approssimazione.

Scaling function e sottospazi V_j

Riprendiamo i concetti sopra esposti in modo più rigoroso.

Una Analisi Multirisoluzione (MRA) di $L^2(\mathbb{R})$ è definita come la sequenza dei sottospazi chiusi $V_j \in L^2(\mathbb{R})$, $j \in \mathbb{Z}$, con le seguenti proprietà:

1. $V_j \subset V_{j+1}$
2. $v(x) \in V_j \Leftrightarrow v(2x) \in V_{j+1}$
3. $v(x) \in V_0 \Leftrightarrow v(x+1) \in V_0$
4. $\bigcup_{j=-\infty}^{\infty} V_j$ è denso in $L^2(\mathbb{R})$ e $\bigcap_{j=-\infty}^{\infty} V_j = \{0\}$
5. Esiste una funzione $\varphi(x) \in V_0$, con integrale non nullo, tale che l'insieme $\{(\varphi(x-k)|k \in \mathbb{Z})\}$ è una base di Riesz per V_0 . $\varphi(x)$ è detta *scaling function*.

Poiché $\varphi(x) \in V_0 \subset V_1$, si dimostra che esiste una sequenza $\{h_k\} \in l^2(\mathbb{Z})$ per cui la scaling function soddisfa [Mal89]:

$$\varphi(x) = 2 \sum_k h_k \varphi(2x - k) \quad (4.6)$$

La (4.6) è nota come *refinement equation* (o anche *dilation equation* o *two-scale difference equation*) ed esprime il fatto che le funzioni base di V_0 possono essere descritte mediante una combinazione lineare di elementi di V_1 . La (4.6) può essere generalizzata in modo da esprimere la stessa relazione tra le funzioni base di due generici spazi V_j e V_{j+1} . Inoltre, si dimostra che, per ogni $j \in \mathbb{Z}$, l'insieme di funzioni:

$$\{\varphi_{j,k}(x) \mid k \in \mathbb{Z}\} \quad \text{con } \varphi_{j,k}(x) = \sqrt{2^j} \varphi(2^j x - k) \quad (4.7)$$

è una base di Riesz per V_j . Risulta quindi che l'insieme $\{\varphi_{j,k}(x) \mid j, k \in \mathbb{Z}\}$ è una base di Riesz per $L^2(\mathbb{R})$.

Le caratteristiche degli spazi V_j possono essere sfruttate per costruire approssimatori di funzioni. Ciò viene realizzato definendo appropriate proiezioni su questi spazi. Poiché l'unione di tutti i V_j è densa in $L^2(\mathbb{R})$ (con un abuso di notazione si potrebbe scrivere $V_\infty \equiv L^2(\mathbb{R})$), ogni data funzione $f(x)$ appartenente a questo spazio può essere approssimata con arbitraria precisione dalle proiezioni $P_j[f(x)]$. In altri termini:

$$\forall f(x) \in L^2(\mathbb{R}), \lim_{j \rightarrow \infty} P_j[f(x)] = f(x) \quad (4.8)$$

Poiché $P_j[f(x)]$ è un elemento di V_j , esso può essere scritto come combinazione lineare delle basi di V_j :

$$\boxed{P_j[f(x)] = \sum_k \lambda_{j,k} \varphi_{j,k}(x)} \quad (4.9)$$

con appropriati coefficienti $\{\lambda_{j,k}\}$.

Funzione Wavelet e spazi dettaglio W_j

Dato che lo spazio V_{j+1} contiene tutte le funzioni dello spazio V_j , è sensato immaginarsi che l'approssimazione di una funzione $f(x)$ in V_{j+1} , $P_{j+1}[f(x)]$, sia più dettagliata di quella nello spazio V_j , cioè sia più vicina ad $f(x)$.

Chiamiamo W_j lo spazio complementare a V_j in V_{j+1} , cioè lo spazio che soddisfa:

$$V_{j+1} = V_j \oplus W_j = \{v_j + w_j \mid v_j \in V_j, w_j \in W_j\}$$

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

dove il simbolo \oplus indica l'operatore di somma diretta. In altre parole, ogni elemento di V_{j+1} può essere scomposto univocamente come la somma di un elemento di V_j (approssimazione) e di un elemento di W_j (dettaglio) (figura 4.1).

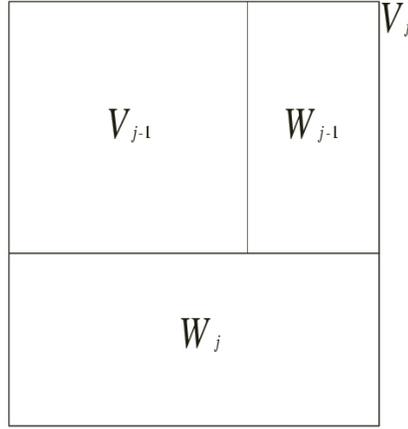


Figura 4.1.: Annidamento degli spazi approssimazione (V_j) e dettaglio (W_j).

Lo spazio W_j contiene le informazioni “dettaglio” necessarie per passare da un'approssimazione a risoluzione j ad una a risoluzione $j + 1$. Conseguentemente:

$$\bigoplus_j W_j = L^2(\mathbb{R})$$

Una funzione $\psi(x)$ è una wavelet se l'insieme delle funzioni $\{\psi(x - k) \mid k \in \mathbb{Z}\}$ è una base di Riesz per W_0 . La wavelet quindi genera una base per descrivere il dettaglio. Si può dimostrare che l'insieme delle funzioni $\{\psi_{j,k}(x) \mid j, k \in \mathbb{Z}\}$ tale che:

$$\psi_{j,k}(x) = \sqrt{2^j} \psi(2^j x - k) \quad (4.10)$$

è una base di Riesz per $L^2(\mathbb{R})$. La (4.10) corrisponde alla (4.1) qualora i parametri a e b giacciono in corrispondenza della griglia diadica, cioè quando $a = 2^{-j}$ e $b = ak$ per $j, k \in \mathbb{Z}$.

Dato che $V_1 = V_0 \oplus W_0$ e $W_0 \subseteq V_1$, la wavelet $\psi(x)$ sarà un elemento di V_1 . Esiste allora una sequenza $\{g_k\} \in l^2(\mathbb{Z})$ tale che:

$$\psi_{0,0}(x) = \psi(x) = 2 \sum_k g_k \varphi(2x - k) \quad (4.11)$$

$\psi_{0,0}$ è detta mother wavelet.

Tutti gli spazi V_j e W_j hanno complemento in $L^2(\mathbb{R})$ indicati rispettivamente con V_j^c e W_j^c . Si ha che:

$$V_j^c = \bigoplus_{i=j}^{\infty} W_i \quad \text{e} \quad W_j^c = \bigoplus_{i \neq j}^{\infty} W_i$$

Definendo Q_j l'operatore proiezione su W_j parallelo a W_j^c , una qualsiasi funzione $f(x)$ appartenente ad $L^2(\mathbb{R})$ può essere scritta come:

$$f(x) = \sum_j Q_j[f(x)] = \sum_{j,k} \gamma_{j,k} \psi_{j,k}(x) \quad (4.12)$$

Questa equazione (da confrontare con le (4.5), (4.8) e (4.9)) esprime una trasformata wavelet discreta inversa, per cui la funzione $f(x)$ si può ottenere come combinazione lineare di funzioni dettaglio. Resta aperto il problema di *trovare in maniera efficiente i coefficienti* $\{\gamma_{j,k}\}$. La facilità con la quale questo compito può essere realizzato dipende dalle caratteristiche della MRA, cioè delle basi utilizzate. Nelle prossime sezioni si descriveranno classi di basi che consentono di utilizzare un algoritmo efficiente per il calcolo dei coefficienti $\{\gamma_{j,k}\}$.

Wavelet Ortogonali

La scelta di una base ortogonale per rappresentare una funzione è particolarmente interessante in quanto permette di scomporre una funzione in elementi tra essi scorrelati. Dalle proprietà della MRA, descritte nel paragrafo 4.1.2, si può dimostrare che le funzioni definite in (4.7) o sono ortonormali, o comunque possono essere utilizzate per calcolare una base ortonormale per lo spazio vettoriale V_j . I coefficienti $\{\lambda_{j,k}\}$ della (4.9) saranno ottenuti mediante la proiezione ortogonale della funzione $f(x)$ sulle corrispondenti scaling function:

$$\lambda_{j,k} = \langle f, \varphi_{j,k} \rangle$$

Pertanto, la proiezione $P_j[f(x)]$ della funzione $f(x)$ sullo spazio V_j (cioè l'approssimazione della funzione f alla scala j),

$$P_j[f(x)] = \sum_k \langle f, \varphi_{j,k} \rangle \varphi_{j,k}(x),$$

è una proiezione ortogonale e definisce la migliore approssimazione in V_j di $f(x)$, nel senso che:

$$\forall g(x) \in V_j \Rightarrow \|g(x) - f(x)\| \geq \|P_j[f(x)] - f(x)\|$$

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

Analogamente, se le funzioni wavelet $\psi_{j,k}(x)$ formano una base ortogonale di $L^2(\mathbb{R})$, le proiezioni $\{Q_j\}$, (4.12), sono proiezioni ortogonali. In tal caso, i coefficienti $\{\gamma_{j,k}\}$ possono essere ottenuti proiettando la funzione $f(x)$ sulle wavelet stesse:

$$\boxed{\gamma_{j,k} = \langle f, \psi_{j,k} \rangle}$$

ed il dettaglio alla scala j viene espresso come:

$$Q_j[f(x)] = \sum_k \langle f, \psi_{j,k} \rangle \psi_{j,k}(x) = \sum_k \gamma_{j,k} \psi_{j,k}(x)$$

Una MRA dove lo spazio wavelet W_j è il complemento ortogonale a V_j in V_{j+1} , si definisce *analisi multirisoluzione ortogonale*. Ne consegue che gli spazi W_j , con $j \in \mathbb{Z}$, sono tutti mutuamente ortogonali, le proiezioni P_j e Q_j , precedentemente definite, sono ortogonali fra loro e l'espansione:

$$\begin{aligned} f(x) &= \sum_j Q_j[f(x)] = \sum_{j,k} \langle f, \psi_{j,k} \rangle \psi_{j,k}(x) = \\ &= \lim_{j \rightarrow \infty} P_j[f(x)] = \lim_{j \rightarrow \infty} \sum_k \langle f, \varphi_{j,k} \rangle \varphi_{j,k}(x) \end{aligned}$$

è un'espansione di funzioni ortogonali.

$P_j[f(x)]$ e $Q_j[f(x)]$ offrono la migliore approssimazione di $f(x)$ nei rispettivi spazi.

Una condizione sufficiente affinché una MRA sia ortogonale è:

$$W_0 \perp V_0 \text{ oppure } \langle \psi(x), \varphi(x - k) \rangle = 0 \text{ per } k \in \mathbb{Z}$$

Wavelet Biortogonali

L'ortogonalità limita fortemente la costruzione di wavelet. Per esempio, la wavelet di Haar è la sola wavelet a valori reali definita su un dominio compatto, che è simmetrica ed ortogonale. La definizione su di un dominio compatto consente una accurata implementazione della trasformata, in quanto i filtri h_k e g_k , associati alla scaling function ed alla wavelet (equazioni (4.6) e (4.11)), sono filtri FIR e quindi caratterizzati da un numero finito di coefficienti non nulli. Come sarà descritto più avanti, è possibile calcolare i valori dei coefficienti della trasformata utilizzando operazioni di convoluzione della funzione da trasformare con i filtri della MRA. L'esistenza di filtri FIR corrispondenti ad una determinata MRA rende quindi possibile una procedura di calcolo efficiente della trasformata secondo la MRA stessa. L'estensione della teoria dell'analisi

multirisoluzione a basi biortogonali consente di ottenere due coppie di filtri, una da usare per la trasformata (per ottenere i coefficienti) e una per l'antitrasformata (per ricostruire la funzione dai coefficienti). Gli schemi in figura 4.2 riassumono questi concetti.

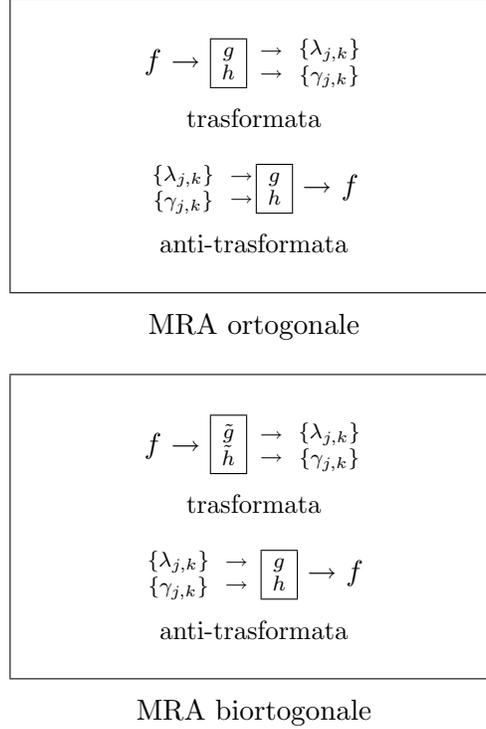


Figura 4.2.: Calcolo della trasformata e della antitrasformata con le differenti famiglie di MRA

La generalizzazione a wavelet biortogonali permette una maggiore flessibilità nella scelta delle funzioni base. In questo caso, esistono le funzioni $\tilde{\varphi}(x)$ e $\tilde{\psi}(x)$, rispettivamente *scaling function duale* e *wavelet duale*, le quali generano una MRA duale con sottospazi \tilde{V}_j e \tilde{W}_j , tali che:

$$\tilde{V}_j \perp W_j \quad \text{e} \quad V_j \perp \tilde{W}_j \quad (4.13)$$

Dato che $V_j = V_{j-1} \oplus W_{j-1}$, risulta:

$$\tilde{W}_j \perp (V_{j-i} \oplus W_{j-1})$$

e, conseguentemente:

$$W_j \perp \tilde{W}_k \quad \text{per } j \neq k$$

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

Esplicitando la (4.13) mediante relazioni tra le basi degli spazi considerati:

$$\langle \tilde{\varphi}(x), \psi(x-k) \rangle = \langle \tilde{\psi}(x), \varphi(x-k) \rangle = 0$$

Devono inoltre essere verificate le seguenti condizioni:

$$\langle \tilde{\varphi}_{j,l}, \varphi_{j,k} \rangle = \delta_{l-k} \quad j, k, l \in \mathbb{Z} \quad (4.14)$$

$$\langle \tilde{\psi}_{j,l}, \psi_{i,k} \rangle = \delta_{j-i} \delta_{l-k} \quad i, j, k, l \in \mathbb{Z} \quad (4.15)$$

In particolare:

$$\langle \tilde{\varphi}, \varphi(-k) \rangle = \delta_k \quad \text{e} \quad \langle \tilde{\psi}, \psi(-k) \rangle = \delta_k$$

La definizione della $\tilde{\varphi}_{j,k}(x)$ e della $\tilde{\psi}_{j,k}(x)$ è simile a quella di, rispettivamente, $\varphi_{j,k}(x)$ e $\psi_{j,k}(x)$. Il ruolo delle funzioni primali e duali è interscambiabile: sostituendo le une alle altre, le formule sono ancora valide. È uso comune denominare “duale” la MRA dedicata all'analisi.

Dalle condizioni di biortogonalità è possibile ottenere:

$$\tilde{h}_{k-2l} = \langle \tilde{\varphi}(\cdot - l), \varphi(2 \cdot - k) \rangle \quad (4.16)$$

$$\tilde{g}_{k-2l} = \langle \tilde{\psi}(\cdot - l), \varphi(2 \cdot - k) \rangle \quad (4.17)$$

In particolare, scrivendo $\varphi(2 \cdot - k) \in V_1$ utilizzando le basi di V_0 e W_0 :

$$\varphi(2 \cdot - k) = \sum_l \tilde{h}_{k-2l} \varphi(\cdot - l) + \sum_l \tilde{g}_{k-2l} \psi(\cdot - l) \quad (4.18)$$

Ricordando le relazioni in (4.7) e (4.10), la (4.18) è generalizzabile: una funzione base per lo spazio V_j è esprimibile combinando le funzioni base di V_{j-1} e di W_{j-1} . Queste relazioni verranno utilizzate per ottenere un algoritmo veloce per il calcolo della trasformata e della antitrasformata wavelet.

Le MRA biortogonali vengono utilizzate per aggirare le limitazioni delle MRA ortogonali, pur mantenendone il maggior pregio: la possibilità di calcolare i coefficienti della trasformata mediante la proiezione ortogonale della funzione sulle funzioni base. Nel caso di una coppia di MRA biortogonali, una delle due è utilizzata per il calcolo dei coefficienti (*analysis MRA*) e l'altra è utilizzata per il calcolo dell'approssimazione (*synthesis MRA*) e del dettaglio. Gli operatori proiezione P_j e Q_j diventano così:

$$P_j[f(x)] = \sum_k \langle f, \tilde{\varphi}_{j,k} \rangle \varphi_{j,k}(x) \quad (4.19)$$

$$Q_j[f(x)] = \sum_k \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k}(x) \quad (4.20)$$

e la trasformata wavelet discreta assume la forma:

$$f = \sum_{j,k} \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k}$$

Per la dualità che sussiste tra le due MRA, è possibile definire una coppia di proiezioni duale $\tilde{P}_j[f(x)]$ e $\tilde{Q}_j[f(x)]$ con caratteristiche analoghe alla coppia P_j e Q_j .

Fast Wavelet Transform

Le relazioni di ortogonalità o biortogonalità consentono di definire un algoritmo veloce per il calcolo della trasformata wavelet. Esso sfrutta la relazione tra le funzioni base ed i filtri ad esse associati per riformulare l'operazione di proiezione della funzione sugli spazi approssimazione e dettaglio come un'operazione di filtraggio con filtri FIR passa-basso e passa-alto.

Dato che V_j è uguale a $V_{j-1} \oplus W_{j-1}$, una funzione $v_j(x) \in V_j$ può essere scritta in modo univoco come somma di una funzione $v_{j-1}(x) \in V_{j-1}$ e di una funzione $w_{j-1}(x) \in W_{j-1}$:

$$\begin{aligned} v_j(x) &= \sum_k \lambda_{j,k} \varphi_{j,k}(x) = v_{j-1}(x) + w_{j-1}(x) = \\ &= \sum_k \lambda_{j-1,k} \varphi_{j-1,k}(x) + \sum_k \gamma_{j-1,k} \psi_{j-1,k}(x) \end{aligned}$$

per opportuni insiemi di coefficienti $\{\lambda_{j,k}\}$, $\{\lambda_{j-1,k}\}$ e $\{\gamma_{j-1,k}\}$.

Abbiamo quindi due rappresentazioni della stessa funzione $v_j(x)$, una come elemento di V_j , associata alla sequenza $\{\lambda_{j,k}\}$, ed una come somma di elementi di V_{j-1} e di W_{j-1} , associata alle sequenze $\{\lambda_{j-1,k}\}$ e $\{\gamma_{j-1,k}\}$. La relazione tra queste due rappresentazioni permetterà di definire una tecnica efficiente per applicare la trasformata wavelet. Si ha infatti che:

$$\lambda_{j-1,l} = \langle v_j, \tilde{\varphi}_{j-1,l} \rangle = \sqrt{2} \left\langle v_j, \sum_k \tilde{h}_{k-2l} \tilde{\varphi}_{j,l} \right\rangle = \sqrt{2} \sum_k \tilde{h}_{k-2l} \lambda_{j,l} \quad (4.21)$$

In modo analogo si ottiene:

$$\gamma_{j-1,l} = \sqrt{2} \sum_k \tilde{g}_{k-2l} \lambda_{j,l} \quad (4.22)$$

La relazione inversa si ottiene dalla *refinement equation* della duale:

$$\lambda_{j,k} = \sqrt{2} \sum_l h_{k-2l} \lambda_{j-1,l} + \sqrt{2} \sum_l h_{k-2l} g_{k-2l} \gamma_{j-1,l} \quad (4.23)$$

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

L'applicazione ricorsiva di queste formule definisce la *Fast Wavelet Transform* (FWT); in particolare, la (4.21) e la (4.22) definiscono la trasformazione diretta (*forward transform*) mentre la (4.23) definisce la trasformazione inversa (*Inverse FWT — IFWT*). L'algoritmo di analisi e sintesi di segnali digitali appena descritto viene anche chiamato *cascade algorithm* [Mal89].

La (4.21) e la (4.22), le quali permettono di calcolare i coefficienti dell'approssimazione e del dettaglio a livello $(j - 1)$, corrispondono quindi alla convoluzione dei vettori dei filtri (eventualmente infiniti) $\tilde{\mathbf{h}}$ e $\tilde{\mathbf{g}}$ con il vettore dei coefficienti approssimazione a livello j , $\boldsymbol{\lambda}_j$, seguita da una sottocampionamento. Trattando con elementi discreti, definiti su un intervallo finito, gli insiemi $\{\lambda_{j-1,k}\}$ e $\{\gamma_{j-1,k}\}$ avranno, quindi, un numero di elementi pari alla metà dell'insieme $\{\lambda_{j,k}\}$. Il numero di coefficienti necessari per rappresentare la funzione, quindi, rimane invariato.

Le trasformata inversa, realizzata mediante la (4.23), corrisponde essenzialmente alla somma della convoluzione dei vettori \mathbf{h} e \mathbf{g} rispettivamente con le successioni ottenute intercalando le successioni $\{\lambda_{j-1,k}\}$ e $\{\gamma_{j-1,k}\}$ con elementi nulli.

Questa procedura è adatta ad essere realizzata mediante circuiti digitali — e questo è uno dei motivi del diffuso utilizzo della FWT — e può essere riformulata a questo scopo in modo più compatto. Data una sequenza \mathbf{f} ottenuta campionando regolarmente il segnale $f(\cdot)$, i coefficienti approssimazione del primo livello wavelet sono quindi ottenuti come $\mathbf{a}_1 = \downarrow 2(\mathbf{f} * \tilde{\mathbf{g}})$, e i coefficienti dettaglio come $\mathbf{d}_1 = \downarrow 2(\mathbf{f} * \tilde{\mathbf{h}})$, dove $\tilde{\mathbf{g}}$ e $\tilde{\mathbf{h}}$ sono rispettivamente il filtro FIR di scomposizione (analisi) passa-basso e passa-alto (i filtri duali) caratteristici della MRA considerata, aventi un numero finito di elementi. L'operatore $\downarrow 2(\cdot)$ è l'operatore di sottocampionamento: scarta un campione ogni due. La procedura viene iterata per ottenere i coefficienti degli strati a scala più elevata utilizzando i coefficienti approssimazione calcolati nello strato precedente: $\mathbf{a}_j = \downarrow 2(\mathbf{a}_{j-1} * \tilde{\mathbf{g}})$ e $\mathbf{d}_j = \downarrow 2(\mathbf{a}_{j-1} * \tilde{\mathbf{h}})$. Dopo n iterazioni, il segnale \mathbf{f} è rappresentato dalla seguente collezione di coefficienti: $\{\mathbf{a}_n, \mathbf{d}_1, \dots, \mathbf{d}_n\}$. Dati \mathbf{a}_n e \mathbf{d}_n , i coefficienti approssimazione dello strato a risoluzione più elevata possono essere calcolati come $\mathbf{a}_{n-1} = \uparrow 2(\mathbf{a}_n) * \mathbf{g} + \uparrow 2(\mathbf{d}_n) * \mathbf{h}$, dove \mathbf{g} e \mathbf{h} sono i filtri FIR di ricostruzione (sintesi) passa-basso e passa-alto, e $\uparrow 2(\cdot)$ è l'operatore di sovracampionamento, il quale inserisce elementi nulli fra gli elementi della sequenza. I filtri \mathbf{g} e \mathbf{h} sono legati ai filtri $\tilde{\mathbf{g}}$ e $\tilde{\mathbf{h}}$ da una relazione di ortogonalità o biortogonalità (pp. 107 e 108).

Le figure 4.3 e 4.4 contengono gli schemi di analisi e sintesi della MRA mediante FWT. Gli schemi riportati prendono in considerazione

quattro livelli di scomposizione. Questa particolarità è esclusivamente legata all'esperimento comparativo riportato nel seguito (§4.3.2). Uno schema più generale dovrebbe avere una struttura iterativa, della quale lo schema riportato rappresenta lo srotolamento eseguito per quattro passi.

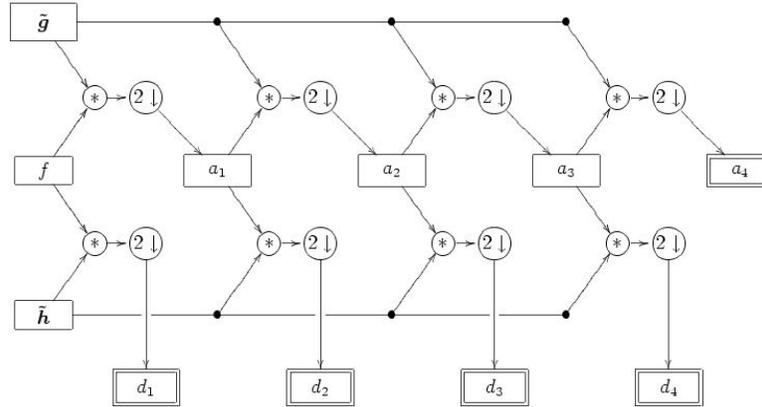


Figura 4.3.: Schema per una implementazione digitale della FWT

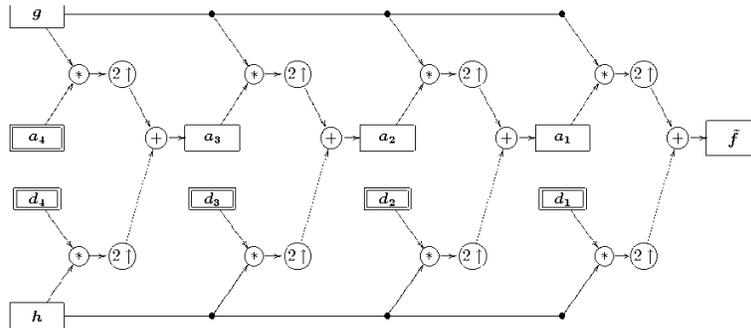


Figura 4.4.: Schema per una implementazione digitale della IFFT

Un problema da risolvere per poter applicare lo schema della trasformata wavelet veloce, consiste nello stimare i coefficienti λ_0 iniziali. Essi corrispondono al prodotto interno della scaling function del livello iniziale con la funzione. Sebbene ciò sia concettualmente differente dal campionamento della funzione, è possibile dimostrare che i campioni della funzione possono essere efficacemente utilizzati allo scopo. Intuitivamente, la scaling function a livello 0 può essere vista come un'approssimazione della funzione δ di Dirac. È importante notare che se i

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

filtri \mathbf{h} e \mathbf{g} sono filtri finiti (hanno cioè un numero finito di elementi non nulli), la FWT definisce un algoritmo per calcolare *esattamente* il prodotto interno della funzione per le scaling function dei livelli *inferiori* utilizzando solamente i coefficienti λ del livello considerato.

Grafici delle funzioni relative alle MRA

Talvolta le funzioni relative ad una MRA non possono essere espresse in forma analitica. È però possibile utilizzare il cascade algorithm per tracciare un grafico approssimato del loro andamento.

