



UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Specialistica in Scienze e Tecnologie dell'Informazione

**ELABORAZIONE TRIDIMENSIONALE
DI IMPRONTE DIGITALI
ACQUISITE SENZA CONTATTO**

RELATORE
Prof. Fabio Scotti

CORRELATORE
Dott. Ruggero Donida Labati

TESI DI LAUREA DI
Genovese Angelo
Matr. 737558

ANNO ACCADEMICO 2009/2010



Nella pagina precedente:
3D render of a fingerprint detected using the scanning Kelvin probe (SKP)
Picture by James Sullivan

"Imagination is more important than knowledge"

Albert Einstein

"I never think of the future. It comes soon enough"

Albert Einstein

Indice

Indice	V
1 Introduzione	1
1.1 La biometria delle impronte digitali	1
1.2 Problematiche attuali e nuovi sviluppi	2
1.3 Il lavoro di tesi	3
2 Tecniche di ricostruzione tridimensionale	5
2.1 Modello della fotocamera	6
2.1.1 Calibrazione di una fotocamera	10
2.2 Sistemi basati su luce strutturata	13
2.3 Sistemi basati su lama laser	15
2.4 Sistemi basati su viste multiple	18
2.4.1 Shape from silhouette	18
2.4.2 Shape from photo-consistency	20
2.4.3 Structure from motion	21
2.4.4 Sistemi basati su immagini stereoscopiche	23
2.5 Rappresentazione della scena nello spazio tridimensionale . .	37
2.6 Ricostruzione tridimensionale con OpenCV	40
2.6.1 Numero di immagini di calibrazione	42
2.6.2 Estrazione dei corner	43
2.6.3 Calcolo dell'omografia	45

2.6.4	Calibrazione delle fotocamere	46
2.6.5	Calcolo dei parametri di distorsione	48
2.6.6	Calibrazione stereoscopica	49
2.6.7	Rettificazione	50
2.6.8	Ricerca delle corrispondenze e costruzione della disparity map	51
2.6.9	Costruzione del modello tridimensionale	53
3	Impronte digitali acquisite senza contatto	55
3.1	Impronte digitali a contatto e senza contatto	55
3.2	Sistemi basati su immagini bidimensionali	57
3.2.1	Sistemi di acquisizione	59
3.2.2	Preprocessing di impronte digitali senza contatto . . .	62
3.2.3	Prestazioni e vulnerabilità dei sistemi senza contatto .	67
3.3	Sistemi basati su modelli tridimensionali	68
3.3.1	Unwrapping del modello tridimensionale	70
3.3.2	Sistemi basati su viste multiple	72
3.3.3	Sistema basato su luce strutturata	78
4	Elaborazioni tridimensionali di impronte digitali acquisite senza contatto	81
4.1	Calibrazione del sistema stereoscopico	82
4.1.1	Miglioramento delle immagini di calibrazione	82
4.1.2	Errore di calibrazione	86
4.2	Ricostruzione tridimensionale basata sulle minuzie	91
4.2.1	Segmentazione dell'immagine	92
4.2.2	Calcolo della regione di interesse	93
4.2.3	Preprocessing delle immagini	96
4.2.4	Estrazione ed elaborazione delle minuzie	100
4.2.5	Triangolazione del bordo del dito e delle minuzie . . .	110
4.3	Ricostruzione tridimensionale basata su modello cilindrico . .	113
4.3.1	Calcolo del punto di massima curvatura	114

4.3.2	Calcolo del modello cilindrico	116
4.3.3	Wrapping dei dettagli sul modello cilindrico	118
4.4	Unwrapping del modello tridimensionale cilindrico	125
4.5	Calcolo di un modello tridimensionale sintetico del dito . . .	129
4.5.1	Calcolo del modello sintetico tridimensionale	129
4.5.2	Definizione del setup stereoscopico	132
4.5.3	Calcolo del modello sintetico del pattern di calibrazione	137
4.5.4	Elaborazione delle immagini sintetiche	138
5	Risultati sperimentali	143
5.1	Descrizione delle immagini	144
5.2	Analisi dell'errore di calibrazione	150
5.3	Confronto tra gli algoritmi di preprocessing	155
5.4	Confronto tra le immagini stereoscopiche	158
5.5	Modelli visuali	163
5.6	Accuratezza biometrica ottenuta	167
6	Conclusioni e sviluppi futuri	173
	Bibliografia	177
	Elenco delle figure	182
	Elenco delle tabelle	190
	Indice analitico	191
	Ringraziamenti	195

INTRODUZIONE

1.1 La biometria delle impronte digitali

La biometria è l'insieme di tecnologie che consentono il riconoscimento di un individuo sulla base delle sue caratteristiche fisiche e comportamentali. A differenza dei sistemi tradizionali, il riconoscimento biometrico non si basa su oggetti posseduti (come una tessera magnetica) o conoscenze dimostrate (come una password), ma su tratti fisici e comportamentali propri della persona.