Un funzione $f(x) \in L^2(\mathbb{R})$ può essere rappresentata come:

$$f(x) = \sum_k \lambda_{j,k} \varphi_{j,k}(x) + \sum_{l \geq j} \sum_k \gamma_{l,k} \psi_{l,k}(x)$$

per un j fissato.

Disponendo dei coefficienti $\{\lambda_{j,k}\}$ e $\{\gamma_{l,k} \mid l \geq j\}$, la funzione $f(x)$ può essere ricostruita utilizzando il cascade algorithm.

L'unico requisito per operare la costruzione appena descritta sono i valori dei coefficienti dei filtri della MRA, $\{\mathbf{g}, \mathbf{h}, \tilde{\mathbf{g}}, \tilde{\mathbf{h}}\}$.

Questa tecnica è utilizzata per ricostruire funzioni smooth a partire da punti campionati (in modo equispaziato): si applica qualche passo del cascade algorithm con coefficienti dettaglio, $\{\gamma_{j,k}\}$, nulli. Se la MRA è interpolante, la funzione ricostruita passa per i punti campionati. Se la MRA non è interpolante, la funzione ricostruita è una versione filtrata dei punti campionati. Analogo effetto di filtraggio lo si può ottenere non utilizzando alcuni coefficienti dettaglio.

4.1.3. Applicazione alle immagini digitali

La trasformata wavelet può essere applicata anche a funzioni multidimensionali, quali le immagini digitali.

Analogamente al caso monodimensionale, si può definire una analisi multirisoluzione di $L^2(\mathbb{R}^2)$, e per essa dimostrare che esiste una funzione $\varphi(x, y)$ tale per cui l'insieme di funzioni $\{\varphi_{j,k,l} \mid j, k, l \in \mathbb{Z}\}$ definite come:

$$\varphi_{j,k,l}(x, y) = \sqrt{2^j} \varphi(2^j x - k, 2^j y - l)$$

forma una base di Riesz per lo spazio V^j .

Sono particolarmente interessanti le MRA separabili, nelle quali ogni sottospazio V_j può essere espresso come prodotto tensoriale di due sottospazi di $V_j^1 \subset L^2(\mathbb{R})$ identici:

$$V_j = V_j^1 \otimes V_j^1.$$

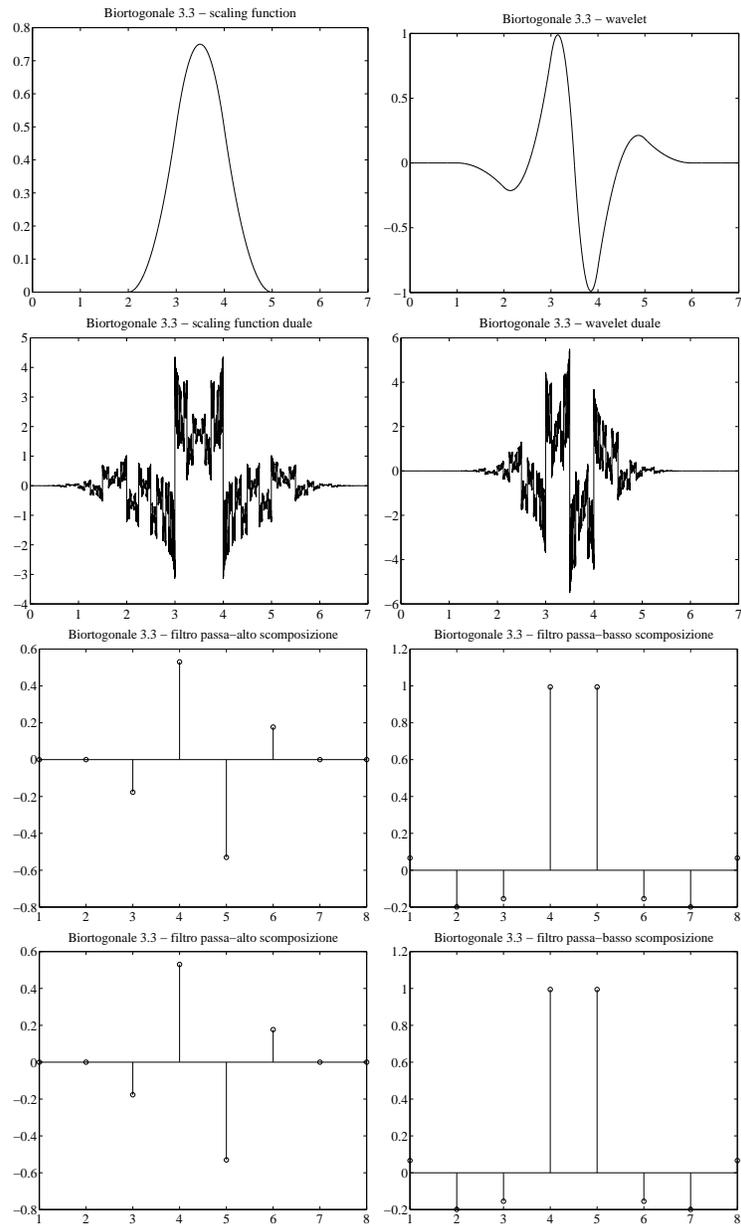


Figura 4.5.: Funzioni base e filtri della MRA biortogonale 3.3

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

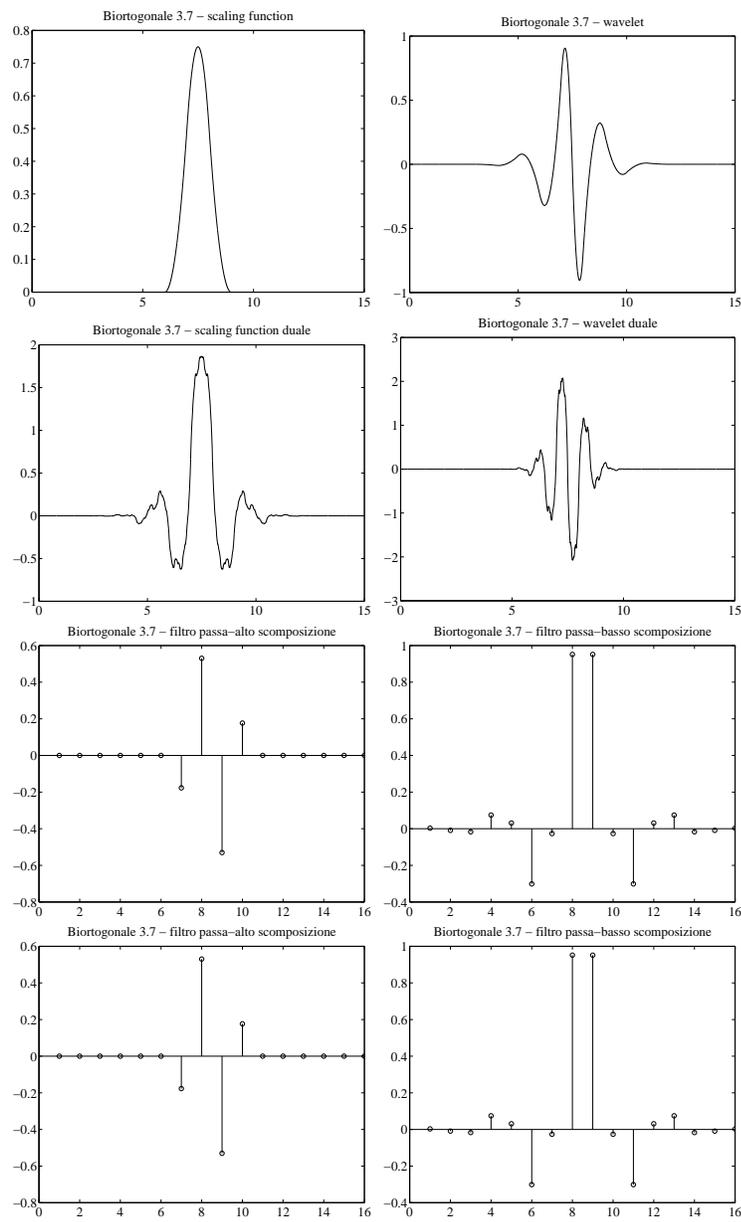


Figura 4.6.: Funzioni base e filtri della MRA biortogonale 3.7

In tal caso, si può dimostrare che la funzione $\varphi_{j,k,l}(x,y)$ può essere espressa come:

$$\varphi(x,y) = \varphi(x)\varphi(y)$$

dove $\varphi(x)$ è la scaling function di V^1 .

Sia $\psi(x)$ la wavelet associata a $\varphi(x)$. Si può dimostrare [Mal89], che esistono tre wavelet associate a $\varphi(x,y)$. Esse sono:

$$\begin{aligned}\psi^1(x,y) &= \varphi(x)\psi(y); \\ \psi^2(x,y) &= \psi(x)\varphi(y); \\ \psi^3(x,y) &= \psi(x)\psi(y).\end{aligned}$$

La scomposizione wavelet di un'immagine digitale produce quindi quattro sottoimmagini: una fornisce l'immagine a risoluzione ridotta, la quale rappresenta le componenti a bassa frequenza dell'immagine originale (approssimazione a risoluzione ridotta), mentre le altre forniscono rispettivamente le alte frequenze nella direzione orizzontale, nella direzione verticale ed in entrambe (il dettaglio nelle tre direzioni).

La separabilità della MRA comporta una semplificazione algoritmica molto vantaggiosa. Utilizzando una MRA separabile, infatti, la trasformata wavelet di immagini digitali può essere effettuata applicando la FWT monodimensionale prima sulle righe e poi sulle colonne. Se la MRA non fosse separabile, si dovrebbero applicare quattro filtri bidimensionali, con corrispondente incremento della complessità computazionale della trasformata.

Le considerazioni fatte per le immagini si applicano direttamente alle superfici del tipo $\mathbb{R}^2 \rightarrow \mathbb{R}$, se i punti della superficie sono campionati in modo equispaziato.

4.1.4. Estensioni

La formulazione originale delle wavelet è stata estesa per evitare le limitazioni che ne rendevano difficile o impossibile l'utilizzo in casi reali di particolare interesse.

In particolare, la ricerca è stata orientata verso il superamento dei seguenti vincoli:

- L'*equispaziatura dei dati* è una condizione necessaria per l'utilizzo degli operatori (campionamento e FFT) usati nella FWT. Nella pratica, però, i *range data* sono un caso molto diffuso.
- La formulazione prevede che lo spazio di definizione delle funzioni base abbia una topologia omeomorfa al piano. Molte applicazioni,

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

però, necessitano di una definizione in un dominio con topologia non-euclidea. Per esempio, la superficie di oggetti reali in generale non ha la topologia del piano.

- La particolare *famiglia di funzioni base* della MRA viene scelta in funzione delle caratteristiche dei dati che l'applicazione deve individuare. Tali caratteristiche sono talvolta in contrasto l'una con l'altra e quindi la scelta della forma delle funzioni base deve sottostare ad un appropriato compromesso. Definendo le caratteristiche degli spazi approssimazione e dettaglio in modo diverso, si può permettere l'uso contemporaneo di più di una mother wavelet.

Recentemente è stata proposta una metodologia, chiamata *Lifting Scheme*, per costruire wavelet adatte all'elaborazione di funzioni definite su un dominio di topologia più generale del piano euclideo (per esempio, la sfera [SS95]) e dati non equispaziati [Swe95], [Swe96], [Swe97]. Si tratta di una generalizzazione del cascade algorithm che consente di adattare la forma delle funzioni base di una MRA ai dati disponibili. Il Lifting Scheme scompone il calcolo della trasformata wavelet in una catena di operazioni di media pesata e differenze, intercalate con operazioni di (sovra/sotto)campionamento. Queste operazioni possono essere facilmente generalizzate o ridefinite per trattare il caso di dati non equispaziati. Se i dati non sono equispaziati, le funzioni base non sono più copie dilatate e traslate di una stessa funzione, ma, pur mantenendo le proprietà di annidamento (§4.1.2) e di complementarità (§4.1.2) degli spazi da esse generati, si adattano ai dati a disposizione. La maggiore elasticità nel trattamento dei dati viene raggiunta a scapito della velocità dell'algoritmo, in quanto il valore dei coefficienti dei filtri debbono essere calcolati per ogni campione.

L'applicazione del Lifting Scheme a spazi di dimensione superiore richiede una riformulazione in grado di trattare dati dotati di una topologia arbitraria, ed è quindi più complessa che non nel caso tradizionale [DGSSed]. Tuttavia, per modelli poligonali, il dominio è localmente topologicamente assimilabile al piano. I vertici del modello poligonale vengono utilizzati come i punti campionati e la relazione di adiacenza data dall'appartenenza allo stesso triangolo, fornisce la topologia del dominio. La mancanza della regolarità nel campionamento, richiede la definizione di un opportuno operatore di sottocampionamento. Esso può essere fornito da una qualsiasi procedura di semplificazione di superfici poligonali, alcuni dei quali sono stati descritti in §2.5.1. In [DGSSed] viene utilizzato *Progressive Meshes* [Hop96]. Questa scelta impone una ulteriore differenza tra le wavelet tradizionali e quelle che vengono chiamate wavelet di seconda generazione. In queste ultime un passo di applicazione della tra-

sformata non ha come argomento l'intera superficie, ma solo un vertice di essa. Tuttavia, una differente scelta dell'operatore di sottocampionamento può comportare un numero di livelli più contenuto, come, per esempio, in [LSS⁺98]. Per il calcolo dei coefficienti viene utilizzata una generalizzazione dello schema piramidale Burt-Adelson [BA83], basato su un'operazione di smoothing per ottenere i coefficienti approssimazione, mentre i coefficienti dettaglio vengono calcolati come differenza tra la superficie originale e la copia a bassa risoluzione. La definizione di tali operazioni nell'ambito delle mesh poligonali con connettività arbitraria, viene descritto in [GSS99].

Un procedimento analogo, in grado di fornire una rappresentazione MRA di una mesh triangolare con connettività arbitraria, viene proposto in [Bon98], dove la connettività delle mesh appartenenti alla gerarchia di approssimazioni viene definita mediante la triangolazione di Delaunay gerarchica [Kir83]. La trasformazione wavelet proposta è una generalizzazione delle wavelet di Haar.

Le tecniche presentate finora si applicano ad una mesh poligonale. In molti casi reali, come visto nel capitolo 2, i dati a disposizione sono costituiti da punti tridimensionali campionati in modo non uniforme. La ricostruzione della mesh poligonale che li approssimi è di importanza critica, perché fissa il genus del modello ricostruito. Diversi metodi per realizzare questo compito sono stati illustrati in §2.4. Alcuni lavori presenti in letteratura si sono occupati del problema della generazione di una rappresentazione MRA di dati non equispaziati. In [FE98] viene introdotta una tecnica *fine to coarse* per ottenere una ricostruzione wavelet a partire da dati campionati in modo non uniforme, limitata al caso monodimensionale. I coefficienti di scaling function e wavelet vengono ottenuti minimizzando l'errore di ricostruzione risolvendo un sistema lineare con il metodo dei minimi quadrati. Una delle ipotesi sulle quali il metodo si basa è che la topologia dell'insieme di dati sia quella della retta reale. In [PR99] i dati sono costituiti da una nuvola di punti tridimensionali. Si ipotizza che la superficie originale abbia la topologia ed il genus di una sfera. La mesh poligonale iniziale viene costruita adattando una sfera geodesica all'insieme di punti $3D$. Alla mesh risultante viene applicata la trasformata wavelet definita su un dominio sferico presentata in [SS95].

Le wavelet a M -bande (o multiwavelet) [SHGB93] [Hel95] [Mag99] [Mag00] rappresentano un'altra possibile generalizzazione della teoria MRA classica. La differenza nasce dalla formulazione più generale della proprietà 2 delle MRA (§4.1.2), dove, nella teoria delle wavelet a M -bande, con $M > 1$, diventa:

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

$$2'. v(x) \in V_j \Leftrightarrow v(Mx) \in V_{j+1}$$

Di conseguenza, la (4.7) viene modificata. È infatti possibile dedurre che, per ogni j , una base di Riesz per V_j è data da:

$$\{\varphi_{j,k}(x) \mid k \in \mathbb{Z}\} \quad \text{con } \varphi_{j,k}(x) = \sqrt{M^j} \varphi(M^j x - k) \quad (4.24)$$

Come nel caso $M = 2$ è possibile costruire spazi W_j che sono il complemento ortogonale di V_j in V_{j+1} , ottenendo così *MRA ortogonali a M -bande*. Passando da 2 a M bande, però, non è più possibile trovare una base per gli spazi W_j che sia costituita dalle traslazioni e dilatazioni di un'unica funzione. In questo caso, sono necessarie $M - 1$ wavelet, ortogonali tra di loro, non legate da relazioni di traslazione e dilatazione:

$$\psi_{j,k,l}(x) = \sqrt{M^j} \psi_l(M^j x - k) \quad j, k, \in \mathbb{Z}, l \in \{1, \dots, M - 1\} \quad (4.25)$$

Rimangono valide le relazioni di ortogonalità tra gli spazi W_j e tra le funzioni che ne costituiscono le basi:

$$\langle \psi_{j,k,l}, \psi_{j',k',l'} \rangle = \delta_{j,k,l} \delta_{j',k',l'} \quad j, k, j', k' \in \mathbb{Z}, l, l' \in \{1, \dots, M - 1\} \quad (4.26)$$

Una analoga generalizzazione si può operare per le MRA biortogonali.

Per le MRA spline ortogonali è possibile calcolare i filtri per le wavelet a partire da quelli per le scaling function [CDF92]. Al contrario, [Soa00], per le MRA spline biortogonali i coefficienti dei filtri per le wavelet non sono né facilmente né univocamente determinabili a partire da quelli relativi alle scaling function. Essi presentano però il vantaggio, rispetto alle MRA ortogonali di una regolarità delle funzioni sintetizzate che cresce linearmente con la lunghezza dei filtri, contro una crescita solo logaritmica della regolarità nel caso ortogonale. A fronte di una maggiore difficoltà nella progettazione, le wavelet a M -bande presentano una maggiore flessibilità di utilizzo. È possibile quindi imporre differenti condizioni sulle $M - 1$ wavelet in modo da ricavare informazioni di differente natura sul segnale, quali, per citarne alcune, simmetria, antisimmetria e regolarità. Poiché il sistema visivo umano sembra essere meno sensibile a errori simmetrici che a quelli antisimmetrici, uno schema *lossy* di compressione per immagini dovrebbe tendenzialmente scartare informazione dovuta alla simmetria piuttosto che a quella di natura antisimmetrica. Una coppia di wavelet, una simmetrica ed una antisimmetrica, possono essere utilizzate per separare questo tipo di informazione [Mag00]. Inoltre, il comportamento delle due wavelet sarà differente in prossimità di uno spigolo, il quale potrà quindi essere facilmente individuato in una immagine digitale.

4.2. Reti HRBF

4.2.1. Introduzione

L'approssimazione di una superficie della quale si conoscano un limitato numero di punti affetti da rumore è visto in ambito connessionistico come un problema di apprendimento. È un caso particolare dell'approssimazione di superfici in spazi n -dimensionali.

In questo lavoro, siamo interessati ad un particolare tipo di rete neurale della famiglia dei perceptron [HKP91]. All'interno di questa famiglia di modelli, i diversi tipi di rete si differenziano tra loro per la struttura e per la funzione di attivazione (o di risposta) delle unità neurali. In particolare, le unità a simmetria radiale consentono di ottenere una buona approssimazione con un numero ridotto di unità. Inoltre, queste unità sono generalmente locali. Esse sono caratterizzate da una posizione nello spazio di input e da un "raggio" che definisce la loro regione di influenza: solamente input all'interno di questa regione producono una risposta significativa dell'unità. La proprietà di località contribuisce a migliorare il processo di apprendimento, evitando la collinearità degli errori e la produzione di cattive soluzioni associate a minimi locali. Inoltre, è stato proposto come reti di questo tipo possano essere uno dei moduli di base computazionali implementati nel nostro sistema nervoso centrale [PG90] [Pla91] [MD89]. Le reti che utilizzano queste unità prendono il nome di *Radial Basis Functions Networks* (RBF) [Pog90].

Usando unità Gaussiane, una RBF è, in generale, esprimibile come funzione $s(\mathbf{x}) : \mathbb{R}^D \rightarrow \mathbb{R}$ di seguito descritta:

$$s(\mathbf{x}) = \sum_{k=1}^M w_k g(\mathbf{x}; \mathbf{c}_k, \sigma_k) = \sum_{k=1}^M w_k \frac{e^{-\frac{\|\mathbf{x}-\mathbf{c}_k\|^2}{\sigma_k^2}}}{\sqrt{\pi^D \sigma_k^D}} \quad (4.27)$$

dove M è il numero di gaussiane utilizzate, $w_k \in \mathbb{R}$ è il coefficiente moltiplicativo (o peso), $\sigma_k \in \mathbb{R}$ e $\mathbf{c}_k \in \mathbb{R}^D$ sono rispettivamente, la deviazione standard e la posizione della k -esima gaussiana, D è la dimensione dello spazio di input. σ_k determina l'ampiezza della regione di influenza: la k -esima gaussiana, infatti, assume un valore significativamente diverso da zero nella regione circolare di centro \mathbf{c}_k e raggio funzione di σ_k . Le gaussiane qui utilizzate hanno norma unitaria.

Il problema dell'apprendimento può essere formulato nel seguente modo: dato un insieme di punti $X = \{(\mathbf{x}_j, z_j) | \mathbf{x} \in \mathbb{R}^D, z \in \mathbb{R}, 1 \leq j \leq N\}$ individuare l'insieme di parametri $\{M, \mathbf{c}_k, \sigma_k, w_k\}$ tali che una opportuna funzione errore $E(s(\mathbf{x}), \{z\})$ sia minimizzata.

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

Il problema è di per sé mal posto perché esistono infinite superfici che passano per i punti dati (che sono in numero finito). Occorre pertanto introdurre dei vincoli che aiutino a scegliere fra le infinite soluzioni la più adatta al compito di approssimazione. Un vincolo naturale è quello di smoothness: a piccole variazioni della posizione di un punto devono corrispondere piccole variazioni della superficie.

La selezione dei parametri ottimali è un problema complesso, e implica una ottimizzazione non lineare in presenza di molti minimi locali. Diverse euristiche sono state proposte per la selezione dei parametri.

Vincoli ed euristiche hanno lo scopo di guidare l'apprendimento della rete in modo da inglobare la conoscenza a priori sulla funzione da apprendere.

Il problema dell'apprendimento può essere riformulato nel campo della regolarizzazione introducendo un termine di smoothness nella funzione costo da minimizzare [GJP95] [Pog90] [Yui88]. Si dimostra che una combinazione lineare di gaussiane centrate nei punti dati è in grado di minimizzare tale funzione costo. A questa chiara formulazione non corrisponde però una soluzione algoritmica calcolabile (cf. [BA95] [NP91] [WD92]). L'uso di gradiente stocastico [GJP95] risente dei minimi locali, i quali possono impedire di ottenere una soluzione accettabile. Per ottenere il minimo globale, devono essere quindi impiegati algoritmi specializzati e complessi (cf. [Ber93] [CBB93] [GST94]). La soluzione può essere cercata operando una ottimizzazione su tutti i parametri utilizzando tecniche alternative [BZ95] [NT89] [KGV83], le quali, sebbene teoricamente in grado di raggiungere il minimo globale, sono basate sulla ricerca esaustiva nello spazio dei parametri e richiedono un notevole sforzo computazionale.

Una strategia differente, chiamata *hybrid learning* [BL88] [MD89], deriva dall'osservazione che i parametri della rete espressa in (4.27) possono essere divisi in due categorie a seconda della funzione da loro svolta: parametri strutturali e pesi sinaptici. Alla prima categoria appartengono i parametri che influenzano il comportamento delle gaussiane, cioè M , $\{c_k\}$ e $\{\sigma_k\}$, mentre alla seconda appartengono i parametri che stanno a valle delle funzioni gaussiane: i pesi $\{w_k\}$. La differente natura di questi insiemi di parametri, suggerisce di usare algoritmi diversi per assegnare loro il valore appropriato. In un approccio di questo tipo, il numero delle gaussiane, M , è solitamente dato a priori e la loro posizione, $\{c_k\}$, è determinata mediante algoritmi di clustering [Llo82] [MBS93]. I parametri $\{\sigma_k\}$, critici in quanto definiscono il comportamento della funzione nelle regioni fra i campioni, sono determinati utilizzando considerazioni euristiche [MD89] [Orr95] [Pla91] [Spe90].

Qualora i parametri strutturali siano fissati, la (4.27) descrive un sistema lineare dove le uniche incognite sono i pesi, $\{w_k\}$. Sebbene questi possano essere calcolati risolvendo il sistema, per reti di grandi dimensioni questa soluzione può causare sia problemi numerici che di allocazione della memoria, e quindi uno schema locale per il calcolo dei pesi può essere preferibile. Un miglioramento degli schemi ibridi è rappresentato dalle strutture ad accrescimento [Fri94] [Fri95] [Pla91], dove il numero di gaussiane (e in generale delle unità) non è dato a priori, ma le unità vengono inserite nella rete una alla volta, fino al soddisfacimento di un certo criterio. Questi schemi sono iterativi e, per trovare valori ragionevolmente ottimali dei parametri, richiedono un apprendimento relativamente lungo.

Un approccio alternativo, che deriva dalla computer vision e dal controllo, consiste nel disporre nello spazio di input una griglia di gaussiane equispaziate [PM90] [SS92]. L'inconveniente di questo approccio sta nella rigidità della struttura adottata, la quale è dotata di un'unica varianza per tutte le unità: la funzione risultante può quindi contemporaneamente soffrire di *overfitting*¹ in alcune regioni dello spazio di input e perdere particolari di interesse in altre. Inoltre, la presenza di overfitting implica uno spreco di risorse dovuto ad un eccesso di unità.

In questo capitolo verrà presentato un paradigma, chiamato *Hierarchical Radial Basis Functions* (HRBF) [BF98] [BMF99], che combina le caratteristiche delle strutture ad accrescimento con la teoria del filtraggio per ottenere una determinazione rapida e stabile sia dei parametri strutturali che dei pesi della rete.

4.2.2. Rete RBF gaussiana regolare e filtro gaussiano

In un primo momento, limiteremo la trattazione al caso $\mathbb{R} \rightarrow \mathbb{R}$. L'estensione agli spazi di dimensionalità superiore è facile per via della fattorizzabilità della gaussiana e verrà introdotta in §4.2.4.

Quando le gaussiane sono equispaziate ed hanno tutte la stessa deviazione standard, σ , la rete opera una convoluzione della sequenza di pesi w_k con un filtro gaussiano $g(x)$. La teoria del filtraggio può essere quindi utilizzata efficacemente per individuare i valori da assegnare ai parametri strutturali ed ai pesi.

$$s(x) = \sum_{k=1}^M w_k g(x; c_k, \sigma) = \sum_{k=1}^M w_k \frac{e^{-\frac{(x-c_k)^2}{\sigma^2}}}{\sqrt{\pi}\sigma} \quad (4.28)$$

¹L'*overfitting* è la situazione per la quale una rete (e più in generale uno stimatore) approssima i dati in modo troppo accurato, fornendo così una cattiva generalizzazione dell'insieme dato.