Affinchè una tecnologia biometrica rappresenti una valida alternativa ai sistemi tradizionali, però, deve possedere delle caratteristiche che la rendano utilizzabile da tutte le persone ed essere sufficientemente sicura. A seconda dello scenario in cui viene impiegata, inoltre, bisogna considerare le performance del sistema e il livello di gradimento degli utenti.

Tra le caratteristiche biometriche di un'individuo, le impronte digitali sono state le prime ad essere sfruttate e, al giorno d'oggi, sono ancora le più utilizzate. L'impiego su larga scala di questo tratto biometrico è dovuto principalmente a tre fattori: in primo luogo, l'accuratezza ottenibile dal riconoscimento basato su impronta digitale è sufficientemente elevata affinché questa tecnica possa essere applicata anche in ambiti in cui

è richiesta un'elevata sicurezza; secondariamente, la scansione di un'impronta digitale è relativamente veloce e poco invasiva; infine, i costi necessari al suo impiego sono relativamente contenuti ed inferiori a quelli richiesti da tecniche biometriche più sofisticate, come per esempio quelle basate sull'iride.

1.2 Problematiche attuali e nuovi sviluppi

La grande maggioranza dei dispositivi utilizzati per la scansione delle impronte digitali richiede la pressione del dito sul sensore. Il contatto tra dito e superficie di acquisizione comporta una serie di problematiche, tra cui la deformazione del dito, l'accumulo sul sensore di polvere e sporco presenti sulle dita, la diffusione di malattie cutanee, la permanenza di impronte latenti.

Le ricerche nel campo del riconoscimento biometrico basato su impronte digitali sono sempre più orientate verso lo sviluppo di tecniche che non prevedano il contatto del dito con il sensore. L'uso di tecnologie con questa caratteristica permette di effettuare un'acquisizione più veloce, assente da sporcizia, invariante alle condizioni del dito e non influenzata dalla posizione e della forza esercitata sul sensore.

Esistono inoltre studi che sfruttano l'applicazione di tecniche di ricostruzione tridimensionale al fine di ottenere un modello virtuale del dito. In questo modo è possibile ottenere una rappresentazione globale del tratto biometrico ed utilizzare l'intera area dell'impronta digitale per l'estrazione delle caratteristiche, aumentando quindi l'accuratezza del sistema di riconoscimento. I principali sistemi di ricostruzione tridimensionale disponibili al giorno d'oggi sfruttano tecniche di ricostruzione tridimensionale basate su luce strutturata e fusione di viste multiple.

1.3 Il lavoro di tesi

Lo scopo di questa tesi è stato il calcolo di modelli tridimensionali del dito e dell'impronta digitale tramite tecniche basate su viste multiple di immagini acquisite senza contatto del dito con un sensore. I modelli tridimensionali sono stati elaborati in modo da essere confrontabili con immagini ottenute da sensori tradizionali.

Il lavoro svolto è suddivisibile in passi distinti. Dopo aver effettuato la calibrazione delle fotocamere, le immagini acquisite in laboratorio sono state elaborate al fine di estrarne le caratteristiche salienti, tra cui bordo del dito, regione di interesse, orientamento e frequenza delle creste, posizione di punti notevoli.

Le immagini sono state inoltre trattate con una serie di algoritmi volti a renderle simili a quelle acquisite con sensori tradizionali. Le immagini così ottenute sono state utilizzate per l'estrazione delle caratteristiche dell'impronta digitale, tra cui la posizione delle minuzie ed il pattern delle creste.

Le informazioni ottenute sono state utilizzate per la costruzione di modelli tridimensionali. Gli approcci proposti sono due: il primo di essi si basa sulla triangolazione delle minuzie per ottenere una ricostruzione discreta delle caratteristiche dell'impronta digitale. Il secondo approccio prevede l'uso di punti notevoli estratti dalle immagini (per esempio il punto di massima curvatura delle creste) per l'approssimazione di un volume tridimensionale, che viene completato con la sovrapposizione delle immagini originali.

Infine, i modelli tridimensionali risultanti vengono elaborati per ottenere immagini planari, le cui caratteristiche sono confrontabili con immagini ottenute tramite sensori tradizionali.

Le immagini utilizzate per lo studio e la valutazione degli algoritmi sono state acquisite in laboratorio con condizioni di luce controllate e con una distanza di circa 20 cm tra il dito ed il sensore. Per poter avere a disposizione in tempi rapidi un dataset di immagini stereoscopiche com-

posto da acquisizioni di impronte digitali di numerosi individui diversi, sono inoltre stati studiati metodi per la generazione di immagini multiple sintetiche. Tali immagini sono state realizzate ricreando virtualmente un modello tridimensionale dell'impronta digitale a partire da immagini acquisite tramite contatto.