Filtro Gaussiano

Il filtro passa-basso ideale lascia inalterate tutte le frequenze inferiori ad $\omega_{\text{cut-off}}$ e annulla il contributo delle componenti a frequenza superiore. Questo filtro non è però fisicamente realizzabile. Nella pratica un filtro passa-basso, $\hat{f}(\omega)$, è caratterizzato da due frequenze, $\omega_{\text{cut-off}}$ e ω_{max} , che identificano tre bande. Definendo due soglie δ_1 e δ_2 , $\omega_{\text{cut-off}}$ e ω_{max} sono identificate dalle seguenti relazioni:

$$\delta_1 \leq |\hat{f}(\omega)| \leq 1 \quad \text{per } 0 \leq \omega \leq \omega_{\text{cut-off}} \quad (4.29)$$

$$0 \leq |\hat{f}(\omega)| \leq \delta_2 \quad \text{per } \omega_{\text{max}} \leq \omega \leq +\infty \quad (4.30)$$

Le tre bande, *Pass Band*, *Stop Band* e *Transition Band*, corrispondono alle regioni dove le frequenze vengono mantenute inalterate, soppresse o progressivamente smorzate.

Il filtro gaussiano (figura 4.7) è espresso dalla seguente formula:

$$g(x; \sigma) = \frac{1}{\sqrt{\pi}\sigma} e^{-\frac{x^2}{\sigma^2}} \quad (4.31)$$

La sua trasformata di Fourier è:

$$\hat{g}(\omega; \sigma) = e^{-\pi^2 \sigma^2 \omega^2} \quad (4.32)$$

Le frequenze caratteristiche sono legate alla σ tramite le seguenti relazioni:

$$\begin{cases} e^{-\pi^2 \sigma^2 \omega_{\text{cut-off}}^2} = \delta_1 \\ e^{-\pi^2 \sigma^2 \omega_{\text{max}}^2} = \delta_2 \end{cases} \Rightarrow \begin{cases} \omega_{\text{cut-off}} = \frac{\sqrt{-\ln \delta_1}}{\pi \sigma} \\ \omega_{\text{max}} = \frac{\sqrt{-\ln \delta_2}}{\pi \sigma} \end{cases} \quad (4.33)$$

Rete RBF regolare continua

La (4.28) descrive una rete neurale, che, per essere fisicamente realizzabile, deve necessariamente essere costituita da un numero finito di gaussiane. È possibile però immaginare una generalizzazione di tale rete, costituita da infiniti elementi contigui. La continuità strutturale di questa nuova RBF è descritta tramite un regione di supporto $\mathbb{U} \subseteq \mathbb{R}$ (intervallo del dominio di input nel quale sono definite le gaussiane), una funzione peso $w(x) : \mathbb{U} \rightarrow \mathbb{R}$ e da una funzione $\sigma(x) : \mathbb{U} \rightarrow \mathbb{R}^+$.

Se la distanza tra le gaussiane tende a zero, l'equazione (4.28) può essere generalizzata mediante una RBF continua con $\mathbb{U} \equiv \mathbb{R}$ e $\sigma(x) = \sigma$:

$$\hat{s}(x) = \int_{\mathbb{R}} w(c) g(x - c|\sigma) dc = w(x) * g(x; \sigma) \quad (4.34)$$

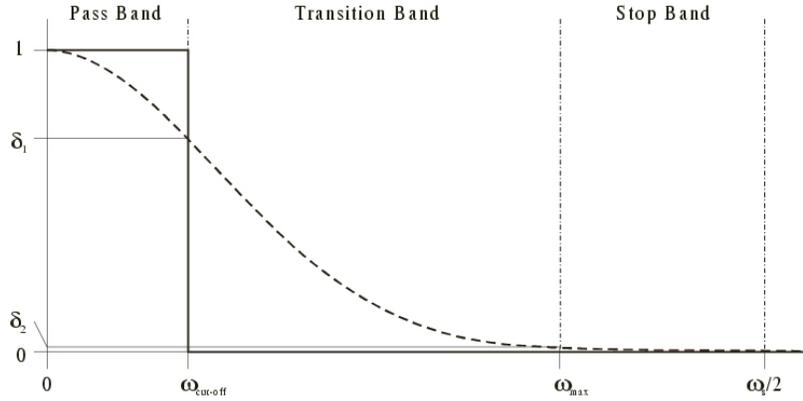


Figura 4.7.: Filtro gaussiano

Per il teorema della convoluzione, la (4.34) implica:

$$\hat{s}(\omega) = \hat{w}(\omega) \hat{g}(\omega) \quad (4.35)$$

Se al posto di $w(x)$ utilizziamo la funzione $s(x)$ stessa, otteniamo la funzione $\tilde{s}(x)$:

$$\tilde{s}(x) = \int_{\mathbb{R}} s(c) g(x - c|\sigma) dc = s(x) * g(x; \sigma) \quad (4.36)$$

e nel dominio della frequenza:

$$\hat{\tilde{s}}(\omega) = \hat{s}(\omega) \hat{g}(\omega; \sigma) \quad (4.37)$$

Dalla (4.37) risulta chiaro che $\tilde{s}(x)$ sarà quindi una versione smooth di $s(x)$ in quanto tutte le componenti in frequenza di $s(x)$ superiori a $\omega_{\text{cut-off}}$ del filtro gaussiano saranno progressivamente attenuate. Per ottenere una approssimazione fedele, se $s(x)$ ha componenti in frequenza non oltre ω_M , la scelta del valore da utilizzare per il parametro σ sarà sottoposta al soddisfacimento del seguente vincolo:

$$\omega_{\text{cut-off}} > \omega_M \quad (4.38)$$

Rete RBF regolare discreta

Nei casi di interesse pratico, si hanno a disposizione solo un numero limitato, N , di campioni. Ipotizziamo di ricostruire la funzione utilizzando una combinazione lineare di gaussiane equispaziate, con passo Δx_p . Si

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

dimostra che la funzione può essere ricostruita in corrispondenza dei punti campionati $\{x_j\}$ come:

$$\tilde{s}(x_j) = \sum_{k=1}^N s_k g(x_j; x_k, \sigma) \Delta x_p = \frac{\Delta x_p}{\sqrt{\pi} \sigma} \sum_{k=1}^N s_k e^{-\frac{(x_j - x_k)^2}{\sigma^2}} \quad (4.39)$$

La (4.39) può essere estesa a tutto l'intervallo continuo usando la seguente interpolazione:

$$\tilde{s}(x) = \sum_{k=1}^N s_k g(x; x_k, \sigma) \Delta x_p = \frac{\Delta x_p}{\sqrt{\pi} \sigma} \sum_{k=1}^N s_k e^{-\frac{(x - x_k)^2}{\sigma^2}} \quad (4.40)$$

In (4.39) anche il filtro gaussiano viene campionato. Questo fatto introduce un ulteriore vincolo:

$$\omega_{\max} < \frac{\omega_s}{2} \quad (4.41)$$

È quindi opportuno ritoccare le relazioni (4.33) nel seguente modo:

$$\begin{cases} e^{-\pi^2 \sigma^2 \omega_{\text{cut-off}}^2} = \delta_1 \\ e^{-\pi^2 \sigma^2 \omega_{\max}^2} = \delta_2/2 \end{cases} \quad (4.42)$$

dove δ_2 è stato dimezzato perché quando $\omega_{\max} = \omega_s/2$, la gaussiana riceve un uguale contributo sia dal lobo principale che dalla sua replica più vicina. Ulteriori repliche potrebbero essere considerate in questo ragionamento, ma sono ritenute troppo distanti per influire significativamente.

Si dimostra che disponendo di un campionamento a frequenza ω_s , la funzione ricostruita non avrà componenti in frequenza superiori a $\bar{\omega}_M$ pari a:

$$\bar{\omega}_M = \sqrt{\frac{\ln \delta_1}{\ln \delta_2/2}} \frac{\omega_s}{2} \quad (4.43)$$

La (4.43) è confrontabile con il teorema di Shannon, secondo il quale la massima frequenza ricostruibile di un segnale campionato è pari alla metà della frequenza di campionamento del segnale stesso. Poiché, per come sono stati definiti, $\delta_1 > \delta_2$, la frequenza massima ricostruibile sarà inferiore a quella indicata da Shannon.

Fissando δ_1 a $\sqrt{2}/2$, in accordo con la pratica comune (attenuazione di 3dB), e δ_2 a 0.01, sostituendo i valori numerici nella (4.43) si ottiene:

$$\omega_s = 7.820 \bar{\omega}_M \quad (4.44)$$

La frequenza di campionamento dovrà essere circa otto volte superiore alla massima frequenza che si vuole ricostruire.

Diminuendo δ_1 o aumentando δ_2 il rapporto si avvicina a due. Così facendo, però si peggiora la qualità della ricostruzione: aumentare δ_2 significa incrementare l'effetto di aliasing, mentre diminuire δ_1 comporta una maggiore attenuazione delle componenti a frequenza più elevata della funzione da ricostruire.

Utilizzando le relazioni sopra descritte, siamo in grado, una volta fissata la massima frequenza da ricostruire, $\bar{\omega}_M$, compatibilmente con la frequenza di campionamento dei dati di input ω_s , di fissare il numero, M , la posizione, $\{c_k\}$, e la σ delle gaussiane. I pesi $\{w_k\}$ saranno dati dalla sottocampionatura dei dati di input risultante. Per esempio, sia $X = \{(x_j, z_j) | x \in \mathbb{R}, z \in \mathbb{R}, j \in \{1, \dots, N\}\}$ un insieme di punti campionati regolarmente, $\Delta x = x_{j+1} - x_j, \forall j$. Per la (4.44), una RBF del tipo sopra descritto è in grado di ricostruire componenti di frequenza non superiore a circa $\frac{1}{8\Delta x}$. Se questa è la massima frequenza contenuta nella funzione dalla quale sono stati campionati i dati X , essa può essere ricostruita ponendo una gaussiana in corrispondenza di ogni punto campionato: $c_k = x_k, M = N$. Dai vincoli precedentemente imposti sulle frequenze, e ponendo $\Delta x = 1/\omega_s$, si può dedurre che il parametro σ non potrà essere inferiore a $1.465 \Delta x$. I pesi w_k corrispondono ai valori assunti dai punti campionati: $w_k = z_k$. Se la massima frequenza contenuta nella funzione da ricostruire fosse inferiore a $\frac{1}{8\Delta x}$, la funzione potrebbe essere ricostruita con un numero inferiore di gaussiane, operando un sottocampionamento dei dati per ottenere i pesi e le posizioni delle gaussiane.

4.2.3. HRBF: la multirisoluzione

La tecnica precedente illustrata è inefficiente qualora la funzione da ricostruire presenti un contenuto in frequenza molto diverso in differenti regioni del suo dominio. La presenza di particolari ad alta frequenza in una porzione limitata del dominio comporta l'utilizzo di un gran numero di gaussiane, anche nelle regioni a basso contenuto in frequenza. Queste regioni potrebbero essere più efficientemente coperte con un numero inferiore di gaussiane aventi una σ più ampia ($\omega_{\text{cut-off}}$ minore).

La ricostruzione può essere effettuata utilizzando più di una rete, ognuna caratterizzata da un valore differente del parametro σ . La prima rete viene costruita con lo scopo di realizzare una buona approssimazione dell'andamento generale della funzione, $a_1(x)$. Il valore della $\omega_{\text{cut-off}}$ viene quindi scelto relativamente basso, ω_1 , e per la stima dei pesi delle gaussiane viene operato un opportuno sottocampionamento dell'insieme

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

di punti dato. Il residuo $r_1(x_k)$ è un insieme costituito dalla differenza tra la funzione ricostruita dalla prima rete e i punti dati:

$$r_1 = \{z_k - a_1(x_k)\} \quad (4.45)$$

Una seconda rete, caratterizzata da una $\omega_{\text{cut-off}}$ di valore ω_2 , $\omega_2 > \omega_1$, è utilizzata per ottenere una approssimazione del residuo r_1 . In questo modo si ottiene una rete che realizza la funzione $a_2(x)$. La procedura può essere ripetuta fino a che la $\omega_{\text{cut-off}}$ della rete rimane inferiore di $\frac{1}{8\Delta x}$ o fino a che il residuo è sufficientemente basso.

La procedura descritta genera una rappresentazione multirisoluzione della funzione da ricostruire. La rappresentazione alla risoluzione j è data dalla somma delle uscite delle prime j reti.

Generalizzazione: dati affetti da rumore e non equispaziati

La procedura sopra descritta presuppone dati equispaziati e non affetti da rumore. Qualora queste ipotesi non siano verificate, deve essere individuata una tecnica per stimare il valore della funzione in corrispondenza delle gaussiane.

La presenza di rumore non è di per sè un problema, in quanto il filtro gaussiano attenua le alte frequenze.

È però possibile migliorare la ricostruzione sfruttando la ridondanza informativa dei dati vicini, informazione che magari andrebbe perduta a causa del sottocampionamento. Inoltre, la spaziatura non uniforme tra i campioni non permette di posizionare le gaussiane in corrispondenza dei punti campionati. Bisogna quindi utilizzare una stima del valore assunto dalla funzione in corrispondenza dei centri delle gaussiane.

A tale scopo, occorre stabilire:

- un criterio di selezione dell'insieme di punti, S_k , da usare nella stima;
- un criterio di stima della funzione.

I criteri per la selezione dei punti sono essenzialmente due, dai quali si possono derivare diverse varianti e combinazioni. Il primo è basato sul numero di elementi per la stima: per ogni gaussiana, si selezionano un numero fissato, \bar{n} , di punti più vicini al centro c_k della gaussiana considerata. Il secondo è legato alla distanza: per ogni gaussiana si selezionano i punti che appartengono all'intorno sferico di raggio R_S (da fissare) del suo centro c_k . Entrambi i metodi hanno vantaggi e si prestano a critiche.

Il criterio di stima proposto in [BF98] è basato su una media pesata dei punti appartenenti all'insieme di stima S_k [AMS99]:

$$\bar{s}(\bar{x}) = \frac{\sum_{x_r \in S_k} z_r \bar{w}(|x_r - \bar{x}|)}{\sum_{x_r \in S_k} \bar{w}(|x_r - \bar{x}|)} \quad (4.46)$$

La funzione peso $\bar{w}(x)$ utilizzata è una funzione gaussiana. L'assunzione implicita è che i punti vicini siano in grado di fornire una stima affidabile con una legge di probabilità normale:

$$s(c_k) = \frac{\sum_{x_r \in S_k} z_r e^{-\frac{(x_r - c_k)^2}{\sigma_w^2}}}{\sum_{x_r \in S_k} e^{-\frac{(x_r - c_k)^2}{\sigma_w^2}}} \quad (4.47)$$

dove l'insieme dei punti per la stima, S_k , è:

$$S_k = \{x_m \mid \|x_m - c_k\| < R_s = \Delta x\} \quad (4.48)$$

ed il parametro σ_w è posto pari alla σ della RBF considerata (sebbene più cautelativamente possa essere posto anche a $\sigma/2$ [BF98]).

Inoltre, la k -esima gaussiana viene inserita se e solo se il valore medio della funzione da approssimare, valutato in un intorno sferico di c_k , supera la soglia ϵ (stima del rumore), cioè se viene soddisfatta la seguente condizione:

$$\frac{\sum_{x_r \in S_k} \|z_r\|}{|S_k|} > \epsilon \quad (4.49)$$

Si ha quindi il seguente algoritmo:

1. Si seleziona un insieme ordinato di L frequenze di cut-off caratteristiche di altrettanti strati di RBF gaussiane $\{\omega_l \mid l = 1, \dots, L\}$.
2. Si calcolano, corrispondentemente, i valori delle σ degli strati: $\{\sigma_l \mid l = 1, \dots, L\}$.
3. Si pone $r_0 = \{r_0(x_k) = z_k\}$.
4. Per ogni strato, l , si posizionano i punti $\{c_{l,j} \mid c_{l,j} - c_{l,j-1} = \Delta x_l = \sigma_l/1.465\}$. Essi sono i centri delle gaussiane dello strato l -esimo.
5. Per ogni gaussiana, j , se la (4.49) è soddisfatta, viene calcolato il coefficiente $w_{l,j} = \Delta x_l r_{l-1}(c_{l,j})$, dove $r_{l-1}(c_{l,j})$ è stimata nel modo indicato dalla (4.47).
6. Si calcola l'insieme $a_l = \{a_l(x_k) = \sum_j w_{l,j} g(x_k; c_{l,j}, \sigma_l)\}$.

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

7. Si pone $r_l = \{r_{l-1}(x_k) - a_l(x_k)\}$.

8. Se $l < L$, si ritorna al punto 4.

La funzione ricostruita risulta quindi:

$$\tilde{s}(x) = \sum_{l=1}^L \sum_j w_{l,j} g(x; c_{l,j}, \sigma_l)$$

Sono quindi disponibili L ricostruzioni a risoluzione via via crescente:

$$\tilde{s}_l(x) = \sum_{h=1}^l a_h(x) = \sum_{h=1}^l \sum_j w_{h,j} g(x; c_{h,j}, \sigma_h)$$

L'algoritmo descritto presenta alcuni aspetti secondari i quali si prestano ad essere adattati a seconda delle peculiarità del problema in esame. Sebbene non vi sia alcun obbligo di natura teorica, l'insieme $\{\omega_1, \dots, \omega_L\}$ viene solitamente calcolato in modo che ogni elemento sia la metà della seguente. In questo modo si copre un ampio spettro con pochi strati. Il passo 1 non è sempre necessario: spesso è più naturale fissare il valore della σ degli strati (cioè la loro scala di risoluzione) direttamente, senza cioè fare uso della nozione di frequenza di taglio del filtro corrispondente allo strato. In tal caso l'algoritmo parte direttamente dal passo 2, dove usualmente si fissano le σ dimezzandone il valore al crescere degli strati. Il numero di strati stesso non deve necessariamente essere fissato a priori: se i dati a disposizione lo consentono, la procedura di inserzione di nuovi strati può essere iterata a piacere. Un criterio di terminazione, in questo caso, può essere basato sul soddisfacimento di un vincolo espresso sul residuo, r_l (per esempio, mediante un test di bianchezza).

Aspetti computazionali

Il numero di unità utilizzate dalla rete raddoppia, all'incirca, ad ogni strato — il fattore è in generale 2^D se la rete realizza una funzione $\mathbb{R}^D \rightarrow \mathbb{R}$. Per decidere se una unità va inserita, bisogna calcolare la distanza dai punti del centro della gaussiana, selezionare solo quelli all'interno del campo recettivo e comparare il residuo medio con l'accuratezza prescelta (4.49). Per ogni unità inserita, deve poi essere calcolato il suo valore di uscita, necessario per calcolare il dettaglio rimanente (passo 6). La complessità è $O(M \times N)$ dove M sono le gaussiane e N sono i dati. Nell'ipotesi che M sia proporzionale ad N , si ottiene una complessità computazionale pari a $O(N^2)$.

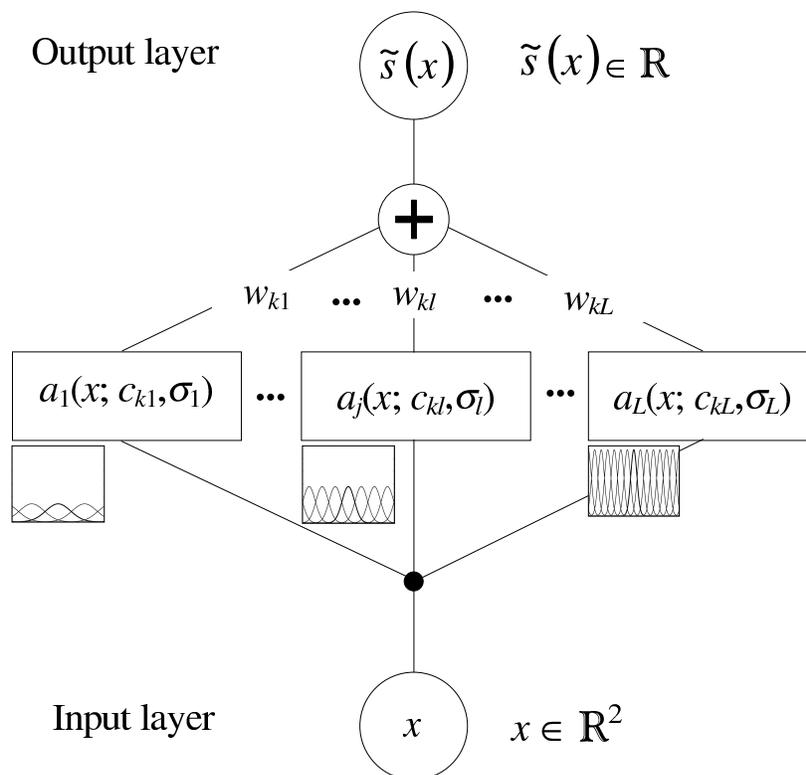


Figura 4.8.: Schema di ricostruzione della HRBF

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

Il costo computazionale potrebbe essere ridotto osservando che le unità dello strato l sono posizionate in un sottoinsieme dei centri delle unità dello strato $l + 1$, mentre le regioni di influenza delle unità dello strato superiore sono incluse in quelle dello strato inferiore. Una fase di preprocessing potrebbe organizzare i dati in modo tale da facilitare la selezione dei punti che cadono nel campo recettivo di ogni gaussiana [Sam84]. Per grandi insiemi di dati o per dimensioni elevate dello spazio di input rappresenterebbe un sostanzioso incremento di prestazioni (la ricerca sarebbe logaritmica in N). Un altro fattore da sfruttare per diminuire il costo computazionale è la località delle gaussiane. L'influenza di una gaussiana ad una distanza superiore di tre volte σ può considerarsi insignificante. Organizzando i dati in modo gerarchico, anche il calcolo dell'output degli strati potrebbe essere reso meno costoso. Il costo del preprocessing verrebbe quindi facilmente assorbito dall'aumento di prestazioni.

Una volta fissati i parametri strutturali, la stima dei pesi è equivalente alla soluzione di un sistema lineare e quindi appropriate tecniche (per esempio, LMS, scomposizione SVD) possono essere utilizzate per configurare i pesi. Per reti di grandi dimensioni, tuttavia, si può andare incontro a problemi di instabilità numerica (a causa della ridondanza delle basi, vedi §4.2.5) e di mancanza di risorse del sistema di calcolo (in quanto la soluzione è globale). La località dell'algoritmo di apprendimento presentato, pur sfruttando la correlazione tra dati vicini, evita che la stima di un peso risenta dell'influenza di quelli adiacenti.

Inoltre, i costi computazionali dell'algoritmo HRBF potrebbero essere ridotti utilizzando uno strumento di calcolo parallelo e con caratteristiche neurali [PG90]. La funzione gaussiana potrebbe essere in tal caso implementata mediante le diramazioni dendritiche o tramite insiemi di neuroni, i pesi nelle connessioni sinaptiche e le medie pesate locali attraverso i meccanismi di inibizione laterale.

4.2.4. Estensione al caso bidimensionale

Se si considera la gaussiana a simmetria radiale (4.27), l'estensione delle HRBF al caso bidimensionale risulta abbastanza semplice. Il filtro gaussiano bidimensionale risulta:

$$g(x; \sigma) = \frac{1}{\pi\sigma^2} e^{-\frac{\|x\|^2}{\sigma^2}} \quad (4.50)$$

dove $x \in \mathbb{R}^2$. Esplicitando le coordinate di $x = (x_1, x_2)$, si può riscrivere la (4.50) mettendo in evidenza la fattorizzazione della gaussiana:

$$g((x_1, x_2); \sigma) = \frac{1}{\pi\sigma^2} e^{-\frac{x_1^2+x_2^2}{\sigma^2}} = \frac{1}{\sqrt{\pi}\sigma} e^{-\frac{x_1^2}{\sigma^2}} \frac{1}{\sqrt{\pi}\sigma} e^{-\frac{x_2^2}{\sigma^2}} = g(x_1; \sigma) g(x_2; \sigma) \quad (4.51)$$

Questa proprietà consente di esprimere la trasformata di Fourier del filtro gaussiano bidimensionale come:

$$\hat{g}(\omega; \sigma) = e^{-\pi^2\sigma^2\omega_1^2} e^{-\pi^2\sigma^2\omega_2^2} = e^{-\pi^2\sigma^2\|\omega\|^2} \quad (4.52)$$

in quanto la trasformata di Fourier di una funzione separabile è uguale al prodotto delle trasformate monodimensionali dei singoli fattori.

Pertanto, tutte le considerazioni e le relazioni che permettono di fissare i parametri strutturali della rete ed i suoi pesi rimangono validi anche nel caso bidimensionale.

L'unica modifica è al passo 5 dell'algoritmo descritto in §4.2.3, dove la costante di proporzionalità tra il valore stimato della funzione e il coefficiente della gaussiana diventa Δx_l^2 invece di Δx_l .

4.2.5. Approssimazioni Multiscala

È intuitivamente chiaro che le reti HRBF realizzano una approssimazione multirisoluzione di funzioni $R^D \rightarrow R$. Non è però chiaro quanto queste strutture possano essere ritenute dei buoni approssimatori, per quale classe di funzioni e sotto quali condizioni possano essere utilizzate. In questo paragrafo si analizzeranno le proprietà delle HRBF in termini di unicità e stabilità della scomposizione, e di convergenza della approssimazione e della velocità con la quale viene raggiunta.

Queste caratteristiche vengono più facilmente indagate nel caso continuo e non limitato, ma i risultati possono essere estesi anche al caso discreto.

Unicità e Stabilità

Si può dimostrare [BMF99] che un insieme di gaussiane

$$G_j = \{g(\cdot - k\Delta x_j; \sigma_j) \mid k \in \mathbb{Z}\}$$

costituisce una base di Riesz per lo spazio V_j da esse generato:

$$V_j = \langle \{g(\cdot - k\Delta x_j; \sigma_j) \mid k \in \mathbb{Z}\} \rangle \quad (4.53)$$

Ciò implica che la scomposizione di una funzione $f(x) \in V_j$ attraverso una rete che utilizzi lo strato G_j (cioè uno strato con parametro di scala

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

σ_j) è unica e stabile. La stabilità della scomposizione dipende, nel caso delle HRBF, dal rapporto $\sigma_j/\Delta x_j$. Poiché tale rapporto è costante per tutte le scale (cioè per tutti gli strati di una rete HRBF), la scomposizione è uniformemente stabile in tutte le scale. Il rapporto $\sigma_j/\Delta x_j$ controlla di fatto il tasso di sovrapposizione delle gaussiane adiacenti dello stesso strato. Si dimostra che le scelte implementative effettuate per la scelta dei valori di σ_j e Δx_j garantiscono un elevato grado di ridondanza.

Approssimazione e Convergenza

La famiglia dei sottospazi $\{V_j\}$ associati alle HRBF *continue* (cf. §4.2.2) costituiscono un buon approssimatore, nel senso che per ogni funzione $f(x) \in L^2(\mathbb{R})$ è possibile scegliere un j sufficientemente elevato per il quale l'approssimazione alla scala j dista dalla funzione $f(x)$ (in norma L^2) meno di una soglia prefissata [BMF99]:

$$\forall f(x) \in L^2, \forall \epsilon > 0, \exists j : \|f(x) - a_j(x)\|_{L^2} < \epsilon \quad (4.54)$$

dove $a_j(x)$ è l'approssimazione ottenuta come:

$$a_j(x) = f(x) * g(x; \sigma_j) \quad (4.55)$$

Questo risultato è valido per funzioni che appartengono ad uno spazio L^p qualsiasi, e può essere dimostrato valido anche per una classe degli spazi di Banach.

Nel caso discreto, non si dispone della funzione $a_j(x)$ della (4.55), ma di una sua approssimazione. Questa approssimazione migliora via via che Δx_j diventa più piccolo, cioè al diminuire della scala (all'aumentare di j). Il che permette di formulare una proprietà di approssimazione analoga a quella per le HRBF continue espressa nella (4.54).

La proprietà di approssimazione precedente fa uso di un solo strato HRBF. L'algoritmo di apprendimento descritto nel paragrafo 4.2.3 agisce in modo differente: esso non va a scegliere la scala adatta a ricostruire la funzione con l'errore desiderato, ma procede per approssimazioni successive. Ad ogni passo, la funzione da approssimare cambia (è infatti il residuo ad essere considerato) e pertanto la approssimazione globale potrebbe anche non convergere ad una funzione sufficientemente vicina alla funzione che si vuole approssimare. La convergenza è tuttavia assicurata per le funzioni che:

- hanno un supporto compatto,
- sono C^∞

- hanno le derivate equilimitate (cioè esiste un M tale che per ogni n vale: $\|s^{(n)}(x)\|_\infty < M$).

Molte superfici reali rientrano in questa classe di funzioni.

4.2.6. Riformulazione HRBF in chiave digitale

L'algoritmo di configurazione di una rete HRBF descritto in §4.2.3 permette la costruzione di una rappresentazione HRBF di un insieme di dati campionati in modo non uniforme.

Se i dati a disposizione sono equispaziati, tuttavia, esso può essere notevolmente semplificato, fino ad assumere una struttura simile a quella dell'algoritmo per il calcolo della trasformata wavelet per una implementazione digitale efficiente.

In questo paragrafo, verrà descritta una riformulazione delle reti HRBF come risultato dell'applicazione di filtri in cascata. Ciò consente una più facile comparazione con le wavelet applicate mediante FWT. In particolare saranno confrontate le prestazioni di ricostruzione dei due paradigmi a fronte di perturbazioni nei loro coefficienti dovute alla rappresentazione numerica. Sebbene la gaussiana abbia un supporto infinito, il suo valore è trascurabile all'esterno di un intorno del suo centro, a causa del decadimento esponenziale. Il filtro FIR corrispondente ad una gaussiana può quindi essere implementato con accuratezza arbitraria trascurando i coefficienti inferiori di una soglia prefissata.

Utilizzando i caratteri in grassetto per denotare i vettori e i caratteri normali per gli scalari, i coefficienti del j -esimo livello di una rete HRBF possono essere calcolati come:

$$\mathbf{a}_j = \downarrow 2^{M-j}(\mathbf{r}_{j-1} * \mathbf{g}_j) \quad (4.56)$$

dove M è il massimo numero di strati della HRBF e \mathbf{r}_{j-1} è il residuo dello strato $(j-1)$ -esimo. I campioni del segnale $f(\cdot)$, \mathbf{f} , sono considerati il residuo per il primo livello, $\mathbf{r}_0 = \mathbf{f}$. Il residuo del livello j , \mathbf{r}_j , viene calcolato come differenza tra il residuo al livello $(j-1)$ e l'uscita del livello j della rete. Quest'ultimo vettore, la ricostruzione del livello j , \mathbf{f}_j , viene ottenuta come convoluzione del filtro gaussiano del livello considerato, \mathbf{g}_j , con i coefficienti, \mathbf{a}_j , intercalati con sequenze di elementi nulli di opportuna lunghezza:

$$\mathbf{f}_j = \uparrow 2^{M-j}(\mathbf{a}_j) * \mathbf{g}_j \quad (4.57)$$

Il filtro \mathbf{g}_j nella (4.56) è il filtro gaussiano del livello j -esimo, ottenuto campionando la gaussiana caratterizzata da σ_j — descritta nella (4.31) — con lo stesso passo di campionamento della funzione $f(\cdot)$. Tenuto

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

contando che $\sigma_{j+1} = \frac{1}{2} \sigma_j$, il filtro del livello j -esimo può essere ottenuto dal filtro gaussiano del livello precedente, utilizzando una relazione simile a quella della scaling function della MRA (§4.1.2):

$$g_{j+1,k}(x) = 2 g_{j,k}(2x) \quad (4.58)$$

Quindi, a meno di un fattore di scala, il filtro al livello $j + 1$, si ottiene sottocampionando il filtro al livello (j)-esimo:

$$\mathbf{g}_{j+1} = \downarrow 2(\mathbf{g}_j)$$

La procedura di configurazione della rete viene quindi iterata fino al soddisfacimento di un opportuno criterio, per esempio l'errore di ricostruzione inferiore ad una data soglia, o fino al raggiungimento del numero prestabilito di strati, M .

La ricostruzione della funzione \mathbf{f} viene operata sommando le ricostruzioni dei vari strati, \mathbf{f}_j , ottenute come descritto in (4.57). Bisogna notare che questa operazione può essere effettuata in parallelo, in quanto, a differenza della trasformata MRA, per ricostruire un dato livello, non è necessario disporre della ricostruzione del livello precedente.

Le figure 4.9 e 4.10 rappresentano gli schemi di analisi e sintesi descritti in questo paragrafo. Anche per questi schemi valgono le considerazioni fatte per gli schemi MRA a pagina 113: gli schemi in figura forniscono lo schema utilizzato per l'esperimento comparativo riportato in §4.3.2, nel quale vengono usati quattro livelli di scomposizione.

A differenza dell'algoritmo originale, questa riformulazione non contiene alcun criterio di inserimento delle unità, come per esempio quello descritto in (4.49). L'analogia operazione digitale è di facile realizzazione, ma non è stata inserita per facilitare la comparazione con l'implementazione digitale della MRA.

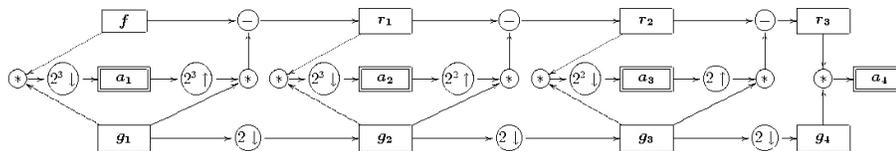


Figura 4.9.: Schema per una implementazione digitale della procedura di analisi HRBF

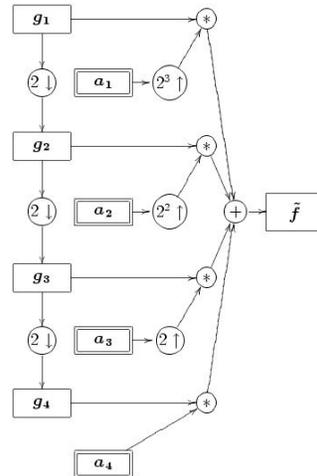


Figura 4.10.: Schema per una implementazione digitale della procedura di sintesi HRBF

4.3. Confronto tra wavelet e HRBF

Le tecniche multirisoluzione sono largamente usate nell'elaborazione del segnale grazie alla loro abilità nell'analizzare la proprietà del segnale e nel produrre una descrizione con caratteristiche di località sia nel dominio del tempo (o spazio) che della frequenza.

La scomposizione wavelet è molto usata come strumento per l'analisi multirisoluzione grazie alla rapidità di calcolo della FWT con la quale si possono calcolare i coefficienti wavelet.

Una interessante alternativa è rappresentata dalle HRBF, le quali, sebbene utilizzando uno schema diverso, forniscono anche loro una scomposizione multirisoluzione, con caratteristiche di approssimazione asintotica paragonabili al quelle delle wavelet (§4.2.5).

In questo paragrafo si intende confrontare questi due paradigmi in termini sia di caratteristiche di approssimazione che di robustezza alla perturbazione dei grandezze che intervengono nella ricostruzione.

Nel paragrafo §4.3.1 saranno comparate le caratteristiche dei due paradigmi dedotte dalla teoria dell'approssimazione. Nel paragrafo §4.3.2 si effettueranno delle prove di ricostruzione di funzioni con lo scopo di quantificare le caratteristiche dei due paradigmi in termini di accuratezza e robustezza. In particolare, verranno confrontate sia le implementazioni analogiche che quelle digitali.

4.3.1. Confronto teorico

Sebbene sia HRBF che MRA offrano una approssimazione multirisoluzione, esse operano in modo diverso. MRA scompone il segnale diminuendo il livello del dettaglio strato per strato, con il primo strato che costituisce i campioni misurati dal segnale. HRBF, invece, opera in modo opposto: la approssimazione meno dettagliata è ottenuta per prima e i dettagli vengono via via aggiunti al crescere del numero di strati utilizzati per la ricostruzione.

Velocità di convergenza Le wavelet generano (usualmente) spazi ortogonali a quelli delle scaling function della stessa scala e quindi la differenza tra l'approssimazione ed il residuo è massimizzata. Nelle reti HRBF, invece, gli spazi generati da strati successivi formano un angolo piuttosto piccolo e quindi la ricostruzione del residuo è più lenta. Per alcune applicazioni (ed entro certi limiti), questo fatto non è necessariamente una cosa negativa. Ciò, infatti, consente di suddividere su più livelli il contenuto informativo dei dati, ottenendo ricostruzioni che recuperano il dettaglio più gradualmente delle MRA. In questo modo è possibile raggiungere in modo più preciso il grado di accuratezza desiderato e risulta più facile evitare di inglobare il rumore nella ricostruzione.

La prova di ricostruzione di dati 3D (figure 4.15 e 4.19) è un buon esempio di come una velocità di convergenza meno spinta possa risultare più utile. L'arricchimento di dettaglio dei quattro livelli wavelet è graduale ed evidente. La MRA biortogonale 3.3, invece, presenta una ricostruzione iniziale (livello 1) nella quale tutti i tratti caratteristici del viso (occhi, naso e bocca) sono presenti. Già dal secondo livello sono bene delineati. La differenza tra il terzo ed il quarto livello è minima. Il primo livello, però risente fortemente dell'effetto di bordo, e gli altri tre presentano vistose oscillazioni dovute a rumore a bassa frequenza.

La velocità di convergenza può essere legata al rapporto tra i fattori di scala dei livelli consecutivi. Le MRA impongono che tale rapporto sia costante (usualmente pari a due). Con le HRBF è solo una scelta di comodo: nulla impedisce di utilizzare un rapporto diverso per ogni strato. Questa proprietà può essere sfruttata per inglobare più facilmente la conoscenza a priori difficilmente formalizzabile (per esempio, la differenza "visuale" tra dettaglio e rumore). Inoltre, nelle MRA il rapporto tra i fattori di scala corrisponde al tasso di sottocampionamento del cascade algorithm. Esso deve essere per forza di cose un intero. Il numero di livelli per le HRBF è virtualmente illimitato. Esso è fissato dall'algoritmo di apprendimento (errore, #unità, disponibilità di dati per le stime

della media locale). La ricostruzione può essere fermata quando il livello di dettaglio raggiunto è soddisfacente.

Le gaussiane rivestono il doppio ruolo di scaling function e wavelet. Esse vengono utilizzate, cioè, sia per approssimare che per esprimere il residuo. Malgrado ciò, è ragionevole apparentarle più alle scaling function che alle wavelet, in quanto non sono generalmente in grado di rappresentare il residuo con accuratezza assoluta, ma solo una sua approssimazione. A differenza delle MRA, inoltre, l'angolo tra gli spazi approssimazione di una HRBF non è nullo [BMF99], e questo consente ad uno strato a risoluzione più fine di catturare il residuo (figura 4.12).

Il cascade algorithm necessita, in fase di analisi, dei coefficienti dell'approssimazione a risoluzione maggiore per poter scomporre la funzione. Tali coefficienti sono il risultato della proiezione della funzione da approssimare sulle scaling function a risoluzione maggiore. Se i coefficienti iniziali sono esatti, la trasformata wavelet è in grado di scomporre e ricostruire la funzione in modo esatto. Eventuali errori presenti sui coefficienti iniziali non sono recuperabili: la ricostruzione risentirà sempre degli errori iniziali. Generalmente questo tipo di errore è contenuto e viene commesso utilizzando campioni equispaziati della funzione come coefficienti del primo livello.

La scomposizione operata dalle HRBF, invece, ha un errore iniziale molto elevato, dovuto alla proiezione della funzione sullo spazio approssimazione a risoluzione più grossolana. Gli errori nel calcolo della approssimazione ad un dato livello possono però essere recuperati dallo strato seguente, il quale è in grado di ricostruire con maggior accuratezza rispetto al livello precedente. Gli errori di calcolo dei coefficienti sono essenzialmente di due tipi:

- la proiezione della funzione sulla funzione base viene calcolata approssimandola con la media pesata dei valori in un intorno del centro della gaussiana;
- una gaussiana non viene inserita perché l'errore locale è mediamente basso.

La prima causa è paragonabile all'utilizzo, nelle wavelet, dei campioni al posto delle proiezioni. La seconda causa di errore non sempre è una cosa negativa: se si intende filtrare dati rumorosi o risparmiare risorse di calcolo, può essere una pratica vantaggiosa. L'eliminazione di funzioni base ritenute non necessarie è possibile sia nelle HRBF che nelle wavelet. In queste ultime è, però, possibile solo a posteriori. Il calcolo dei coefficienti nelle MRA mediante il cascade algorithm (§4.1.2) avviene

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

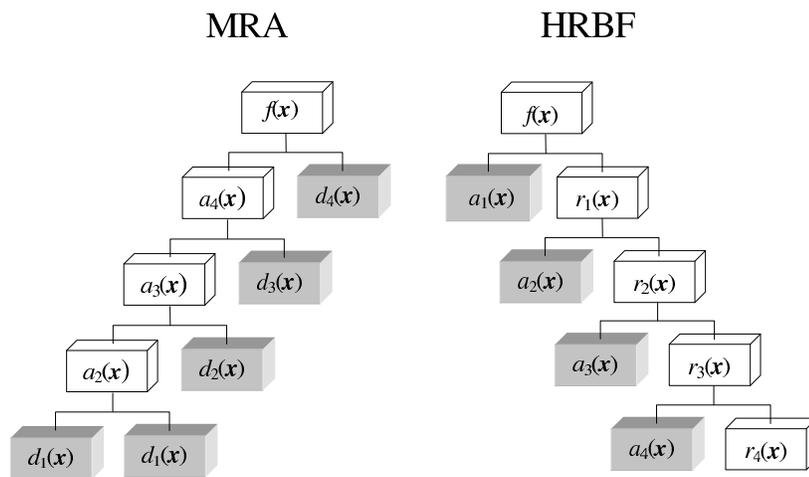


Figura 4.11.: Struttura *fine to coarse* della MRA e *coarse to fine* della HRBF.

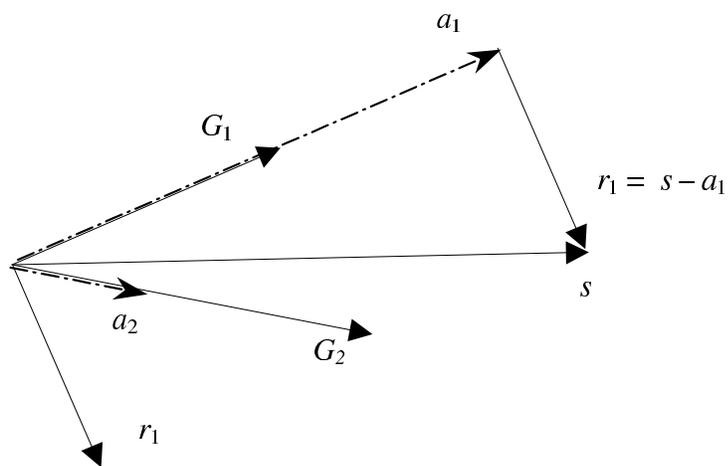


Figura 4.12.: Rappresentazione della ricostruzione HRBF mediante analogia con vettori del piano \mathbb{R}^2 : spazi non paralleli (G_1 e G_2) consentono la ricostruzione del residuo (r_1).

considerando solo i coefficienti della approssimazione ad un scala superiore. Risulta quindi difficile valutare l'impatto dell'eliminazione di un coefficiente. Nelle HRBF i coefficienti sono calcolati in base al residuo dello strato precedente, e quindi gli effetti dell'eliminazione (o meglio del non utilizzo) di alcune funzioni base possono essere propriamente valutati.

Formulazione digitale Molte funzioni base utilizzate per le MRA non possono avere una descrizione analitica e non possono quindi essere campionate in modo continuo. La funzione ricostruita può essere ottenuta iterativamente per approssimazioni successive, calcolandone il valore in corrispondenza di punti sempre più vicini (IFWT, §4.1.2). Lo spazio tra due punti consecutivi deve però essere interpolato, mentre l'uscita delle HRBF può essere calcolata direttamente. Per questo motivo la MRA è adatta ad una realizzazione digitale, mentre le HRBF può essere realizzata sia con una tecnologia digitale, sia con una tecnologia ibrida, dove i coefficienti sono memorizzati in modo digitale e le funzioni base sono calcolate analogicamente.

I filtri MRA sono generalmente più corti delle implementazioni FIR digitali delle approssimazioni a supporto finito dei filtri gaussiani, i quali, a causa della loro larga banda di transizione occupano almeno otto campioni [BMF99]. Inoltre, il numero di coefficienti di una scomposizione MRA è sempre pari a al numero dei campioni originali, mentre per una rappresentazione HRBF il numero di coefficienti è circa il doppio. Comunque, per le applicazioni reali, dopo aver posto a zero i coefficienti più piccoli, il numero di coefficienti significativi è generalmente lo stesso. La creazione di un'applicazione basata su HRBF è più semplice, in quanto richiede un solo filtro, caratterizzato da un unico parametro, il quale è legato da legge di proporzionalità al contenuto in frequenza delle informazioni che si vogliono elaborare, mentre MRA richiede la scelta di due coppie di filtri $(\mathbf{g}, \mathbf{h}, \tilde{\mathbf{g}}, \tilde{\mathbf{h}})$, una per l'analisi e una per la sintesi.

Per particolari applicazioni, HRBF può soffrire di mancanza di flessibilità: il (massimo) numero di livelli deve essere definito all'inizio, in quanto questo valore è in relazione con i parametri caratteristici del primo filtro gaussiano. Ciò si contrappone, nel caso MRA, alla possibilità di decidere il numero di strati run-time, semplicemente interrompendo la procedura di analisi. Infatti, sebbene il numero di livelli di una MRA sia limitato dal numero di campioni originali — partendo da N campioni si possono ottenere circa $\log_2 N$ livelli, generalmente i dati a disposizione sono in numero elevato e l'interruzione della procedura di analisi è dovuta a scelte progettuali.

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

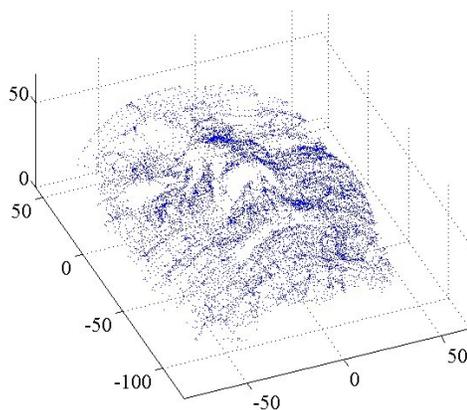


Figura 4.13.: Dati 3D.

Inoltre, il parametro σ delle gaussiane dell'ultimo livello non deve essere inferiore a 1.465 volte il passo di campionamento. Questa limitazione si propaga verso gli strati precedenti, moltiplicata per il rapporto tra i fattori di scala degli strati adiacenti.

4.3.2. Confronto sperimentale

Implementazione ibrida

Per confrontare le caratteristiche di HRBF (nella formulazione ibrida digitale/analogica) ed MRA come approssimatori, si proveranno i due paradigmi nella ricostruzione della superficie di un oggetto reale, a partire da punti 3D campionati su di essa, non equispaziati ed affetti da rumore.

I dati sono stati acquisiti con Autoscan [BF00] e rappresentano il volto di una persona (figura 4.13).

Le particolari condizioni di acquisizione (e la natura della superficie campionata) consentono di considerare le prime due componenti (X ed Y) come il dominio e la terza coordinata (Z) come il codominio di una funzione $\mathbb{R}^2 \rightarrow \mathbb{R}$. Tramite triangolazione sul piano XY o un'interpolazione spline cubica (figura (figura 4.14a-b) è possibile fornire una ricostruzione della superficie. È evidente che, a causa degli errori di misura presenti sui dati, entrambe le ricostruzioni sono poco soddisfacenti.

Il rumore ha nel caso in esame due fonti: l'incertezza della misura e piccoli spostamenti involontari del soggetto. In entrambi i casi è ragio-

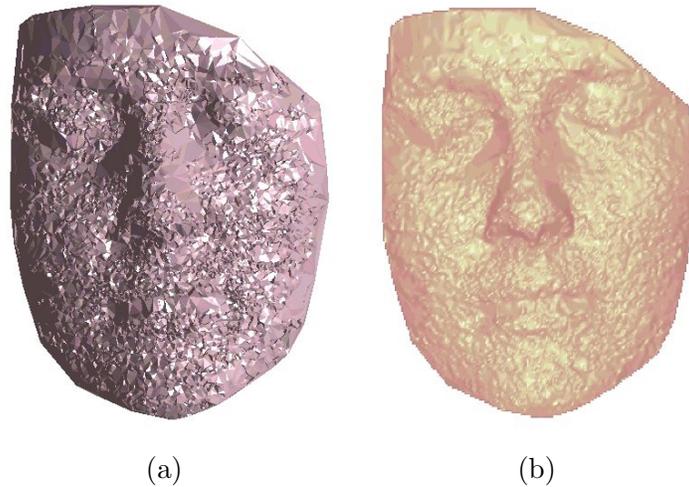


Figura 4.14.: Triangolazione di Delaunay (a) e interpolazione mediante spline cubica (b) dei dati in figura 4.13.

nevole pensare che le componenti in frequenza del rumore si manifestino principalmente su frequenze molto più elevate di quelle caratterizzanti la forma del viso. Il contenuto in frequenza della superficie da ricostruire, inoltre, varia notevolmente a seconda delle parti del viso rappresentate (per esempio la fronte ha un basso contenuto in frequenza, mentre le labbra presentano delle rapide variazioni della superficie). L'impiego di un modello multirisoluzione può quindi essere indicato per la ricostruzione di una superficie come quella in esame.

La ricostruzione ottenuta con le HRBF è riportata in figura 4.15. L'errore di misura (stimato come deviazione standard dell'errore di ricostruzione) è risultato di 0.8 mm. Per la ricostruzione, sono stati usati quattro livelli di HRBF, caratterizzati da parametro di scala, σ , fra i 16 mm ed i 2 mm. La qualità dell'output aumenta col crescere del numero di strati, principalmente nelle regioni con frequenza spaziale più elevata, come il naso, gli occhi e le labbra. Dalla figura 4.16, nella quale sono riportate le uscite dei singoli livelli, e dalla figura 4.17, che raffigura la posizione delle gaussiane effettivamente inserite, si può notare che il contributo degli strati a risoluzione più elevata è concentrato proprio su queste regioni.

I risultati quantitativi che mostrano la progressiva riduzione dell'errore di ricostruzione, sono riportati nella tabella 4.1.

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

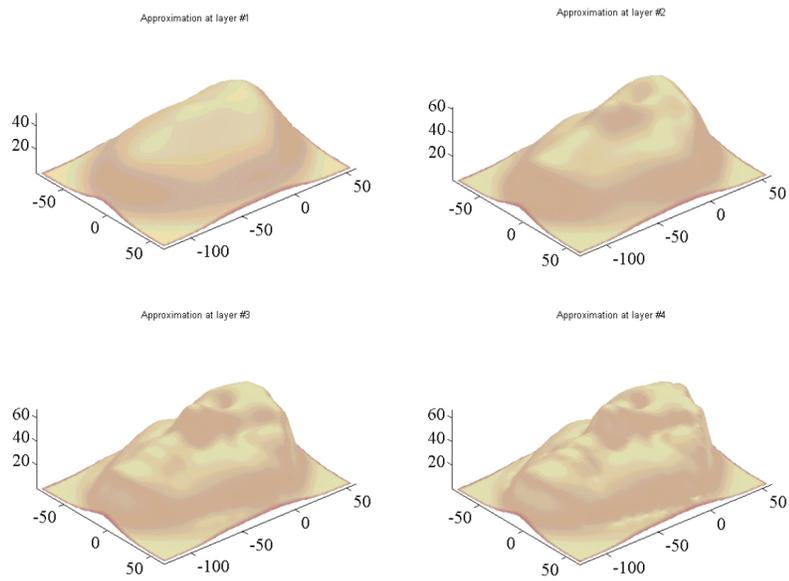


Figura 4.15.: Ricostruzione HRBF dei dati in figura 4.13

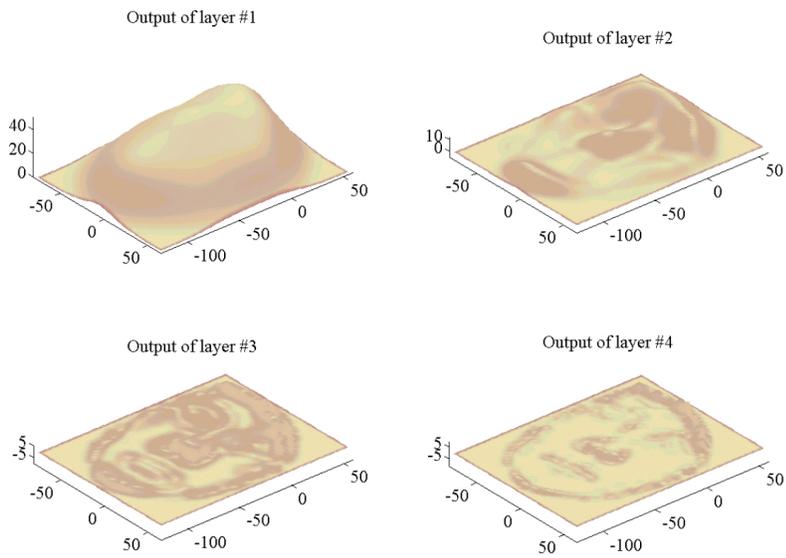


Figura 4.16.: Uscite dei singoli livelli HRBF

4.3. Confronto tra wavelet e HRBF

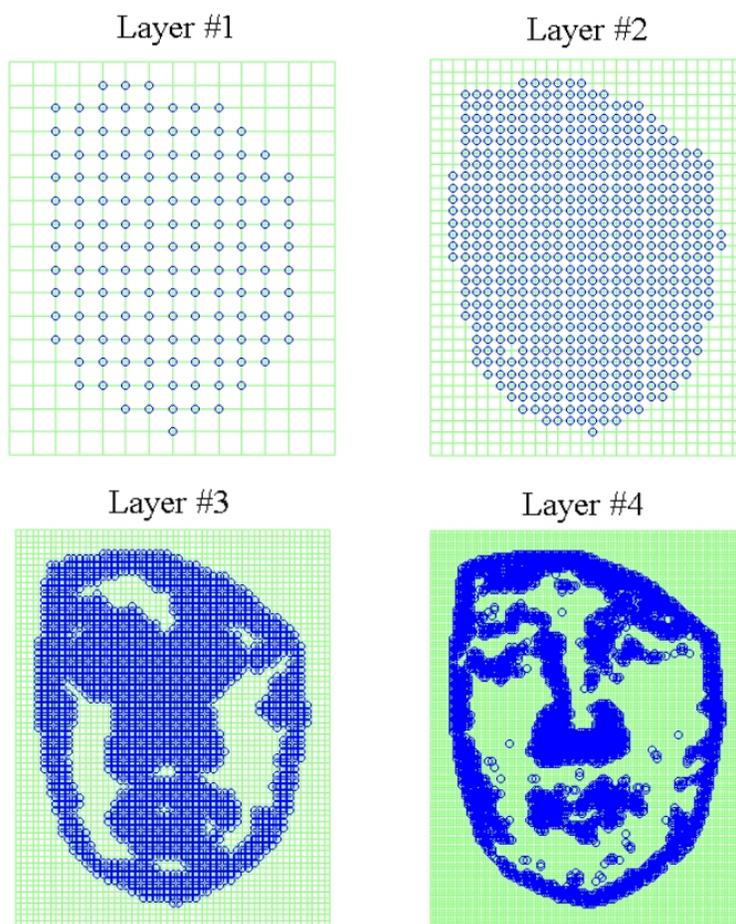


Figura 4.17.: Posizione delle gaussiane nei livelli HRBF

Tabella 4.1.: Cifre di merito della ricostruzione mediante HRBF

#Livelli	Errore medio [mm]	std errore [mm]	# di gaussiane	σ [mm]
1	2.33	4.38	141/270	16
2	0.03	2.06	573/980	8
3	0.04	1.01	1,780/3,795	4
4	0.06	0.76	3,637/14,933	2

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

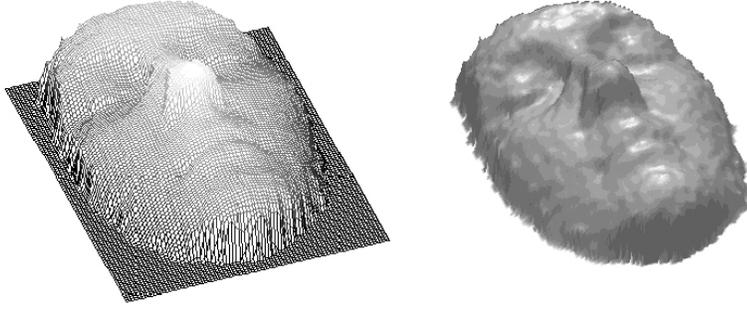


Figura 4.18.: Ricampionamento su griglia regolare

Bisogna notare che la ricostruzione del volto alla scala più piccola (2 mm) mediante un singolo strato HRBF avrebbe richiesto un totale di 14933 unità (109×137). L'approssimazione finale ha richiesto solo 6131 unità, distribuite nei quattro strati.