Le immagini sintetiche sono state poi elaborate in modo da valutare l'accuratezza dei sistemi di ricostruzione tridimensionale. I risultati hanno permesso di confrontare le prestazioni biometriche dei modelli tridimensionali ottenuti con quelle dei sistemi tradizionali.

L'elaborato di tesi è suddiviso come segue: nel Capitolo 2 è descritta la calibrazione delle fotocamere di un sistema a viste multiple e lo stato dell'arte delle tecniche di ricostruzione tridimensionali; nel Capitolo 3 è presentato il confronto tra le tecniche biometriche basate su impronte digitali tradizionali e le tecniche che operano senza il contatto tra dito e sensore. Sono descritte, inoltre, le tecnologie per il calcolo di rappresentazioni tridimensionali del dito. Nel Capitolo 4 sono presentati i metodi proposti per l'elaborazione delle immagini, l'estrazione delle caratteristiche dell'impronta digitale, la costruzione di modelli tridimensionali e la trasformazione dei modelli tridimensionali in immagini planari.

Nel Capitolo 5 sono descritti i test effettuati ed i risultati ottenuti. Nel Capitolo 6, infine, sono riassunte le conclusioni a cui ha portato il lavoro svolto ed alcuni possibili sviluppi futuri.

TECNICHE DI RICOSTRUZIONE TRIDIMENSIONALE

In questo capitolo vengono trattate in modo sintetico le principali tecniche di ricostruzione tridimensionale, ovvero i metodi utilizzati per generare una rappresentazione tridimensionale virtuale di un oggetto fisico partendo da immagini bidimensionali. Le tecniche descritte appartengono a tre grandi categorie: luce strutturata, lama laser e tecniche basate su viste multiple.

La maggior parte dei metodi proposti ha in comune i concetti di triangolazione, nuvola di punti e *mesh*:

- La triangolazione è un procedimento che consente di determinare la profondità di un oggetto osservando la sua posizione da diverse angolazioni.
- Una nuvola di punti è un modello tridimensionale descritto da un insieme di coordinate in uno spazio tridimensionale.
- Una mesh è un insieme di vertici, spigoli e facce che definiscono la forma di un oggetto poliedrico rappresentato in tre dimensioni.

Prima di prendere in esame queste tecniche viene proposta un'introduzione ai concetti geometrici e matematici propri dell'acquisizione di immagini, oltre ad una descrizione dell'operazione di calibrazione dei dispositivi di acquisizione stessi.

2.1 Modello della fotocamera

Un modello geometrico di acquisizione di immagini, come spiegato in [1] e [2], descrive come un oggetto reale (in tre dimensioni) sia proiettato su un piano bidimensionale, cioè mappato da uno spazio $\mathbb{R}^3$ ad uno spazio $\mathbb{Z}^2$. La conoscenza di questo modello, ottenuta tramite il procedimento di calibrazione, fornisce gli strumenti necessari per ottenere una rappresentazione tridimensionale partendo da immagini bidimensionali, con un procedimento inverso alla calibrazione.