Per operare una ricostruzione MRA, i dati a disposizione non sono direttamente utilizzabili. L'insieme di dati è stato quindi elaborato mediante una procedura di ricampionamento tale da produrre punti 3D equispaziati nel dominio (il piano \mathbb{R}^2). La superficie è stata ricampionata in corrispondenza dei punti di intersezione di una griglia regolare di dimensioni 109×137 , le maglie della quale hanno lato, Δx , di 1.365 mm. Le dimensioni sono le stesse dell'ultimo livello HRBF. L'altezza della superficie è stata stimata utilizzando una procedura simile a quella utilizzata per le HRBF, 4.47, considerando un intorno sferico iniziale di raggio Δx . Se meno di cinque punti cadono nell'intorno sferico, il raggio viene aumentato finché almeno cinque punti vengano inclusi nella regione usata per la stima. Il risultato di questa procedura di ricampionamento è riportato nella figura 4.18.

I dati ricampionati sono stati quindi elaborati utilizzando la MRA biortogonale 3.3, le cui funzioni base sono riportate in figura 4.5, la quale è caratterizzata da filtri di lunghezza contenuta (otto elementi). Per poter operare un confronto più significativo, solo le 6131 wavelet con i coefficienti di valore più elevato sono stati tenuti. Gli altri sono stati posti a zero. Tutti i coefficienti approssimazione delle scaling function al primo livello sono stati tenuti. I risultati sono riportati in figura 4.19 ed in tabella 4.2

Risultati simili sono stati ottenuti anche limitando la selezione dei coefficienti da annullare ai primi due strati (quelli con contenuto in frequenza più elevato). Come può essere notato dalla figura 4.19, una ricostruzione smooth è ottenibile solo con gli ultimi due strati, mentre

4.3. Confronto tra wavelet e HRBF

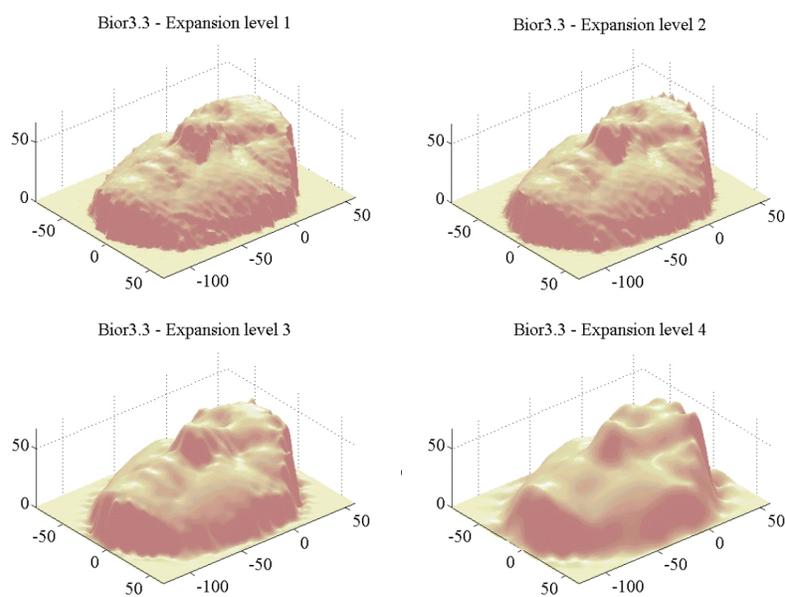


Figura 4.19.: Ricostruzione con MRA biortogonale 3.3

Tabella 4.2.: Cifre di merito relative alla ricostruzione con MRA biortogonale 3.3

#Livelli	Errore medio [mm]	std errore [mm]	# di basi	Δx [mm]
1	-4.37×10^{-5}	0.066	2 933/12 528	1.365
2	5.21×10^{-4}	0.24	2 021/3 744	2.73
3	0.019	0.76	720/1 311	5.46
4	0.15	1.47	285/585	10.92

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

gli ultimi due esibiscono un comportamento molto rumoroso. D'altra parte, gli ultimi due livelli non ricostruiscono i dettagli del volto (labbra ed occhi) con sufficiente accuratezza, e risentono molto dell'effetto del bordo.

Per evitare di considerare anche effetti estrapolativi della ricostruzione, le cifre di merito sono state calcolate prendendo in considerazione solo la regione interna al convex hull dei dati.

Implementazione digitale

La riformulazione delle HRBF in chiave digitale è stata qui utilizzata per provare la robustezza delle procedure di ricostruzione HRBF e MRA.

La qualità della ricostruzione del segnale dipende dalla accuratezza dei parametri determinati nella fase di configurazione. La robustezza è la capacità di un approssimatore di degradare le proprie prestazioni, quando i suoi parametri vengano perturbati, in modo proporzionale alla perturbazione stessa. La robustezza è una qualità importante soprattutto per le implementazioni digitali, dove le perturbazioni possono essere sia transitorie (disturbi elettromagnetici) che stabili (utilizzo di un numero ridotto di bit).

In particolare, nella prova in esame, è stato simulato l'effetto della quantizzazione che influisce sui calcoli in una implementazione di MRA e HRBF sia in virgola fissa (*fixed point*) che mobile (*floating point*).

La notazione in virgola fissa è usata in modo esteso nell'implementazione hardware per via della semplicità circuitale. L'uso di questa notazione implica — dato un prefissato numero di bit disponibili — un compromesso tra l'intervallo di valori rappresentabili e la risoluzione (cioè la distanza tra due numeri consecutivi). Per simulare l'implementazione a virgola fissa i valori dei parametri sono stati vincolati — tramite arrotondamento — ad assumere un numero finito di valori equispaziati. L'accuratezza della ricostruzione è stata valutata per i seguenti passi di quantizzazione, espressi come frazione del massimo valore assoluto dei vettori di parametri considerati: $\{\frac{1}{10000}, \frac{1}{1000}, \frac{1}{100}, \frac{1}{50}, \frac{1}{10}\}$.

La notazione a virgola mobile permette di rappresentare i dati con un prefissata accuratezza relativa. Nella rappresentazione a virgola mobile, i valori sono arrotondati in base al numero di bit dedicati alla mantissa ed all'esponente. Nella simulazione in esame, la restrizione sul numero di bit è stata limitata alla mantissa. Ciò equivale a fornire una precisione relativa del valore vero dei parametri, tramite troncamento. Sono stati usati da due a sei cifre decimali per la mantissa delle grandezze rappresentate.

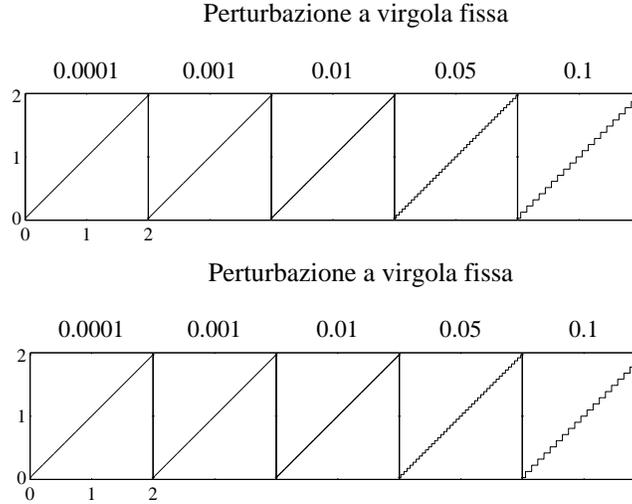


Figura 4.20.: Effetti delle notazioni in virgola fissa e mobile sulla rappresentazione di un segmento di linea retta

Nella figura 4.20 sono stati resi graficamente gli effetti della rappresentazione, usando le codifiche appena descritte, dei punti di un segmento lineare.

Ogni sessione di simulazione è stata caratterizzata dal tipo di perturbazione (a virgola fissa o mobile) e dall'insieme di parametri perturbato. Gli insiemi di parametri considerati in questo esperimento sono:

- i dati di input, \mathbf{f} ,
- i coefficienti dei filtri, \mathbf{g} per la HRBF e $\{\mathbf{g}, \mathbf{h}, \tilde{\mathbf{g}}, \tilde{\mathbf{h}}\}$ per la MRA.
- i coefficienti delle funzioni base, $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ e $\{\mathbf{a}_n, \mathbf{d}_1, \dots, \mathbf{d}_n\}$ per, rispettivamente, la HRBF e la MRA.

In ogni sessione, HRBF e MRA ricevono i dati di input \mathbf{f} e calcolano i coefficienti della rappresentazione (figura 4.21). I calcoli vengono eseguiti ad una precisione considerata infinita per gli scopi di questo esperimento. Con ciò si intende che solo gli effetti della rappresentazione numerica delle entità vengono presi in considerazione, mentre gli errori numerici dovuti ai calcoli — dato che questi sono stati eseguiti in doppia precisione — vengono ritenuti trascurabili.

L'errore di ricostruzione viene valutato su un differente insieme di dati, \mathbf{f}_V . \mathbf{f} e \mathbf{f}_V sono ottenuti campionando in modo equispaziato la

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

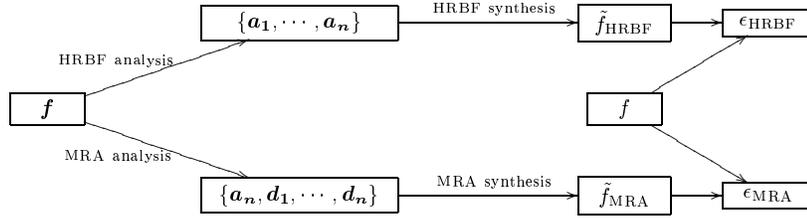


Figura 4.21.: Set-up per la prova di robustezza rispetto alla notazione usata

funzione $f(\cdot)$ riportata in figura 4.22 in 1 001 e 32 000 punti rispettivamente. A causa dell'elevata densità di campionamento usata per ottenerlo, f_V può essere considerato approssimazione ragionevole di $f(\cdot)$. Quindi, esso può essere usato per misurare l'errore di generalizzazione delle due ricostruzioni. Una volta ottenute le ricostruzioni HRBF, \tilde{f}_{HRBF} , e MRA, \tilde{f}_{MRA} , alla stessa frequenza di campionamento di f_V , l'errore di ricostruzione può quindi essere calcolato come differenza con segno:

$$\epsilon_{\text{HRBF}} = f_V - \tilde{f}_{\text{HRBF}}$$

$$\epsilon_{\text{MRA}} = f_V - \tilde{f}_{\text{MRA}}$$

Discussione dei risultati

L'accuratezza della ricostruzione è misurata come deviazione standard dell'errore di ricostruzione ϵ_{HRBF} e ϵ_{MRA} . Nelle prove riguardanti la MRA sono state utilizzate le wavelet biortogonali 3.7 (figura 4.6). Nelle figure 4.23 e 4.24 sono stati riportati i risultati ottenuti perturbando i tre insiemi di parametri nelle due modalità (virgola fissa e mobile).

Le HRBF resistono bene alla perturbazione dei vari insiemi di parametri. Questa robustezza è dovuta essenzialmente a due ragioni. Il primo motivo è che la MRA può perdere le proprietà di biortogonalità delle sue funzioni base. La seconda (e principale) ragione è che il calcolo del residuo nella procedura di configurazione HRBF per lo strato j (4.45) incorpora gli errori introdotti come output del livello precedente, sia che gli errori siano causati da filtri non perfetti, dai coefficienti a_j o dall'elaborazione stessa. Ciò permette la correzione degli errori nei livelli superiori. Solo gli errori introdotti nell'ultimo livello non possono essere corretti. Poiché il residuo decresce con il numero di livelli, il suo valore non è elevato il suo effetto sulla ricostruzione è di lieve entità.

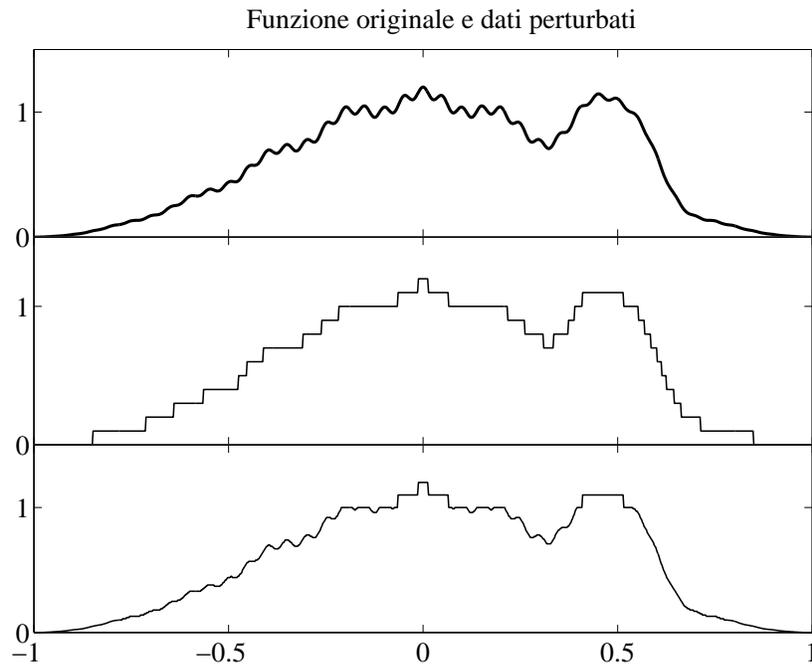


Figura 4.22.: La funzione utilizzata per le prove

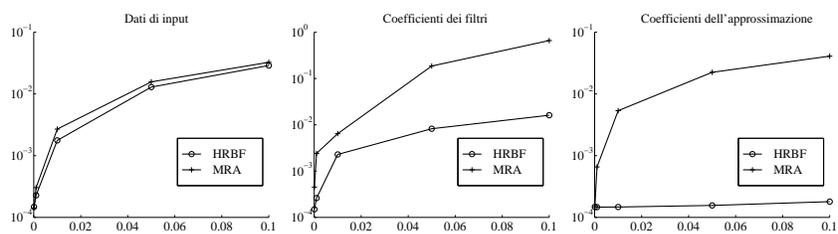


Figura 4.23.: Errore di generalizzazione nel caso di perturbazioni dovute alla notazione in virgola fissa

4. Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D

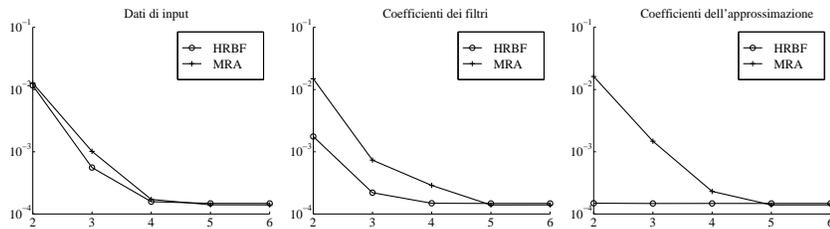


Figura 4.24.: Errore di generalizzazione nel caso di perturbazioni dovute alla notazione in virgola mobile

Questo non accade con la MRA, dove l'errore nei coefficienti si propaga attraverso il *cascade algorithm*. Inoltre, un fattore che contribuisce a contenere l'errore di generalizzazione è la ridondanza delle funzioni base. In ogni punto, più di una gaussiana contribuisce alla costruzione delle funzione. I valori dei coefficienti di gaussiane adiacenti sono, cioè, correlati. Quindi, se il rumore dovuto alla notazione è uniforme, gli errori tendono a compensarsi. Questo non accade per la MRA, dove l'ortogonalità delle funzioni base è sfruttata per ottenere dei coefficienti tra loro decorrelati.

4.4. Conclusioni

In questo capitolo sono stati presentati due paradigmi gerarchici di rappresentazione multirisoluzione di segnali: Multi-Resolution Analysis e reti Hierarchical Radial Basis Functions.

La trasformata wavelet è un utile strumento matematico per rappresentare una funzione in base al comportamento della funzione stessa in finestre spazio-frequenza di dimensione finita. La sua riformulazione in un dominio discreto di intervalli spazio-frequenza ha portato alla ideazione della MRA.

Una MRA consente di descrivere una funzione appartenente ad $L^2(\mathbb{R})$ a diversi gradi di dettaglio. In particolare essa fornisce, per una data grado di approssimazione, una scomposizione additiva di ogni funzione formata da una sua approssimazione alla scala desiderata e dal dettaglio che differenzia l'approssimazione dalla funzione originale. Queste funzioni (approssimazione e dettaglio) sono espresse come combinazione lineare delle basi degli spazi corrispondenti. Le basi hanno la proprietà di essere composte da copie traslate di un'unica funzione, la scaling function per gli spazi approssimazione e la wavelet per gli spazi dettaglio. Le basi di spazi a scala diversa si differenziano per un fattore di scala.

Grazie all'algoritmo FWT, illustrato nel paragrafo 4.1.2, è possibile ottenere i coefficienti della combinazione lineare a partire da una sequenza di campioni della funzione da rappresentare. Una MRA è pertanto adatta ad essere utilizzata per analizzare, elaborare e ricostruire un segnale a partire da un suo campionamento.

Basate su concetti afferenti sia alla teoria del filtraggio che delle reti neurali, le HRBF forniscono un valido strumento alla elaborazione ed alla ricostruzione di segnali. Una rete HRBF è in grado di ricostruire una funzione continua a partire da un insieme di campione della funzione stessa. Nessuna ipotesi è richiesta né sul tipo di campionamento (uniforme o sparso) né sulla sua accuratezza (eventuale presenza di rumore).

Una rete HRBF ricostruisce la funzione data mediante una approssimazione ad un grado di risoluzione grossolano (approssimazione operata dal primo strato HRBF) e una sequenza di approssimazioni del residuo (forniti attraverso i restanti strati).

Dal punto di vista teorico, una sequenza di strati HRBF costituiscono una struttura multirisoluzione in grado di approssimare con accuratezza arbitraria qualsiasi funzione appartenente ad uno spazio $L^p(\mathbb{R})$, con qualsiasi p . In particolare, uno strato HRBF costituisce una base (stabile) di Riesz per lo spazio da essa generato.

Le reti HRBF presentano quindi molte proprietà in comune con le MRA. Questi due paradigmi sono stati confrontati sia dal punto di vista teorico che in base a prove sperimentali, con lo scopo di evidenziare le loro capacità di ricostruzione e di robustezza alle perturbazioni dei vari fattori che intervengono nella ricostruzione. Le ricostruzioni basate su MRA sono risultate più adatte alla compressione e alla ricostruzione in condizioni ambientali caratterizzate da assenza di rumore, mentre le HRBF sono maggiormente robuste alle condizioni di rumore e si prestano ad una ricostruzione progressiva.

Per la costruzione di modelli digitali da *range data*, HRBF sembra essere uno strumento più adatto della MRA per una approssimazione multirisoluzione. Poiché l'algoritmo di configurazione HRBF fa uso di operazioni locali, può essere facilmente parallelizzato e implementato per un uso in tempo reale su una macchina dotata di hardware parallelo (per esempio FPGA).

Inoltre, HRBF richiede poche informazioni *a priori* ed è autostrutturante.

4. *Tecniche gerarchiche per l'analisi e la sintesi di immagini 3D*

5. Un esempio applicativo: Autoscan

In questo capitolo verrà descritto il sistema di digitalizzazione *Autoscan* [BFB⁺98] [BF00], il quale rappresenta un esempio di applicazione della metodologia studiata in questa tesi. La catena di elaborazione da esso implementata rientra nello schema generale per sistemi di digitalizzazione presentata in §3.1.

Le sue caratteristiche principali sono la portatilità e la flessibilità di utilizzo. Autoscan è costituito da quattro moduli principali. Il primo modulo è dedicato all'acquisizione degli insiemi di punti tridimensionali di una superficie mediante il rilevamento della posizione di un puntatore laser. Quando richiesto, un modulo ausiliario può essere affiancato per la rilevazione del colore della superficie. Il secondo modulo ricostruisce una superficie $3D$, filtrando il rumore di misura. Come spesso capita nella pratica, alcune applicazioni richiedono modelli poligonali. Un terzo modulo è quindi dedicato alla conversione della superficie continua in una superficie poligonale. Eventualmente, un quarto modulo provvede all'integrazione delle informazioni di colore con la superficie.

L'elaborazione necessaria per ottenere un modello finale con elevata accuratezza può essere realizzata in tempo reale utilizzando una opportuna scheda per PC di fascia bassa.

Il modulo di acquisizione della geometria ed il modulo ausiliario per l'acquisizione delle informazioni di colore producono una collezione di dati che possono essere utilizzati per la generazione di un modello tridimensionale dell'oggetto acquisito.

Il modulo di ricostruzione si occupa della generalizzazione dei dati geometrici, fornendo una superficie $2\frac{1}{2}D$ descritta mediante una rete HRBF. Il quarto modulo è dedicato alla registrazione delle informazioni di colore ed alla loro fusione con i dati geometrici. Essi costituiscono l'implementazione della fase di generazione del modello. Ad essi va affiancato un altro modulo (attualmente sostituito da operazioni semi-manuali) per la registrazione e la fusione delle informazioni geometriche provenienti da *viste* diverse. Questo modulo dovrebbe colloquiare

5. Un esempio applicativo: Autoscan

con il modulo di ricostruzione, fornendo topologia e trasformazioni di registrazione e fusione ed ottenendo superfici continue (possibilmente multirisoluzione).

La fase di ottimizzazione (e conversione) è realizzata dal modulo di poligonalizzazione della superficie. La sua attuale implementazione utilizza un algoritmo di quantizzazione vettoriale ottimizzato per dati con le caratteristiche dei modelli tridimensionali. Se necessario è in grado di operare un affinamento della rappresentazione *per-vertex color* del modello.

5.1. Fase di acquisizione

5.1.1. Acquisizione

Il modulo di acquisizione è costituito da un puntatore laser commerciale da 5 mW di potenza, un insieme di videocamere digitali, ognuna delle quali fornisce un'immagine (in toni di grigio) di 256×256 pixel con un *frame rate* di 100 frame/secondo, una scheda per elaborazione di immagini dedicata ed un computer di interfaccia.

Il processore della scheda dedicata all'elaborazione delle immagini stima la posizione del punto illuminato dal puntatore laser con una accuratezza, stimata sperimentalmente [FP90], di 0.1 pixel. L'elaborazione consiste essenzialmente in una cross correlazione con una maschera di 6 pixel (un template del punto luminoso) seguita da una soglia. Essa può quindi essere facilmente realizzato su una scheda *general purpose* (per esempio, una scheda FPGA), più economica, senza intaccare le prestazioni del sistema. La tecnica utilizzata per riconoscere il punto permette di raggiungere un elevato rapporto segnale/rumore (SNR) (pochi falsi punti rilevati) ed un'accuratezza sub-pixel (calcolando la posizione del punto luminoso come media dei pixel sopra soglia, pesando ogni elemento con il relativo valore di cross-correlazione).

La posizione 3D del punto proiettato dal laser è calcolato mediante procedure fotogrammetriche a partire dalla posizione del punto su due TV-camere [BRFP90]. Il sistema è calibrato riprendendo una barra rigida, di lunghezza conosciuta, avente agli estremi due marker sferici rifrangenti della stessa dimensione dei punti luminosi [BC00].

Il campionamento della superficie viene realizzato muovendo il puntatore laser manualmente (figura 5.2). Un riscontro visivo dell'operazione di campionamento effettuata viene fornito riportando (in tempo reale) su un monitor i punti acquisiti dal sistema.

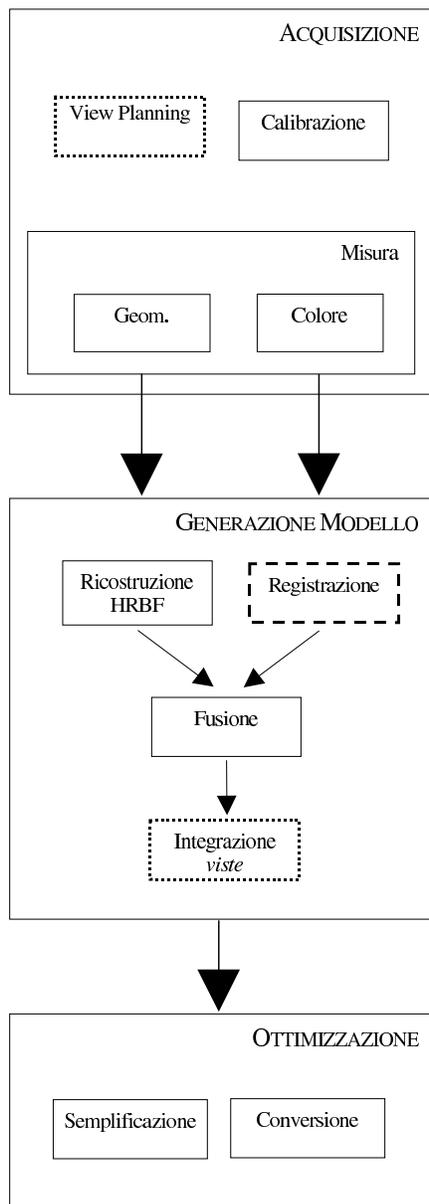


Figura 5.1.: Schema ad alto livello dello scanner tridimensionale Autoscan.

5. Un esempio applicativo: Autoscan

Alla fine del processo di campionamento, un insieme di punti tridimensionali è ricostruito per ogni coppia di TV-camere che condividano parte del loro campo visivo¹.

Allo stesso tempo, mediante una videocamera dotata di CCD a colori, un'immagine² dell'oggetto ripreso è acquisita e convertita in una bitmap di 576×768 (standard PAL) a 24 bit/pixel.

5.1.2. Accuratezza della acquisizione

L'accuratezza con cui Autoscan acquisisce i punti è stata valutata sperimentalmente, utilizzando un metodo simile a quello riportato in [GR99].

L'esperimento è costituito da una serie di 100 acquisizioni di punti proiettati dal laser su un piano posizionato al centro del campo di visione delle TV-camere. Per ogni acquisizione, composta da circa 1000 punti, è stato stimato il piano ottimale e calcolata la distanza (con segno) dai punti acquisiti da esso. Per un volume di dimensione media ($0.8 \text{ m} \times 0.6 \text{ m} \times 1 \text{ m}$), l'accuratezza rilevata, misurata mediante lo scarto quadratico medio (RMS), è stata di $0.27 \text{ mm} \pm 0.0063 \text{ mm}$. Questa cifra di merito varia linearmente con le dimensioni del volume, ed è consistente con le misure di accuratezza riportate nel campo dell'analisi del movimento [BC00]. La prova è stata effettuata con due TV-camere poste quasi ortogonalmente tra di loro (86.9°).

I risultati dell'esperimento sono riportati in figura 5.3. La distribuzione dell'errore di misura ha media nulla ed ha distribuzione normale, come è lecito attendersi, dal momento che è il risultato degli errori commessi in diverse fasi dell'elaborazione, sorgenti di errore di vario genere tra le quali: errore di quantizzazione delle TV-camere, approssimazioni numeriche nei calcoli e aspetto non uniforme e non circolare del punto luminoso.

L'accuratezza diminuisce quando l'angolo tra la coppia di TV-camere usata per la ricostruzione dei punti differisce dai novanta gradi ottimali seguendo una legge approssimativamente quadratica [BF90]. In situazioni reali, dove l'angolo tra le TV-camere è il risultato di un compromesso tra l'ampiezza del campo visivo comune e l'accuratezza, un angolo di $40^\circ \div 60^\circ$ è generalmente usato, con una perdita in accuratezza tra 1.5 e 3 volte. Se un'accuratezza maggiore è necessaria, sono necessarie un numero maggiore di sessioni di acquisizione, per diminuire il volume

¹Una *vista* è quindi costituita dall'insieme di punti raccolti da una coppia di TV-camere.

²Più immagini possono essere acquisite per ottenere un'immagine più affidabile, come media di immagini.



Figura 5.2.: Sessione di acquisizione con Autoscan.

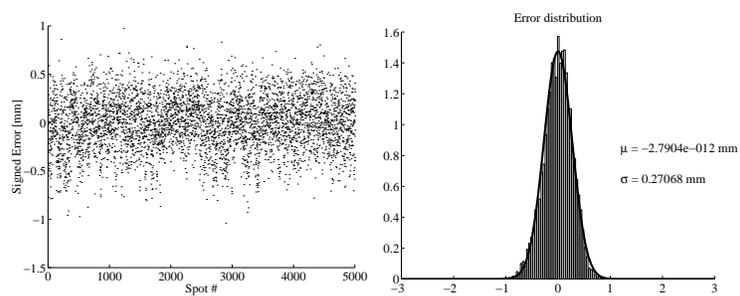


Figura 5.3.: Distribuzione dell'errore di misura.

5. Un esempio applicativo: Autoscan

del campo di ogni singola acquisizione o per aumentare l'angolo tra le TV-camere.

5.2. Fase di ricostruzione

5.2.1. Ricostruzione della superficie

La ricostruzione della superficie è ottenuta utilizzando una rete HRBF (*Hierarchical Radial Basis Functions* [BF00]). Le HRBF, descritte in maggiore dettaglio in §4.2, realizzano una approssimazione sparsa con risoluzione spaziale variabile e possiedono la capacità di filtrare il rumore in modo adattativo. Inoltre, essendo auto-organizzanti consentono di operare la ricostruzione in modo automatico, necessitando solo di alcune informazioni sulla qualità della ricostruzione desiderata dall'utente.

I punti raccolti da ogni coppia di TV-camere, formano una nuvola di punti tridimensionali. Generalmente, scegliendo come piano \mathbb{R}^2 di riferimento il piano ortogonale al piano di giacenza delle linee di visione delle camere e ortogonale alla bisettrice dell'angolo da esse formato, i punti di ogni *vista* possono essere interpretati come punti campionati da una superficie $\mathbb{R}^2 \rightarrow \mathbb{R}$ (§6). Essi possono quindi essere ricostruiti mediante HRBF. Essendo i punti campionati in modo non equispaziato, l'algoritmo utilizzato per la configurazione della rete è quello ibrido, riportato in §4.2.3.

Per l'esperimento di ricostruzione si è acquisito un volto umano. I dati raccolti da Autoscan sono riportati in figura 5.4.

Autoscan è stato spesso utilizzato per la ricostruzione di parti del corpo umano, volti in particolare. Quest'ultimo caso rappresenta un compito particolarmente difficile, in quanto il contenuto in frequenza è altamente variabile nelle diverse regioni del viso. Inoltre, si tratta di una superficie senza spigoli vivi. Quindi, le proprietà multirisoluzione delle HRBF ben si prestano per la ricostruzione di questa superficie.

I risultati della ricostruzione sono stati riportati in figura 5.5. È stata utilizzata una rete HRBF a quattro livelli. Nella prova, le TV-camere formavano un angolo di circa 50° , con una accuratezza stimabile intorno ai 0.65 mm. La superficie è stata ricostruita da 17 080 punti. Come la figura 5.6 riporta, negli strati più elevati le gaussiane sono inserite solo nelle regioni ricche di dettagli.

I risultati quantitativi sono riportati in tabella 5.1, dalla quale si può notare come l'errore abbia una media praticamente nulla, ed una deviazione standard che decresce all'aumentare dei livelli, fino a raggiungere il valore di 0.72 mm, che ben approssima la stima dell'errore di misura sui dati. Inoltre, come si può notare dalla figura 5.5, la distribuzione dell'er-

re assume via via una legge gaussiana, avvicinandosi alla distribuzione dell'errore di misura.

Registrazione

Se la sessione di acquisizione è relativamente lunga e se il soggetto della digitalizzazione può muoversi (come nel caso di acquisizione del volto di una persona), il movimento va rilevato e la misura corretta. A tale scopo, è possibile sfruttare le caratteristiche del modulo di acquisizione. Esso è stato infatti sviluppato originariamente nel campo del *motion capture* per rilevare la traiettoria di speciali marcatori rifrangenti. Per rilevare il movimento della testa è quindi sufficiente piazzare sul soggetto dell'acquisizione alcuni marcatori rifrangenti (almeno tre e di dimensione simile a quella del punto luminoso) e rilevarne la traiettoria seguite durante l'acquisizione. Questa tecnica è utilizzabile sia per eliminare piccoli movimenti non volontari nel caso di acquisizioni di un soggetto umano, sia per la registrazione di acquisizioni multisessione.

La correzione dei movimenti è infatti un caso particolare della registrazione di molteplici *viste*. In questo caso, ogni *frame* acquisito può essere considerato una *vista* dove il sistema di riferimento (individuato dai marcatori) è solidale con l'oggetto. I differenti *frame* hanno una posizione diversa rispetto al sistema di riferimento. Essa è però conosciuta in quanto identificabile a partire dai tre (o più) punti visti in ogni *frame*. È quindi possibile, per ogni coppia di TV-camere collasare l'asse temporale, dopo aver effettuato la registrazione rispetto ai marker di riferimento, ottenendo così la nuvola di punti corrispondente ad una *vista*.

Assegnamento delle caratteristiche cromatiche

Quando anche il colore dell'oggetto è importante, una (comune) videocamera a colori può essere utilizzata per acquisire le informazioni di natura cromatica.

La registrazione dell'informazione cromatica avviene in modo simile alla registrazione delle informazioni geometriche di *viste* differenti, e cioè calibrando la videocamera.

Attualmente, la procedura di calibrazione della videocamera (e quindi della registrazione) è di tipo semiautomatico [Rig00]. Essa è basata sull'identificazione sull'immagine digitale e sulla superficie di alcuni punti corrispondenti. Cinque punti sono sufficienti per stimare i parametri della videocamera quindi essere in grado di proiettare il colore dei pixel dell'immagine digitale sulla superficie. Le proprietà cromatiche del-

5. Un esempio applicativo: Autoscan

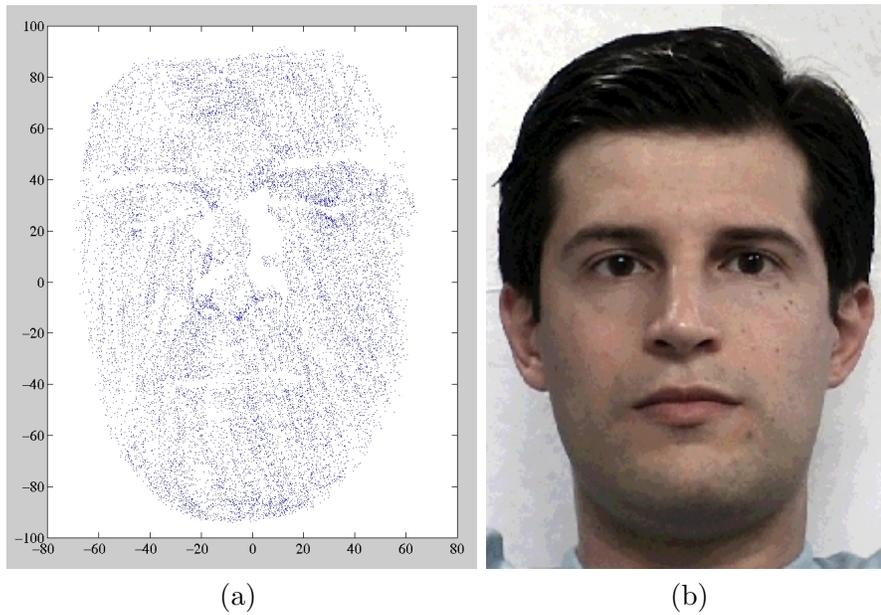


Figura 5.4.: Dati geometrici (a) e cromatici (b) acquisiti con Autoscan.

Tabella 5.1.: Risultati quantitativi della ricostruzione

#Livello	Errore medio [mm]	std errore [mm]	# di gaussiane	ω_{co} [Hz]	dimensione griglia
1	1.89	3.66	282/315	0.015	15 × 21
2	0.07	1.45	1 124/1 333	0.03	31 × 43
3	-0.005	0.96	3 020/5 394	0.06	62 × 87
4	-0.002	0.72	5 929/21 750	0.12	125 × 174

5.2. Fase di ricostruzione

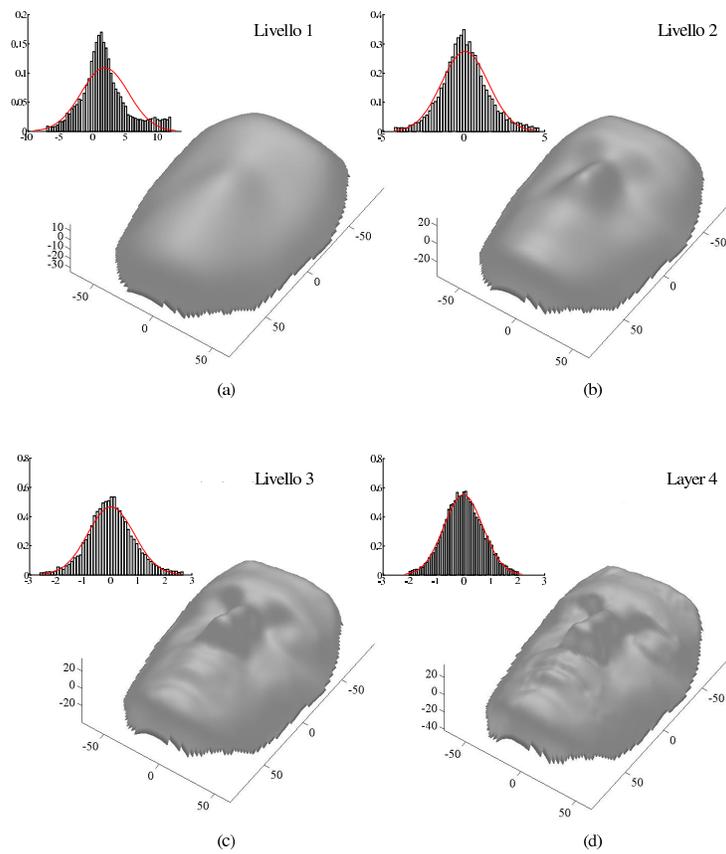


Figura 5.5.: Ricostruzione con HRBF a quattro livelli.

5. Un esempio applicativo: Autoscan

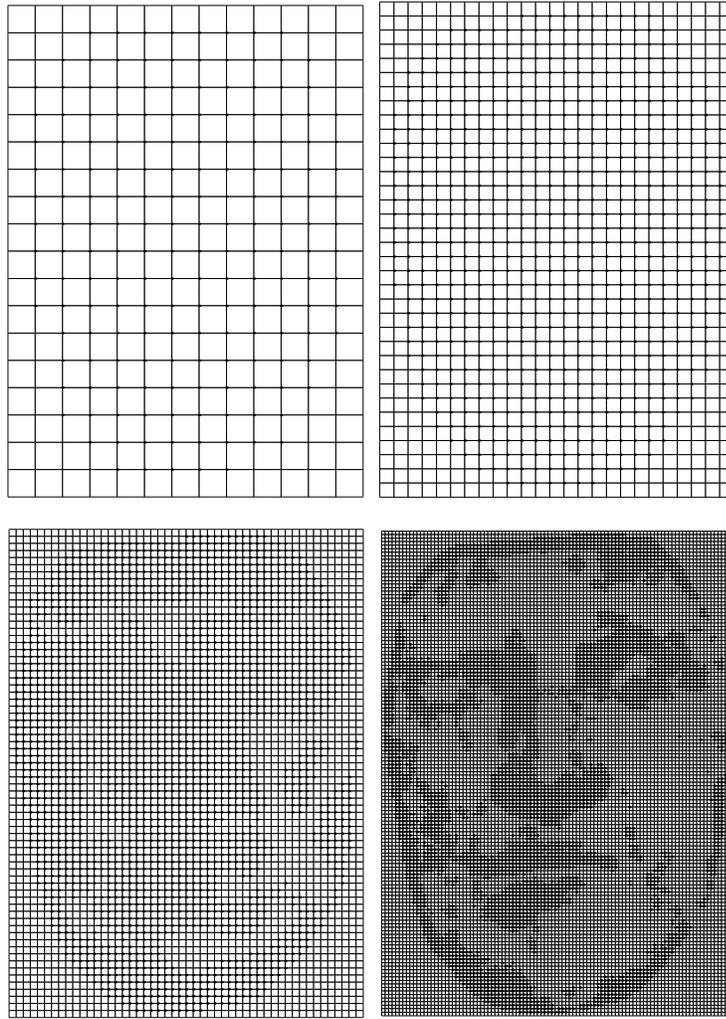


Figura 5.6.: Posizione delle unità utilizzate nei quattro livelli.

la superficie vengono assegnate ai vertici attraverso *texture mapping* o assegnando un colore a ciascuno dei vertici della mesh (§2.4.5).

5.2.2. Conversione in superficie poligonale

La ricostruzione HRBF della superficie fornisce una descrizione continua e infinitamente derivabile. Se una descrizione poligonale è richiesta dall'applicazione che utilizzerà il modello, bisogna applicare un ricampionamento della superficie ricostruita. Le modalità operative sono varie. Un semplice approccio consiste nel ricampionare la superficie HRBF con la risoluzione richiesta dal livello di dettaglio desiderato, e formare dei triangoli con i punti ricampionati adiacenti (in figura 5.7 a 1 mm di passo di campionamento). La mesh così ottenuta può essere quindi sottoposta ad una procedura di *mesh simplification* [CMS98] (si veda anche §2.5.1). Tuttavia, una procedura più raffinata può sfruttare la conoscenza analitica della curvatura della superficie per ottenere una mesh ottimale direttamente dalla descrizione HRBF [HG99].

5.3. Conclusioni

Autoscan è uno scanner altamente flessibile, scalabile ed accurato: può essere utilizzato per acquisire oggetti di dimensioni differenti con la stessa precisione relativa.

Le TV-camere dispongono di un'ottica intercambiabile e possono essere posizionate in modo completamente arbitrario, avendo come unico vincolo la copertura della regione da acquisire con almeno una coppia di TV-camere. Inoltre la procedura di calibrazione è semplice e può essere realizzata anche da personale senza particolari conoscenze tecniche.

La procedura di scansione manuale consente di intensificare la frequenza di campionamento spaziale laddove vi sia una maggiore ricchezza di dettaglio. Inoltre, l'utilizzo di una sorgente multipunto (più puntatori messi a pettine, un proiettore con pattern adeguato) al posto di un singolo puntatore laser può aumentare la velocità di campionamento.

La doppia elaborazione (sogliatura e cross-correlazione) utilizzata per il riconoscimento del punto luminoso rende particolarmente robusta l'operazione di misura. Ciò rende possibile l'utilizzo di Autoscan anche in condizioni ambientali non controllate.

Come tutti i dispositivi ottici, Autoscan non richiede alcun contatto con la superficie, ed è quindi (a meno di utilizzarlo per acquisire oggetti sensibili a fasci luminosi, anche se di bassa intensità) poco invasivo.

Il filtraggio del rumore di misura sui dati e la ricostruzione della superficie vengono compiuti in un singolo passo, mediante un algoritmo

5. Un esempio applicativo: Autoscan

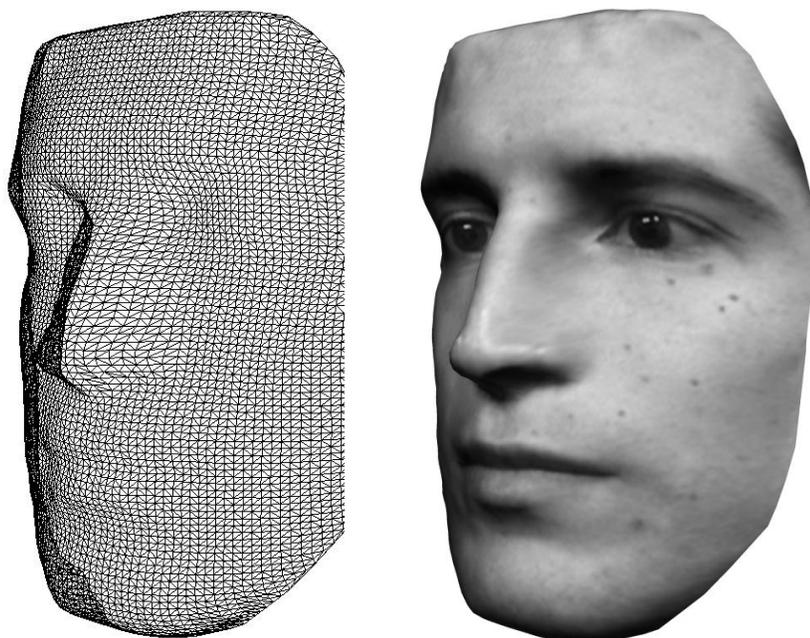


Figura 5.7.: Ricampionamento della superficie HRBF e assegnamento del colore tramite texture mapping.

adattativo. Il paradigma HRBF fornisce una descrizione multi-scala la cui procedura di configurazione può essere facilmente parallelizzata e realizzata in real-time.

Le operazioni sono attualmente effettuate in modo semiautomatico, ma la modularità dell'approccio consente di sviluppare opportunamente una fase dell'elaborazione lasciando invariate le altre.

Le apparecchiature di cui Autoscan è composto sono dispositivi di comune diffusione. Le procedure attualmente realizzate via software potranno essere realizzate su schede general purpose (per esempio, FPGA), abbassando ulteriormente il costo del sistema e migliorandone le prestazioni.

L'utilizzo del sistema non richiede particolari capacità tecniche o conoscenze specifiche.

Possibili sviluppi di Autoscan riguardano principalmente l'implementazione dei moduli dedicati alle attività previste dalla metodologia, ma non ancora presenti in Autoscan.

La registrazione di superfici provenienti da differenti viste può essere facilitata dalla presenza di marcatori sull'oggetto dell'acquisizione. Lo stesso principio utilizzato per la registrazione dei singoli *frame* è utilizzabile per fornire a differenti coppie di sensori un comune sistema di riferimento sul quale basare la registrazione dei dati provenienti da differenti sessioni. In generale, questo procedimento è utilizzabile anche per la registrazione di informazioni provenienti da differenti *viste*. Tuttavia, nel caso specifico di Autoscan, ciò equivale a registrare le differenti coppie di TV-camere. Questo compito è pertanto già espletato dalla procedura di calibrazione a barra.

Le HRBF realizzano una funzione $\mathbb{R}^2 \rightarrow \mathbb{R}$. Esse vengono infatti utilizzate per ricostruire i dati ricavati da una coppia di sensori, i quali costituiscono una *vista*. Opportune procedure di *zippering* (come ad esempio quella proposta in [DWJM98]) possono essere impiegate per integrare le superfici risultanti dalle diverse viste.

L'assegnamento del colore è attualmente una procedura che necessita dell'intervento dell'utente. Diverse tecniche sono utilizzabili per automatizzare questa procedura. Una possibilità è l'implementazione del riconoscimento dei marcatori utilizzati per la registrazione delle viste. In questo caso la procedura di integrazione sarebbe più organicamente strutturata. Alternativamente, tecniche di *computer vision* possono essere impiegate. A seconda della conoscenza a priori che si ha sull'oggetto dell'acquisizione, tecniche *ad hoc* (per esempio, riconoscimento di labbra, sopracciglia, narici) o tecniche di applicazione più generale (per esempio, *shape from shading*) possono essere impiegate per estrarre le *caratteristiche* da allineare con la superficie.

5. *Un esempio applicativo: Autoscan*

6. Conclusioni

La disponibilità di elaboratori equipaggiati con dispositivi a basso costo per la grafica 3D e la disponibilità di dispositivi di memorizzazione hanno reso realizzabili e largamente fruibili applicazioni basate sul trattamento delle informazioni e delle caratteristiche legate alla realtà tridimensionale. Queste applicazioni hanno però bisogno di essere dotate di modelli tridimensionali sui quali operare. Due sono gli approcci per la generazione dei modelli tridimensionali: la modellazione manuale e la copia digitale.

Benché sia fondato su tecniche fotogrammetriche, il problema della digitalizzazione di oggetti reali è quindi relativamente recente, evidenziato dal progresso tecnologico degli ultimi decenni e dal crescente mercato legato alla multimedialità. Gli approcci utilizzati ad oggi per ottenere una copia digitale di oggetti reali sono numerosi. Sebbene spesso essi siano sviluppati *ad hoc*, per la soluzione di problemi particolari, è possibile individuare alcune fasi dell'elaborazione comuni a tutti i procedimenti: acquisizione, ricostruzione, ottimizzazione.

Nella fase di acquisizione vengono effettuate le procedure per mettere a punto il sistema di misura e per effettuare le misurazioni. Queste procedure includono il posizionamento dei sensori e la loro calibrazione. Le procedure si differenziano in base al tipo di sensori utilizzati (cioè al fenomeno fisico da essi sfruttato per la misura) e dal metodo usato per la misura. Quest'ultimo è generalmente riconducibile ad una triangolazione, ma i diversi metodi si differenziano per le variabili indipendenti (un angolo, un lato). La fase di misura produce una collezione di dati che costituiscono un campionamento delle caratteristiche della superficie dell'oggetto acquisto.

Nella fase di generazione del modello tridimensionale i dati forniti dalla fase di acquisizione vengono generalizzati ed integrati tra loro. In particolare, questa fase si compone di tre sottofasi (ricostruzione, registrazione e fusione), tra loro cooperanti. Nella fase di ricostruzione, i dati delle singole viste vengono generalizzati per fornire una descrizione continua della superficie. La fase di registrazione, invece, individua la trasformazione da applicare ai dati per renderli relativi ad un unico si-

6. Conclusioni

stema di riferimento. La fase di fusione unisce le varie informazioni (geometriche e fotometriche), raccordandole. Questo blocco di elaborazione produce quindi un modello tridimensionale dell'oggetto acquisito.

La fase di ottimizzazione e conversione si occupa di trasformare il modello in un formato aderente alle caratteristiche richieste dall'applicazione obiettivo.

La ricerca condotta nell'ambito di questa tesi è stata rivolta principalmente verso la fase centrale di questa catena di elaborazione, ed in particolare verso la definizione di una metodologia per la generazione di un modello tridimensionale a partire da dati campionati da un oggetto reale. L'innovazione metodologica è stata qui formulata con l'intento di fornire un approccio strutturato e modulare, in grado di includere anche la maggior parte delle tecniche presenti in letteratura, ma anche sufficientemente flessibile da accogliere in modo organico le specificità di ogni applicazione reale.

La generazione del modello avviene per generalizzazione ed integrazione delle informazioni acquisite sull'oggetto. Essa è stata scomposta in sottofasi ben caratterizzate da uno scopo, dalle informazioni ad esse necessarie e da quelle che devono produrre. La generazione del modello tridimensionale avviene per approssimazioni successive applicando iterativamente le procedure di generalizzazione ed integrazione fino ad ottenere un unico modello.

Sebbene non sia strettamente richiesto dalla metodologia individuata, l'utilizzo di un paradigma gerarchico per la rappresentazione della superficie da ricostruire offre vantaggi che comportano un miglioramento delle prestazioni in tutte le fasi dell'elaborazione.

La rappresentazione multirisoluzione — intrinseca nei paradigmi gerarchici — consente di tarare il livello di dettaglio (ed il costo computazionale della rappresentazione) in base alla numerosità di dati ed alla necessità di ricchezza di dettagli. Le informazioni sono suddivise nei vari livelli di rappresentazione a seconda della scala alla quale si manifesta. Ciò consente di limitare le elaborazioni al livello di interesse per il compito da svolgere. In generale i paradigmi gerarchici manifestano anche una località spaziale. Ciò consente di economizzare le risorse di calcolo, impegnandole dove vi sia effettivamente informazione.

A tal proposito, è stato ideato un paradigma gerarchico, denominato HRBF, il quale possiede le caratteristiche appena esposte. Le sue proprietà sono state indagate sia in termini di accuratezza e robustezza della ricostruzione, che di economicità della realizzazione.

La validità della metodologia e dell'uso delle HRBF è stata provata con l'implementazione di un sistema di digitalizzazione tridimensionale, *Autoscan*. La sua struttura modulare ricalca quella dello schema gene-

rale studiato. Il modulo che si occupa della ricostruzione del modello da punti campionati utilizza HRBF come paradigma di rappresentazione della superficie.

6.1. Risultati ottenuti

Nell'ambito di questa tesi, è stata sviluppata una metodologia modulare e strutturata, in grado di considerare in modo organico i passi necessari ad affrontare il problema della generalizzazione di informazioni provenienti da differenti sorgenti. Essa è stata esposta nel capitolo 3. La definizione delle fasi elaborative ad alto livello di astrazione consente di inglobare le tecniche finora proposte in letteratura, senza che ciò costituisca un limite a possibili future innovazioni. Per ogni fase elaborativa di cui la metodologia è composta, sono stati individuati gli elementi caratteristici. Essi consentono di definire le proprietà dei paradigmi utilizzati per la rappresentazione del modello tridimensionale che possono semplificare e rendere efficienti le elaborazioni da svolgere.

Grazie alle loro proprietà multirisoluzione, i paradigmi gerarchici sembrano i più adatti allo scopo.

Allo scopo è stato sviluppato un paradigma gerarchico, chiamato HRBF, del tipo *coarse to fine*. Esso è stato presentato in §4.2. È stato inoltre ideato un algoritmo di configurazione di un modello HRBF a partire da dati non equispaziati ed affetti da rumore. Esso quindi ben si presta ad essere utilizzato nella metodologia studiata.

Le proprietà di HRBF sono state confrontate con quelle del paradigma MRA, basato sulla trasformata wavelet. L'utilizzo di MRA è molto diffuso e le sue proprietà sono ormai ben conosciute nella comunità scientifica. Tuttavia, rispetto alle MRA — quantomeno nell'ambito dei problemi affrontati in questa tesi — le HRBF presentano alcuni vantaggi.

La ricostruzione *coarse to fine* consente di modulare lo sforzo computazionale richiesto nella generazione del modello in base alla risoluzione effettivamente necessaria.

Le HRBF consentono di variare più flessibilmente i parametri della ricostruzione multiscala. Infatti, il parametro di scala di ciascuno strato delle HRBF non deve necessariamente essere dimezzato come nelle MRA, ma può essere adattato al contenuto in frequenza dei dati. Inoltre, se una norma differente da quella in L^2 è richiesta, è sufficiente cambiare l'espressione dell'errore nella formula utilizzata nella valutazione del residuo, (4.49). Per esempio, se l'insieme di dati contiene punti che in

6. Conclusioni

realtà non appartengono alla superficie (*outliers*), una norma più adatta alla misura del residuo può essere L^1 [BF99].

La rappresentazione HRBF fornisce una descrizione continua. In ogni punto del dominio di definizione può essere calcolata la posizione del corrispondente punto della superficie. Il calcolo può avvenire in modo diretto, a differenza di quanto accade con una rappresentazione MRA, dove è generalmente necessario ricorrere ad approssimazioni successive ed, in definitiva, ad una interpolazione.

Così come per MRA, anche l'algoritmo di configurazione e ricostruzione HRBF può essere riformulato in chiave digitale, nell'ipotesi di dover elaborare solo dati equispaziati. La formulazione digitale sfrutta la convoluzione (operata tramite FFT) per operare il calcolo dei coefficienti HRBF o della ricostruzione dei dati in modo rapido e robusto. Le prove effettuate hanno dimostrato che HRBF ha prestazioni comparabili a quelle di MRA, in condizioni normali, ma che esibisce una maggiore robustezza a fronte di perturbazioni delle grandezze che intervengono nell'elaborazione.

Il maggiore limite delle HRBF — limite condiviso con molti paradigmi per l'approssimazione di funzioni — appare essere l'incapacità, dovuta alle proprietà di filtraggio, di ricostruire brusche variazioni nella superficie. Ciò rende difficoltosa la ricostruzione della superficie di oggetti reali dotati di spigoli ben definiti ed angoli acuti. Per oggetti di questo genere, una procedura di ricostruzione basata su una elaborazione intermedia che si occupi di estrarre primitive geometriche può fornire una ricostruzione più fedele [HDD⁺94] [Hop94] [DTM96].

Autoscan risulta essere una implementazione della metodologia tracciata in questa tesi. Ad esso è stato dedicato il capitolo 5, nel quale sono stati riportati gli esperimenti eseguiti per valutare le sue caratteristiche. Le prove effettuate hanno quantificato la sua accuratezza sia in fase di acquisizione che in fase di ricostruzione del modello. Il paradigma di rappresentazione della superficie utilizzato è stato HRBF, il quale consente di ricostruire una superficie a partire dal suo campionamento. Autoscan è quindi in grado di fornire una dettagliata ricostruzione tridimensionale delle superfici, utilizzando una elaborazione adattativa dei dati. Le operazioni impiegate sia nel rilevamento del punto luminoso che nella ricostruzione della superficie sono realizzate analizzando i dati solo localmente. Esse possono quindi essere facilmente parallelizzate ed implementate su schede *general purpose* (per esempio, FPGA), migliorando il sistema in termini di prestazioni, di portatilità, e di costo.

6.2. Spunti per ricerche future

Una rappresentazione tramite HRBF fornisce una descrizione continua della superficie. Nei casi pratici, però, come illustrato nel capitolo 2, una descrizione poligonale è richiesta da molte applicazioni. La HRBF può essere ricampionata alla frequenza desiderata, grazie alla continuità della sua descrizione. Tuttavia, il numero di poligoni generato con un campionamento uniforme non è solitamente ottimale. Sebbene una procedura di ottimizzazione (come quelle descritte in 2.5.1) possa essere applicata alla mesh generata dal campionamento uniforme operato alla massima frequenza desiderata, una procedura più raffinata può sfruttare la conoscenza analitica della curvatura della superficie per ottenere una mesh ottimale direttamente dalla descrizione HRBF [HG99].

Analogamente, la descrizione HRBF può essere sfruttata per individuare le *caratteristiche* della superfici, come previsto dalla metodologia. Un'implementazione veloce delle HRBF dovrebbe consentire anche la ricerca di una opportuna *signature* di una superficie in un'altra. Una procedura del genere troverebbe immediata applicazione nella fase di registrazione, ma potrebbe essere sfruttata anche per realizzare *morphing* tra diverse superfici HRBF, o il riconoscimento tramite *pattern matching*.

A tal proposito val la pena di notare che le derivate di una combinazione lineare di esponenziali sono ancora combinazioni lineari di esponenziali. Questo significa che le proprietà differenziali delle HRBF possono essere calcolate in contemporanea alla superficie, riutilizzando parzialmente i risultati per il calcolo della superficie. Una definizione delle *caratteristiche* basata sull'analisi delle derivate della superficie (e delle grandezze ad esse legate) potrebbe essere realizzata in modo, in modo economico.

Le gaussiane hanno un supporto infinito. Il calcolo dell'uscita della rete, quindi, coinvolge (almeno in linea teorica) tutte le unità di cui la rete è composta. Dato il rapido decadere della gaussiana (località), è però generalmente possibile accettare un'approssimazione per cui oltre una certa distanza dalla posizione del centro della gaussiana stessa, il suo valore viene ritenuto nullo. Se i dati sono organizzati in una adeguata struttura di ricerca (per esempio il quadtree), questa approssimazione consente di ottenere enormi vantaggi a livello computazionale. Prove preliminari hanno mostrato un aumento di un ordine di grandezza delle prestazioni delle procedure di configurazione e di ricostruzione, senza variazioni sostanziali dell'accuratezza.

L'approssimazione appena citata opera per troncamento della gaussiana. Se ciò può essere ritenuto accettabile nella maggior parte dei casi

6. Conclusioni

nella pratica, dal punto di vista teorico questo equivale ad introdurre delle componenti ad alta frequenza. Ciò diviene più evidente quanto più il troncamento è operato vicino al centro della gaussiana. In generale può essere rendersi visibile qualora nella rappresentazione HRBF vi sia una unità con il coefficiente molto più elevato delle unità ad essa vicine. Per evitare questo effetto è necessario che la gaussiana venga raccordata a zero oltre una certa distanza dal centro. Questa pratica non è innovativa, in quanto è già stata applicata per velocizzare il calcolo di reti neurali ad unità gaussiane [Pla91] [ADRY94] [AR95]. Nei lavori citati, la funzione gaussiana viene sostituita con spline approssimanti la forma della gaussiana.

Questa ulteriore approssimazione potrebbe migliorare le senz'altro le prestazioni di velocità di calcolo delle procedure che coinvolgono la HRBF. Le prestazioni di accuratezza non verrebbero modificare sensibilmente, almeno nella maggior parte dei casi di interesse pratico. Dal punto di vista teorico, però, questa soluzione implementativa ha il pregio ulteriore di aprire nuovi interrogativi.

Oltre alla spline, infatti, altre funzioni potrebbero essere utilizzate per la multirisoluzione *coarse to fine*. Alcune proprietà delle HRBF sono però state derivate direttamente da proprietà della gaussiana. Per esempio, l'estensione dal caso monodimensionale a multidimensionale è stato basato sulla proprietà di scomposizione della gaussiana. Per altre funzioni questa proprietà non vale più. Una formulazione più generale del paradigma HRBF tale da comprendere altre famiglie di funzioni può essere quindi estremamente interessante.

Su questa linea di ricerca, può innestarsi anche un'estensione sulla traccia delle wavelet a M -bande. Esse consentono alla approssimazione MRA di catturare differenti caratteristiche del segnale attraverso l'uso di differenti wavelet. Le HRBF, come è già stato noto, non sono in grado di ricostruire superfici di oggetti che presentano spigoli vivi. L'utilizzo di unità di natura mista dovrebbe rendere possibile la rappresentazione di tali superfici.

Altri spunti di ricerca possono derivare dall'analisi dei progressi ottenuti nel campo delle MRA. Il *Lifting Scheme* ha permesso di operare la trasformazione wavelet su punti con una struttura non uniforme ed una topologia arbitrariamente complessa. Esso ha comportato per le wavelet la rinuncia alla somiglianza tra le funzioni base ed all'uso della trasformata di Fourier come strumento per il calcolo della trasformata wavelet. Un analogo procedimento potrebbe essere applicato alle HRBF. Ciò comporterebbe la perdita della radialità delle gaussiane. Inoltre, la sostanziale equivalenza di una rete di gaussiane equispaziate con la convoluzione con un filtro gaussiano ha permesso l'ideazione di un algoritmo

di configurazione delle HRBF senza iterazioni (come invece è necessario per altri algoritmi dell'area neurale). L'impossibilità di invocare la trasformata di Fourier per validare la procedura di configurazione impone una più approfondita analisi della natura della procedura stessa.

Le *Normal mesh* codificano una superficie mediante un modello gerarchico nel quale uno strato fornisce la topologia allo strato superiore. La disponibilità di un algoritmo di ottimizzazione poligonale per le HRBF potrebbe consentire lo sviluppo di una formulazione "normale" delle HRBF. L'approssimazione poligonale di uno strato potrebbe essere infatti usata come descrizione della topologia dello strato seguente.

Possibili sviluppi di Autoscan riguardano principalmente l'implementazione dei moduli dedicati alle attività previste dalla metodologia, ma non ancora presenti in Autoscan.

La registrazione di superfici provenienti da differenti viste può essere facilitata dalla presenza di marcatori sull'oggetto dell'acquisizione. Lo stesso principio utilizzato per la registrazione dei singoli *frame* è utilizzabile per fornire a differenti coppie di sensori un comune sistema di riferimento sul quale basare la registrazione dei dati provenienti da differenti sessioni. In generale, questo procedimento è utilizzabile anche per la registrazione di informazioni provenienti da differenti *viste*. Tuttavia, nel caso specifico di Autoscan, ciò equivale a registrare le differenti coppie di TV-camere. Questo compito è pertanto già espletato dalla procedura di calibrazione a barra.

Le HRBF realizzano una funzione $\mathbb{R}^2 \rightarrow \mathbb{R}$. Esse vengono infatti utilizzate per ricostruire i dati ricavati da una coppia di sensori, i quali costituiscono una *vista*. Opportune procedure di *zipping* (come ad esempio quella proposta in [DWJM98]) possono essere impiegate per integrare le superfici risultanti dalle diverse viste.

L'assegnamento del colore è attualmente una procedura che necessita dell'intervento dell'utente. Diverse tecniche sono utilizzabili per automatizzare questa procedura. Una possibilità è l'implementazione del riconoscimento dei marcatori utilizzati per la registrazione delle viste. In questo caso la procedura di integrazione sarebbe più organicamente strutturata. Alternativamente, tecniche di *computer vision* possono essere impiegate. A seconda della conoscenza a priori che si ha sull'oggetto dell'acquisizione, tecniche *ad hoc* (per esempio, riconoscimento di labbra, sopracciglia, narici) o tecniche di applicazione più generale (per esempio, *shape from shading*) possono essere impiegate per estrarre le *caratteristiche* da allineare con la superficie.

Le eventuali innovazioni nel campo HRBF avrebbero immediata applicazione nella realizzazione di nuove versioni di Autoscan. La procedura di poligonalizzazione ottimale potrebbe, per esempio, essere utiliz-

6. Conclusioni

zata nel modulo di ottimizzazione e conversione, mentre l'estrazione di *caratteristiche* ha una immediata applicazione nella realizzazione di una procedura automatica di registrazione.

Glossario

α -shapes L' α -shape è una generalizzazione della triangolazione di Delaunay [BBX95]. L' α -shape di un insieme di punti P si ottiene dalla sua triangolazione di Delaunay rimuovendo gli elementi (spigoli, triangoli e tetraedri) che non siano inscrivibili in una sfera di raggio α (figura 2.18).

Acquisizione Procedura mediante la quale vengono misurate le caratteristiche geometriche e cromatiche di un oggetto reale. Generalmente questa fase è realizzata per campionamento.

Color field Il *color field* è una funzione che associa ad un punto della superficie il colore che essa possiede.

Convex hull Il *convex hull* di un insieme di punti $P \subset \mathbb{R}^D$ è l'intersezione dei sottoinsiemi convessi di \mathbb{R}^D contenenti P . Per un insieme finito di punti in \mathbb{R}^2 (\mathbb{R}^3) è il poligono (poliedro) convesso di minore area (volume) che contiene i punti dati. In particolare i vertici del convex hull sono un sottoinsieme di P .

Diagramma di Voronoi Il diagramma di Voronoi di un insieme di punti $P \subset \mathbb{R}^D$ è la partizione dello spazio indotta da P in regioni tali che la regione corrispondente a $p \in P$ contiene tutti i punti di \mathbb{R}^D più vicini a p che ad ogni altro punto di P . È legato da una relazione di dualità alla triangolazione di Delaunay, in quanto ogni vertice del diagramma di Voronoi coincide con il centro di un tetraedro (in $3D$) della triangolazione di Delaunay corrispondente. Uno spigolo del diagramma corrisponde a due facce adiacenti della triangolazione.

Digitalizzazione Insieme di procedure necessarie a generare un modello (tridimensionale) di un oggetto reale. La digitalizzazione è costituita dalle fasi di acquisizione, ricostruzione e ottimizzazione.

Genus A topologically invariant property of a surface defined as the largest number of nonintersecting simple closed curves that can be drawn on the surface without separating it. Roughly speaking, it is the number of holes in a surface.

Immagine tridimensionale Rappresentazione di un oggetto, mediante un opportuno dispositivo di visualizzazione, tale da suscitare nell'osservatore la sensazione di tridimensionalità (e.g, fedeltà della prospettiva, dell'ombreggiatura).

Manifold A manifold is a topological space which is locally Euclidean (i.e., around every point, there is a neighborhood which is topologically the same as the open unit ball in \mathcal{R}^n).

Marching cubes Marching Cubes [LC87] è un algoritmo per l'approssimazione di isosuperfici in dati volumetrici (figura 2.17). Ogni cubo viene caratterizzato dai valori dei voxel che ne costituiscono i vertici. Se, data una soglia definita dall'utente, un cubo ha alcuni vertici sopra soglia e alcuni vertici sotto soglia, significa che l'isosuperficie di valore pari alla soglia attraversa la superficie. Determinando quali vertici del cubo sono intersecati dall'isosuperficie, è possibile creare patch triangolari che suddividono il cubo in regioni interne ed esterne all'isosuperficie. La rappresentazione dell'isosuperficie viene quindi generata connettendo tutte le patch triangolari così ottenute.

Modello tridimensionale Rappresentazione numerica della struttura geometrica e delle caratteristiche cromatiche (colore, tessitura, ombra, trasparenza) di un oggetto.

Octree Un *octree* è una struttura dati ad albero dove ogni nodo ha otto figli. Ogni nodo di un octree rappresenta un cubo nello spazio tridimensionale. Ogni figlio rappresenta un ottante del nodo-padre. È una struttura dati utilizzata per la ricerca dei punti appartenenti ad una regione in quanto consente un accesso logaritmico ai dati dell'insieme. Analogamente, il *quadtree* è una struttura dati utilizzata per organizzare i dati in uno spazio bidimensionale.

Omeomorfismo Un omeomorfismo è un'applicazione biunivoca, continua, e con inverso continuo; due insiemi omeomorfi hanno a comune tutte le proprietà che dipendono solo dalla topologia (ma non necessariamente proprietà metriche, come il volume). In generale, l'omeomorfismo

ad una circonferenza e' un carattere che accomuna tutte le figure piane formate da contorni, comunque complicati, che non si intersechino tra di loro. Il piano metrico possiede un contorno chiuso in quanto esso è delimitato dalla sua retta all'infinito ed è topologicamente equivalente ad un disco piatto.

Raster I dati sono organizzati in una struttura regolare, che ricalca il dispositivo di visualizzazione. Generalmente l'elemento minimo è una cella quadrata (detta pixel) o cubica (voxel), a seconda che il dominio sia il piano immagine o lo spazio tridimensionale. Un'entità geometrica viene rappresentata come collezione di tutte le celle che attraversa.

Ricostruzione Procedura mediante la quale l'insieme di misure effettuate durante l'acquisizione vengono rielaborate per generare un modello dell'oggetto acquisito. Se l'input è costituito da un insieme di campioni, la ricostruzione effettua una generalizzazione dell'informazione posseduta sull'oggetto.

Trasformazione affine Una trasformazione affine è una applicazione definita sullo spazio euclideo che trasforma rette in rette. Hanno la proprietà di conservare il parallelismo, delle linee, ma non le lunghezze ne' gli angoli. Le più comuni sono la traslazione, la rotazione e la variazione di scala. Esse sono rappresentabili mediante una matrice.

Triangolazione di Delaunay La triangolazione di Delaunay di un insieme di punti P , è costituita da triangoli (o tetraedri in tre dimensioni) aventi per vertici i punti di P stessi, tali che nessun vertice cada all'interno della sfera che circoscrive un qualsiasi triangolo. Ciò porta alla realizzazione di una griglia in cui i triangoli (tetraedri) sono "ben dimensionati", cioè il più possibile regolari. L'insieme dei lati più esterni della triangolazione (quelli cioè che appartengono ad un solo triangolo) forma il *convex hull* dell'insieme P . A meno di punti di P coincidenti, la triangolazione di Delaunay è unica.

Vettoriale Forme geometriche (punti, segmenti, curve, solidi) vengono descritti in modo simbolico, insieme ai loro parametri caratteristici.

Glossario

Bibliografia

- [3DS] <http://www.3Dsystems.com/>.
- [ADRY94] Nur Arad, Nira Dyn, Daniel Reifeld, and Yehezkel Yeshurun. Image warping by radial basis functions: Application to facial expressions. *Computer Vision, Graphics, and Image Processing. Graphical Models and Image Processing*, 56(2):161–172, March 1994.
- [AMS99] C. G. Atkeson, A. W. Moore, and S. Schaal. Locally weighted learning. *Artificial Intelligence Review*, 1999.
- [AR95] Nur Arad and Daniel Reifeld. Image warping using few anchor points and radial functions. *Computer Graphics Forum*, 14(1):35–46, March 1995.
- [AS96] Maria-Elena Algorri and Francis Schmitt. Surface reconstruction from unstructured 3D data. *Computer Graphics Forum*, 15(1):47–60, March 1996.
- [BA83] P.J. Burt and E.H. Adelson. Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, 1983.
- [BA95] N. A. Borghese and M. A. Arbib. Generation of temporal sequences using local dynamic programming. *Neural Networks*, 8(1):39–54, 1995.
- [BBCS99] Fausto Bernardini, Chandrajit L. Bajaj, Jindong Chen, and Daniel R. Schikore. Automatic reconstruction of 3D CAD models from digital scans. *Int. J. on Comp. Geom. and Appl.*, 9(4–5):327–370, Oct 1999.
- [BBX95] Chandrajit L. Bajaj, Fausto Bernardini, and Guoliang Xu. Automatic reconstruction of surfaces and scalar fields from

- 3D scans. In Robert Cook, editor, *SIGGRAPH 95 Conference Proceedings*, Annual Conference Series, pages 109–118. ACM SIGGRAPH, Addison Wesley, August 1995. held in Los Angeles, California, 06-11 August 1995.
- [BC00] Nunzio Alberto Borghese and Pietro Cerveri. Calibrating a video camera pair with a rigid bar. *Pattern Recognition*, 33(1):81–95, 2000.
- [Ber93] A.B. Berg. Locating global minima in optimisation problems by a random-cost approach. *Nature*, 361:708–710, 1993.
- [BF90] Nunzio Alberto Borghese and Giancarlo Ferrigno. An algorithm for 3-D automatic movement analysis by means of standard tvcameras. *IEEE Transactions on Biomedical Engineering*, 37:1221–1225, 1990.
- [BF98] Nunzio Alberto Borghese and Stefano Ferrari. Hierarchical RBF networks in function approximation. *Neurocomputing*, 19:259–283, 1998.
- [BF99] Nunzio Alberto Borghese and Stefano Ferrari. Automatic acquisition of 3-D objects through portable modular system. In V. Piuri and M. Savino, editors, *Proceedings of the 16th IEEE Instrumentation and Measurement and Technology Conference 1999, IMTC/99*, volume 3, pages 1823–1827, 24–26 may 1999.
- [BF00] Nunzio Alberto Borghese and Stefano Ferrari. A portable modular system for automatic acquisition of 3D objects. *IEEE Transaction on Instrumentation and Measurement*, 49(5):1128–1136, October 2000.
- [BFB⁺98] Nunzio Alberto Borghese, Giancarlo Ferrigno, Guido Baroni, Riccardo Savarè, Stefano Ferrari, and Antonio Pedotti. Autoscan: A flexible and portable scanner of 3D surfaces. *IEEE Computer Graphics & Applications*, 18(3):38–41, may–jun 1998.
- [BH94] A. Baader and G. Hirzinger. A self-organizing algorithm for multisensory surface reconstruction. In *IROS '94. Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems. Advanced Robotic Systems and the Real World*, volume 1, pages 81–8, New York, NY,

- USA, 1994. Inst. for Robotics & Syst. Dynamics, German Aersp. Res. Establ. , Wessling, Germany, IEEE.
- [BL88] D. S. Broomhead and D. Lowe. Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2:321–355, 1988.
- [BM92] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and machine Intelligence*, 14(2):239–258, February 1992.
- [BMF99] Nunzio Alberto Borghese, Mauro Maggioni, and Stefano Ferrari. Multi-scale approximation with Hierarchical Radial Basis Functions networks. Technical Report 99-67, Politecnico di Milano - Dipartimento di Elettronica e Informazione, Piazza Leonardo da Vinci, 32 20133 Milano, Italia, 1999.
- [BMR⁺99] Fausto Bernardini, Joshua Mittleman, Holly Rushmeier, Claudio Silva, and Gabriel Taubin. The ball-pivoting algorithm for surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 5(4):349–359, Oct–Dec 1999.
- [BN76] James F. Blinn and Martin E. Newell. Texture and reflection in computer generated images. *Communications of the ACM*, 19(10):542–547, October 1976.
- [Boi84] Jean-Daniel Boissonnat. Geometric structures for three-dimensional shape representation. *ACM Transactions on Graphics*, 3(4):266–286, October 1984.
- [Bon98] Georges-Pierre Bonneau. Multiresolution analysis on irregular surface meshes. *IEEE Transactions on Visualization and Computer Graphics*, 4(4), October 1998.
- [BRFP90] N.A. Borghese, M. Di Rienzo, G. Ferrigno, and A. Pedotti. Elite: a goal-oriented vision system for moving objects detection. *Robotica*, 9:275–282, 1990.
- [BTG95] Eric Bittar, Nicolas Tsingos, and Marie-Paule Gascuel. Automatic reconstruction of unstructured 3D data: Combining medial axis and implicit surfaces. *Computer Graphics Forum*, 14(3):C/457–C/468, September 1995.

Bibliografia

- [BV99] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3D faces. *Proceedings of SIGGRAPH 99*, pages 187–194, August 1999. ISBN 0-20148-560-5. Held in Los Angeles, California.
- [BZ95] S. A. Billings and G. L. Zheng. Radial basis function network configuration using genetic algorithms. *Neural Networks*, 8(6):877–890, 1995.
- [Cat74] E. Catmull. A subdivision algorithm for computer display of curved surfaces. Technical Report UTEC-CSc-74-133, University of Utah, Salt Lake City, December 1974.
- [CBB93] B.C. Cetin, J. Barhen, and J.W. Burdick. Terminal repeller unconstrained subenergy tunnelling (trust) for fast global optimization. *Journal of Optimization Theory and Application*, 77(1):97–126, 1993.
- [CCMS97] Andrea Ciampalini, Paolo Cignoni, Claudio Montani, and Roberto Scopigno. Multiresolution decimation based on global error. *The Visual Computer*, 13(5):228–246, 1997. ISSN 0178-2789.
- [CDF92] A. Cohen, Ingrid Daubechies, and J.C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45:485–560, 1992.
- [Cha95] Y. T. Chan. *Wavelet Basics*. Kluwer Academic Publishers, Boston, 1995.
- [CL96] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In Holly Rushmeier, editor, *SIGGRAPH 96 Conference Proceedings*, Annual Conference Series, pages 303–312. ACM SIGGRAPH, Addison Wesley, August 1996. held in New Orleans, Louisiana, 04-09 August 1996.
- [CM91] Yang Chen and Gérard Medioni. Object modeling by registration of multiple range images. In *Proceedings of the IEEE International Conference on Robotics and Automation, 1991*, volume 3, pages 2724–2729, April 1991.
- [CM92a] Yang Chen and Gérard Medioni. Object modelling by registration of multiple range images. *International Journal of Image and Vision Computing*, 10(3):145–155, April 1992.

- [CM92b] Steven D. Cochran and Gérard Medioni. 3-D surface description from binocular stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):981–994, Oct 1992.
- [CMS98] P. Cignoni, C. Montani, and R. Scopigno. A comparison of mesh simplification algorithms. *Computers & Graphics*, 22(1):37–54, February 1998. ISSN 0097-8493.
- [Dau92] Ingrid Daubechies. *Ten Lectures on Wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [Dau96] Ingrid Daubechies. Where do wavelets come from? A personal point of view. *Proceedings of the IEEE*, 84(4):510–513, April 1996.
- [DGSSed] I. Daubechies, I. Guskov, P. Schröder, and W. Sweldens. Wavelets on irregular point sets. *Phil. Trans. R. Soc. Lond. A*, To be published.
- [DIGa] <http://www.cyberware.com/products/digisize.html>.
- [DIGb] <http://www.digibotics.com/3Product.htm>.
- [DM98] Douglas DeCarlo and Dimitris Metaxas. Shape evolution with structural and topological changes using blending. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 20(11):1186–1205, Nov 1998.
- [Dob97] Adrian Dobson. Architectural composition in the electronic design studio: conceptual design using CAD visualisation and virtual reality modelling. In *Proceedings of IEEE Conference on Information Visualization, 1997*, pages 179–185, 27–29 aug 1997.
- [DTM96] Paul E. Debevec, Camillo J. Taylor, and Jitendra Malik. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. *Computer Graphics*, 30(Annual Conference Series):11–20, 1996.
- [DWJM98] Chitra Dorai, Gang Wang, Anil K. Jain, and Carolyn Mercer. Registration and integration of multiple object views for 3D model construction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):83–89, Jan 1998.

Bibliografia

- [EDD⁺95] Matthias Eck, Tony DeRose, Tom Duchamp, Hugues Hoppe, Michael Lounsbery, and Werner Stuetzle. Multiresolution analysis of arbitrary meshes. In Robert Cook, editor, *SIGGRAPH 95 Conference Proceedings*, Annual Conference Series, pages 173–182. ACM SIGGRAPH, Addison Wesley, August 1995. held in Los Angeles, California, 06-11 August 1995.
- [EH96] Matthias Eck and Hugues Hoppe. Automatic reconstruction of B-Spline surfaces of arbitrary topological type. In Holly Rushmeier, editor, *SIGGRAPH 96 Conference Proceedings*, Annual Conference Series, pages 325–334. ACM SIGGRAPH, Addison Wesley, August 1996. held in New Orleans, Louisiana, 04-09 August 1996.
- [EM94] H. Edelsbrunner and E. P. Mücke. Three-dimensional alpha shapes. *ACM Trans. Graphics*, 13(1):43–72, January 1994.
- [EMP⁺94] David Ebert, Kent Musgrave, Darwyn Peachey, Ken Perlin, and Worley. *Texturing and Modeling: A Procedural Approach*. Academic Press, October 1994. ISBN 0-12-228760-6.
- [FAR] <http://www.faro.com/>.
- [FB88] David R. Forsey and Richard H. Bartels. Hierarchical B-spline refinement. *Computer Graphics*, 22(4):205–212, August 1988.
- [FB95] David Forsey and Richard H. Bartels. Surface fitting with hierarchical splines. *ACM Transactions on Graphics*, 14(2):134–161, April 1995.
- [FBP00] Stefano Ferrari, Nunzio Alberto Borghese, and Vincenzo Piuri. Multi-resolution models for data processing: an experimental sensitivity analysis. In *Proceedings of the 17th IEEE Instrumentation and Measurement Technology Conference (IMTC 2000)*, volume 2, pages 1056–1060, 1–4 May 2000.
- [FE98] C. Ford and D.M. Etter. Wavelet basis reconstruction of nonuniformly sampled data. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 45(8):1165–1168, Aug 1998.

- [Fer99] Stefano Ferrari. MRA e HRBF: Comparazione tra modelli multirisoluzione. Technical Report 99-54, Politecnico di Milano — Dipartimento di Elettronica e Informazione, Piazza Leonardo da Vinci 32, 20133 Milano Italia, 1999.
- [FP90] Giancarlo Ferrigno and Antonio Pedotti. Modularly expandible system for real-time processing of a TV display, useful in particular for the acquisition of coordinates of known shapes objects. U.S. patent 4,706,296, 1990.
- [Fri94] Bernd Fritzke. Growing cell structures — A self-organizing network for unsupervised and supervised learning. *Neural Networks*, 7(9):1441–1460, 1994.
- [Fri95] Bernd Fritzke. Growing grid - a self-organizing network with constant neighborhood range and adaptation strength. *Neural Processing Letters*, 2(5):9–13, 1995.
- [FvFH90] J. Foley, A. van Dam, S. Feiner, and J. Hughes. *Computer graphics — Principles and Practice—*. Addison-Wesley, 2nd edition, 1990.
- [FW98] David Forsey and David Wong. Multiresolution surface reconstruction for hierarchical B-splines. In *Graphics Interface*, pages 57–64, June 1998.
- [FW00] Peter Fröst and Peter Warren. Virtual reality used in a collaborative architectural design process. In E. Banissi, M. Bannatyne, C. Chen, F. Khosrowshahi, M. Sarfraz, and A. Ursyn, editors, *Proceedings of IEEE International Conference on Information Visualization, 2000*, pages 568–573, 19–21 jul 2000.
- [GH97] Michael Garland and Paul S. Heckbert. Surface simplification using quadric error metrics. In Turner Whitted, editor, *SIGGRAPH 97 Conference Proceedings*, Annual Conference Series, pages 209–216. ACM SIGGRAPH, Addison Wesley, August 1997. ISBN 0-89791-896-7.
- [GH98] Michael Garland and Paul S. Heckbert. Simplifying surfaces with color and texture using quadric error metrics. In David Ebert, Hans Hagen, and Holly Rushmeier, editors, *IEEE Visualization '98*, pages 263–270. IEEE, 1998.

Bibliografia

- [GJP95] Federico Girosi, Michael Jones, and Tomaso Poggio. Regularization theory and neural networks architectures. *Neural Computation*, 7(2):219–269, 1995.
- [GMW97] B. Guo, J. Menon, and B. Willette. Surface reconstruction using alpha shapes. *Computer Graphics Forum*, 16(4):177–190, 1997. ISSN 1067-7055.
- [GR99] E. Gagnon and J.F. Rivest. Laser range imaging using the self-mixing effect in a laser diode. *IEEE Trans. Instrum. Meas.*, 48(3):693–699, 1999.
- [GSS99] Igor Guskov, Wim Sweldens, and Peter Schröder. Multiresolution signal processing for meshes. *Computer Graphics Proceedings (SIGGRAPH 99)*, pages 325–334, 1999.
- [GST94] D. Gorse, A. Sheperd, and J.G. Taylor. Avoiding local minima by a classical range expansion algorithm. In *Proceedings of ICANN 94*, 1994.
- [GVSS00] Igor Guskov, Kiril Vidimce, Wim Sweldens, and Peter Schröder. Normal meshes. *Proceedings of SIGGRAPH 2000*, pages 95–102, July 2000. ISBN 1-58113-208-5.
- [GW92] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Addison-Wesley, Reading, MA, USA, 3 edition, 1992.
- [HDD⁺92] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. Surface reconstruction from unorganized points. In Edwin E. Catmull, editor, *Computer Graphics (SIGGRAPH '92 Proceedings)*, volume 26, pages 71–78, July 1992.
- [HDD⁺94] Hugues Hoppe, Tony DeRose, Tom Duchamp, Mark Halstead, Hubert Jin, John McDonald, Jean Schweitzer, and Werner Stuetzle. Piecewise smooth surface reconstruction. *Computer Graphics*, 28(Annual Conference Series):295–302, July 1994.
- [Hec86] Paul S. Heckbert. Survey of texture mapping. In M. Green, editor, *Proceedings of Graphics Interface '86*, pages 207–212, May 1986.

- [Hel95] Peter Niels Heller. Rank M wavelets with N vanishing moments. *SIAM Journal on Matrix Analysis and Applications*, 16(2):502–519, April 1995.
- [HG99] Paul S. Heckbert and Michael Garland. Optimal triangulation and quadric-based surface simplification. *Computational Geometry*, 14(49–65), 1999.
- [HBJB⁺96] Adam Hoover, Gillian Jean-Baptiste, Xiaoyi Jiang, Patrick J. Flynn, Horst Bunke, Dmitry B. Goldgof, Kevin Bowyer, David W. Eggert, Andrew Fitzgibbon, and Robert B. Fisher. An experimental comparison of range image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(7):673–689, July 1996.
- [HKP91] John Hertz, Anders Krogh, and Richard G. Palmer. *An Introduction to the Theory of Neural Computation*. Lecture Notes Volume I. Addison Wesley, 1991.
- [Hop94] Hugues Hoppe. Surface reconstruction from unorganized points. Technical Report TR-94-06-01, University of Washington, Department of Computer Science and Engineering, June 1994.
- [Hop96] Hugues Hoppe. Progressive meshes. In Holly Rushmeier, editor, *SIGGRAPH 96 Conference Proceedings*, Annual Conference Series, pages 99–108. ACM SIGGRAPH, Addison Wesley, August 1996. held in New Orleans, Louisiana, 04-09 August 1996.
- [Hop97] Hugues Hoppe. View-dependent refinement of progressive meshes. In Turner Whitted, editor, *SIGGRAPH 97 Conference Proceedings*, Annual Conference Series, pages 189–198. ACM SIGGRAPH, Addison Wesley, August 1997. ISBN 0-89791-896-7.
- [Hop99] Hugues H. Hoppe. New quadric metric for simplifying meshes with appearance attributes. In David Ebert, Markus Gross, and Bernd Hamann, editors, *IEEE Visualization '99*, pages 59–66, San Francisco, 1999. IEEE.
- [IBS98] F. Issselhard, G. Brunnett, and T. Schreiber. Polyhedral approximation and first order segmentation of unstructured

- point sets. In Franz-Erich Wolter and Nicholas M. Patrikalakis, editors, *Proceedings of the Conference on Computer Graphics International 1998 (CGI-98)*, pages 433–443, Los Alamitos, California, June 22–26 1998. IEEE Computer Society.
- [jM96] Francesco Introna jr. and Giuseppe Mastronardi. On human face identification methods. In M. de Sario, B. Macione, P. Pugliese, and M. Savino, editors, *Proceedings of 8th Mediterranean Electrotechnical Conference, 1996 (MELECON '96)*, volume 2, pages 1101–1103, 13–16 may 1996.
- [JS94] Björn Jawerth and Wim Sweldens. An overview of wavelet based multiresolution analyses. *SIAM Rev.*, 36(3):377–412, 1994.
- [KGV83] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimisation by simulated annealing. *Science*, 220:671–680, 1983.
- [Kir83] D. Kirkpatrick. Optimal search in planar subdivisions. *SIAM J. Computing*, 12(1):28–35, February 1983.
- [Koh95] Teuvo Kohonen. *Self-Organizing Maps*, volume 30 of *Springer Series in Information Sciences*. Springer, Berlin, Heidelberg, 1995. (Second Extended Edition 1997).
- [LC87] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, 21(4):163–169, July 1987.
- [LDSS99] Aaron W. F. Lee, David Dobkin, Wim Sweldens, and Peter Schröder. Multiresolution mesh morphing. *Computer Graphics*, 33(Annual Conference Series):343–350, 1999.
- [LDW97] Michael Lounsbery, Tony D. DeRose, and Joe Warren. Multiresolution analysis for surfaces of arbitrary topological type. *ACM Transactions on Graphics*, 16(1):34–73, January 1997. ISSN 0730-0301.
- [Lev99] Adi Levin. Interpolating nets of curves by smooth subdivision surfaces. *Proceedings of SIGGRAPH 99*, pages 57–64, August 1999. ISBN 0-20148-560-5. Held in Los Angeles, California.

- [LJ94] Henry J. Lamousin and Warren N. Waggenspack Jr. NURBS-based free-form deformations. *IEEE Computer Graphics and Applications*, 14(6):59–65, November 1994.
- [Llo82] S. P. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28:128–137, 1982.
- [LM94] Chia-Wei Liao and Gérard Medioni. Simultaneous segmentation and approximation of complex patterns. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 617–623, Los Alamitos, CA, USA, June 1994. IEEE Computer Society Press.
- [LM98] Bruno Lévy and Jean-Laurent Mallet. Non-distorted texture mapping for sheared triangulated meshes. In Michael Cohen, editor, *SIGGRAPH 98 Conference Proceedings*, Annual Conference Series, pages 343–352. ACM SIGGRAPH, Addison Wesley, July 1998. ISBN 0-89791-999-8.
- [LMH00] Aaron Lee, Henry Moreton, and Hugues Hoppe. Displaced subdivision surfaces. *Proceedings of SIGGRAPH 2000*, pages 85–94, July 2000. ISBN 1-58113-208-5.
- [LPC⁺00] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, Jonathan Shade, and Duane Fulk. The Digital Michelangelo project: 3D scanning of large statues. *Proceedings of SIGGRAPH 2000*, pages 131–144, July 2000. ISBN 1-58113-208-5.
- [LSS⁺98] Aaron W. F. Lee, Wim Sweldens, Peter Schröder, Lawrence Cowsar, and David Dobkin. MAPS: Multiresolution adaptive parameterization of surfaces. In Michael Cohen, editor, *SIGGRAPH 98 Conference Proceedings*, Annual Conference Series, pages 95–104. ACM SIGGRAPH, Addison Wesley, July 1998. ISBN 0-89791-999-8.
- [LT97] Kok-Lim Low and Tiow-Seng Tan. Model simplification using vertex-clustering. In Michael Cohen and David Zeltzer, editors, *1997 Symposium on Interactive 3D Graphics*, pages 75–82, New York, April 27–30 1997. ACM SIGGRAPH, ACM Press. ISBN 0-89791-884-3.

Bibliografia

- [LWS97] Seungyong Lee, George Wolberg, and Sung Yong Shin. Scattered Data Interpolation with Multilevel B-Splines. *IEEE Transactions on Visualization and Computer Graphics*, 3(3):228–244, July 1997.
- [Mag99] Mauro Maggioni. Ondine biortogonali a m bande a supporto compatto e ondine Burt-Adelson. Master’s thesis, Università degli Studi di Milano, Milano, July 1999.
- [Mag00] Mauro Maggioni. m -band Burt-Adelson biorthogonal wavelets. *Applied and Computational Harmonic Analysis*, 9(3):286–311, October 2000.
- [Mal89] Stéphane Mallat. A theory for multiscale signal decomposition: The wavelet representation. *IEEE Trans. on Pattern and Machine Intelligence*, 11(7):674–693, 1989.
- [Mal98] Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1998.
- [MBL⁺91] James V. Miller, David E. Breen, William E. Lorensen, Robert M. O’Bara, and Michael J. Wozny. Geometrically deformed models: a method for extracting closed geometric models form volume data. *Computer Graphics*, 25(4):217–226, July 1991.
- [MBS93] T.M. Martinetz, S.G. Berkovich, and K.J. Schulten. ”neural-gas” network for vector quantization and its application to time-series prediction. *IEEE Trans. on Neural Networks*, 4(4):558–568, 1993.
- [MD89] J. Moody and C. J. Darken. Fast learning in networks of locally-tuned processing units. *Neural Computation*, 1(2):281–294, 1989.
- [MM98a] R. Mencl and H. Müller. Graph-based surface reconstruction using structures in scattered point sets. In Franz-Erich Wolter and Nicholas M. Patrikalakis, editors, *Proceedings of the Conference on Computer Graphics International 1998 (CGI-98)*, pages 298–311, Los Alamitos, California, June 22–26 1998. IEEE Computer Society.
- [MM98b] Robert Mencl and Heinrich Müller. Interpolation and approximation of surfaces from three-dimensional scattered data points. In *STAR — State of The Art Report — Eurographics ’98*, pages 51–67, 1998.

- [MPE] <http://www.cselt.it/mpeg/standards/mpeg-4/mpeg-4.htm>.
- [NBC⁺99] S.G. Nikolov, D.R. Bull, C.N. Canagarajah, M. Halliwell, and P.N.T. Wells. Image fusion using a 3-D wavelet transform. In *Seventh International Conference on Image Processing And Its Applications*, volume 1, pages 235–239, July 1999.
- [NP91] K. S. Narendra and K. Parthasarathy. Gradient methods for the optimization of dynamical systems containing neural networks. *IEEE Transactions on Neural Networks*, 2(2):252–262, March 1991.
- [NT89] K. S. Narendra and M. A. L. Thathachar. *Learning Automata - An Introduction*. Prentice Hall, 1989.
- [Orr95] Mark J. L. Orr. Regularization in the selection of radial basis function centers. *Neural Computation*, 7(3):606–623, 1995.
- [PARD⁺98] Kari Pulli, Habib Abi-Rached, Tom Duchamp, Linda G. Shapiro, and Werner Stuetzle. Acquisition and visualization of colored 3D objects. In A.K. Jain, S. Venkatesh, and B.C. Lovell, editors, *Proceedings of the Fourteenth International Conference on Pattern Recognition*, volume 1, pages 11–15, Aug 1998.
- [PG90] Tomaso Poggio and Federico Girosi. Network for approximation and learning. *Proceedings of the IEEE*, 78(9):1481–1497, September 1990.
- [PHF99] Emil Praun, Hugues Hoppe, and Adam Finkelstein. Robust mesh watermarking. In Alyn Rockwood, editor, *Siggraph 1999, Computer Graphics Proceedings, Annual Conference Series*, pages 49–56, Los Angeles, 1999. ACM Siggraph, Addison Wesley Longman.
- [Pit99] Richard Pito. A solution to the next best view problem for automated surface acquisition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(10):1016–1030, Oct 1999.
- [Pla91] John Platt. A resource-allocating network for function interpolation. *Neural Computation*, 3:213–225, 1991.

Bibliografia

- [PM90] Pietro Perona and Jitendra Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-12(7):629–639, July 1990.
- [PMWP00] P. Jonathon Phillips, Alvin Martin, C.L. Wilson, and Mark Przybocki. An introduction evaluating biometric systems. *Computer*, 33(2):56–63, February 2000.
- [Pog90] Tomaso Poggio. A theory of how the brain might work. Technical Report AIM-1253, Massachusetts Institute of Technology, December 1990.
- [PR99] Luis Pastor and Angel Rodriguez. Surface approximation of 3D objects from irregularly sampled clouds of 3D points using spherical wavelets. In *International Conference on Image Analysis and Processing, 1999*, pages 70–75, Sep 1999.
- [PST00] Federico Pedersini, Augusto Sarti, and Stefano Tubaro. Visible surface reconstruction with accurate localization of object boundaries. *IEEE Transaction on Circuits and Systems for Video Technology*, 10(2):278–292, March 2000.
- [PSX] <http://psx.ign.com/news/7154.html>.
- [PT97] Les Piegl and Wayne Tiller. *The NURBS Book*. Springer-Verlag, New York, NY, 2 edition, 1997. 646 pages, 578 illustrations.
- [PTR⁺98] Michael Petrov, Andrey Talapov, Timothy Robertson, Alexei Lebedev, Alexander Zhilyaev, and Leonid Polonsky. Optical 3D digitizers: Bringing life to the virtual world. *IEEE Computer Graphics & Applications*, 18(3):28–37, May – June 1998. ISSN 0272-1716.
- [Pul99] Kari Pulli. Multiview registration for large data sets. In *Second International Conference on 3-D Digital Imaging and Modeling*, pages 160–168, Oct 1999.
- [PYL00] In Kyu Park, Il Dong Yun, and Sang Uk Lee. Automatic 3-D model synthesis from measured range data. *IEEE Transaction on Circuits and Systems for Video Technology*, 10(2):293–301, Mar 2000.

- [RB93] Jarek Rossignac and Paul Borrel. Multi-resolution 3D approximations for rendering complex scenes. In B. Falci-dieno and T. Kunii, editors, *Modeling in Computer Graphics: Methods and Applications*, pages 455–465, Berlin, 1993. Springer-Verlag. Proc. of Conf., Genoa, Italy, June 1993. (Also available as IBM Research Report RC 17697, Feb. 1992, Yorktown Heights, NY 10598).
- [RB99] Holly Rushmeier and Fausto Bernardini. Computing consistent normals and colors from photometric data. In *Proceedings of the Second International Conference on 3-D Digital Imaging and Modeling*, pages 99–108, 4–8 oct 1999.
- [RCMS99] Claudio Rocchini, Paolo Cignoni, Claudio Montani, and Roberto Scopigno. Multiple textures stitching and blending on 3D objects. In Dani Lischinski and Greg Ward Larson, editors, *Rendering Techniques '99*, Eurographics, pages 119–130. Springer-Verlag Wien New York, 1999.
- [Rig00] Paolo Rigioli. Sistema per acquisizione di modelli 3D di volti umani e loro animazione ad alto grado di realismo mediante motion-capture. Master's thesis, Politecnico di Milano, Dept. Electronics and Information, 2000.
- [RL00] Szymon Rusinkiewicz and Marc Levoy. Qsplat: A multi-resolution point rendering system for large meshes. *Proceedings of SIGGRAPH 2000*, pages 343–352, July 2000. ISBN 1-58113-208-5.
- [RW97] Gerhard Roth and Eko Wibowoo. An efficient volumetric method for building closed triangular meshes from 3-D image and point data. In Wayne Davis, Marilyn Mantel, and Victor Klassen, editors, *Graphics Interface*, pages 173–180, May 1997.
- [Sam84] Hanan Samet. The quadtree and related data structures. *ACM Computing Surveys*, June 1984, 16(2):188–260, 1984.
- [SB97] Thomas Schreiber and Guido Brunnet. Approximating 3D objects from measured points. In *Proceedings of 30th ISADATA*, 1997.
- [Şen97] Burku Şenyapılı. Visualization of virtual architecture. In *Proceedings of IEEE Conference on Information Visualization, 1997*, pages 260–266, 27–29 aug 1997.

Bibliografia

- [Sha] James Sharman. <http://www.exaflop.org/docs/marchcubes/ind.html>.
- [SHGB93] P. Steffen, P. Heller, R. A. Gopinath, and C. S. Burrus. Theory of regular M -band wavelet bases. *IEEE Trans. Signal Processing*, 41(12), December 1993. Special Transaction issue on wavelets; Rice contribution also in Tech. Report No. CML TR-91-22, November 1991.
- [SI96] Yoichi Sato and Katsushi Ikeuchi. Reflectance analysis for 3D computer graphics model generation. *Graphical models and image processing: GMIP*, 58(5):437–451, September 1996.
- [SN96] Gilbert Strang and Truong Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, MA, 1996.
- [Soa00] Paolo M. Soardi. Biorthogonal m -channel compactly supported wavelets. *Constructive Approximation*, 16(2):283–311, 2000.
- [Spe90] D. F. Specht. Probabilistic neural networks. *Neural Networks*, 3:109–118, 1990.
- [SS92] Robert M. Sanner and Jean-Jacques E. Slotine. Gaussian networks for direct adaptive control. *IEEE Transactions on Neural Networks*, 3(6):837–863, November 1992.
- [SS95] Peter Schröder and Wim Sweldens. Spherical wavelets: Efficiently representing functions on the sphere. In Robert Cook, editor, *SIGGRAPH 95 Conference Proceedings*, Annual Conference Series, pages 161–172. ACM SIGGRAPH, Addison Wesley, August 1995. held in Los Angeles, California, 06-11 August 1995.
- [SS96] W. Sweldens and P. Schröder. Building your own wavelets at home. In *Wavelets in Computer Graphics*, pages 15–87. ACM SIGGRAPH Course notes, 1996.
- [SSS00] Alexei Sourin, Olga Sourina, and Howe Tet Sen. Virtual orthopedic surgery training. *IEEE Computer Graphics and Applications*, 20(3):6–9, May – June 2000.
- [ST92] Richard Szeliski and David Tonnesen. Surface modeling with oriented particle systems. *Computer Graphics*, 26(2):185–194, July 1992.

- [Swe95] W. Sweldens. The lifting scheme: A new philosophy in biorthogonal wavelet constructions. In Andrew F. Laine, Michael A. Unser, and Mladen V. Wickerhauser, editors, *Wavelet applications in signal and image processing III*, volume 2569 of *Proceedings of SPIE*, pages 68–79, 1995.
- [Swe96] W. Sweldens. The lifting scheme: A custom-design construction of biorthogonal wavelets. *Appl. Comput. Harmon. Anal.*, 3(2):186–200, 1996.
- [Swe97] Wim Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29(2):511–546, 1997.
- [Ter88] Demetri Terzopoulos. The computation of visible-surface representations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10(4):417–438, July 1988.
- [TH] Michael Teschner and Christian Henn. <http://www.sgi.com/chembio/resources/texture/>.
- [TL94] Greg Turk and Marc Levoy. Zippered polygon meshes from range images. *Computer Graphics*, 28(Annual Conference Series):311–318, July 1994.
- [TM91] Demetri Terzopoulos and Dimitri Metaxas. Dynamic 3D models with local and global deformations: deformable superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):703–714, Jul 1991.
- [TOdSg⁺99] William B. Thompson, Jonathan C. Owen, H. James de St. germain, Stevan R. Stark, and Thomas C. Henderson. Feature-based reverse engineering of mechanical parts. *IEEE Transactions on Robotics and Automation*, 15(1):57–66, February 1999.
- [Vap99] V. N. Vapnik. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10(5):988, September 1999.
- [Vel95] Remco C. Veltkamp. Boundaries through scattered points of unknown density. *Graphical models and image processing: GMIP*, 57(6):441–452, November 1995.

Bibliografia

- [VZ98] M. Vasiliu and B. Zavidovique. High precision image matching using 4D optimal minimisation. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 5, pages 2793–2796, May 1998.
- [War00] Randall Warniers. Picture perfect modeling. *Computer Graphics World*, pages 54–58, June 2000.
- [WD92] C. J. C. H. Watkins and P. Dayan. Q-learning. *Machine Learning Journal*, 8(3/4), May 1992. Special Issue on Reinforcement Learning.
- [Wel96] Frank Weller. On the stability of edges in planar delaunay triangulations. Research Report 637/1996, Fachbereich Informatik, Universität Dortmund, Fachbereich Informatik, Universität Dortmund, 44221 Dortmund, Germany, December 1996.
- [YA98] Ruigang Yang and Peter K. Allen. Registering, integrating and building CAD models from range data. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-98)*, pages 3115–3120, Piscataway, May 16–20 1998. IEEE Computer Society.
- [YKOL96] Bum-Jae You, Hong Jae Kim, Sang-Rok Oh, and Chong Won Lee. Development of an eye system for three dimensional shape reconstruction. In *5th IEEE International Workshop on Robot and Human Communication*, pages 376–381, Nov 1996.
- [Yui88] A. L. Yuille. A computational theory for the perception of coherent visual motion. *Nature*, 333:71–74, 1988.
- [ZHH98] Hongbin Zha, Tsuyoshi Hoshide, and Tsutomu Hasegawa. A recursive fitting-and-splitting algorithm for 3-D object modeling using superquadrics. In A.K. Jain, S. Venkatesh, and B.C. Lovell, editors, *Proceedings of the International Conference on Pattern Recognition*, volume 1, pages 658–662, August 1998.
- [ZSS96] Denis Zorin, Peter Schroeder, and Wim Sweldens. Interpolating subdivision for meshes with arbitrary topology. In Holly Rushmeier, editor, *SIGGRAPH 96 Conference Proceedings*, Annual Conference Series, pages 189–192. ACM

SIGGRAPH, Addison Wesley, August 1996. held in New Orleans, Louisiana, 04-09 August 1996.

- [ZSS97] Denis Zorin, Peter Schröder, and Wim Sweldens. Interactive multiresolution mesh editing. In Turner Whitted, editor, *SIGGRAPH 97 Conference Proceedings*, Annual Conference Series, pages 259–268. ACM SIGGRAPH, Addison Wesley, August 1997. ISBN 0-89791-896-7.