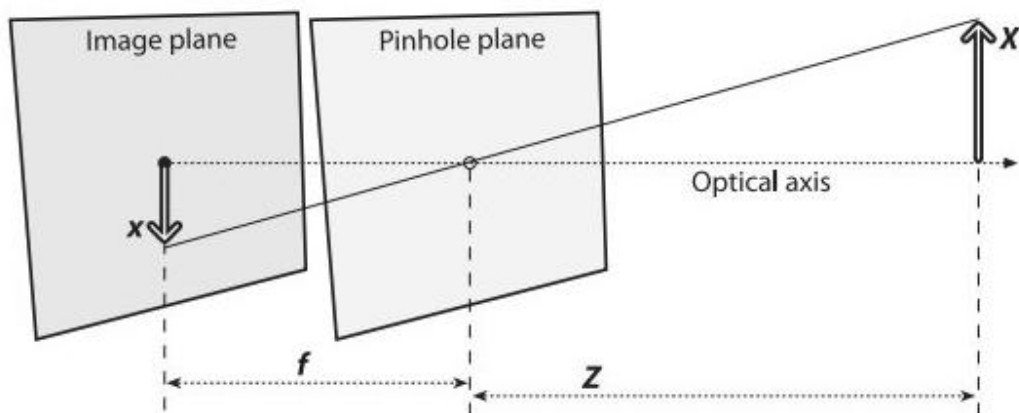
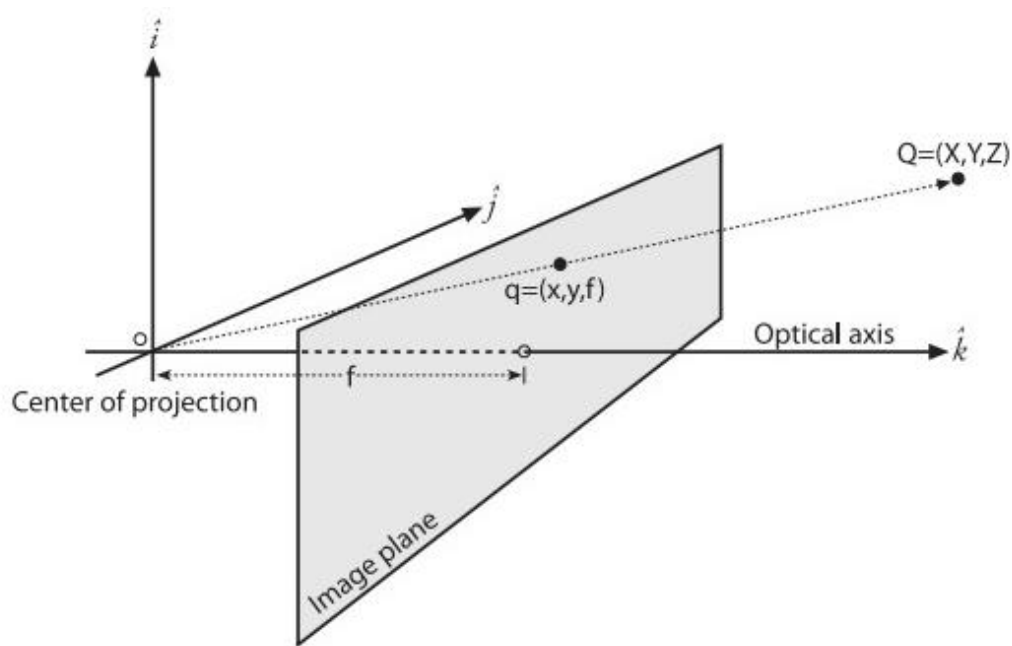
Prima di trattare i parametri che descrivono il modello più comunemente utilizzato, è necessario introdurre un modello semplificato, detto *pinhole model*. Come illustrato in Figura 2.1, in questo modello la fotocamera è priva di lenti ed è caratterizzata da una piccola apertura su una superficie (*pinhole plane*). Ciò permette che un solo raggio luminoso possa attraversare l'apertura per ogni punto dell'oggetto ed essere proiettato sul piano dell'immagine (*image plane*).

Nel *pinhole model*, un punto di altezza X nello spazio tridimensionale è proiettato in un punto di altezza x sul piano dell'immagine bidimensionale secondo la formula:

$$-x = f \frac{X}{Z} \quad , \quad (2.1)$$

dove f è la lunghezza focale della fotocamera, Z è la distanza dell'oggetto dalla fotocamera e X è l'altezza dell'oggetto. Il valore negativo nell'equazione 2.1 descrive il fatto che la proiezione di un oggetto risulta essere inversa rispetto alla profondità reale dell'oggetto stesso.

Tramite astrazione matematica, è possibile eliminare il segno negativo

Figura 2.1: *Pinhole model*Figura 2.2: *Pinhole model semplificato*

dall'equazione 2.1, ottenendo:

$$x = f \frac{X}{Z} \quad . \quad (2.2)$$

L'equazione 2.2 descrive il modello semplificato mostrato in Figura 2.2.

In questa semplificazione del pinhole model, il punto di intersezione tra il piano dell'immagine (*image plane*) e l'asse ottico (*optical axis*) è chiamato punto principale (*principal point*). Nelle fotocamere reali il punto principale non coincide necessariamente con il centro del piano dell'immagine perché ciò richiederebbe una precisione micrometrica nel posizionamento dell'elettronica situata dietro alla lente della fotocamera. È dunque possibile migliorare l'accuratezza del modello precedentemente descritto tramite l'ausilio di due parametri, chiamati c_x e c_y , che descrivono lo scostamento del punto principale rispetto all'asse ottico. Un punto tridimensionale di coordinate $Q = (X, Y, Z)$ è proiettato su un piano bidimensionale nel punto di coordinate $q = (x, y)$ secondo le equazioni:

$$x = f_x \frac{X}{Z} + c_x \quad , \quad y = f_y \frac{Y}{Z} + c_y \quad , \quad (2.3)$$

dove f_x e f_y sono le lunghezze focali relative all'asse delle ascisse e delle ordinate. Non è possibile utilizzare un solo parametro per la descrizione della lunghezza focale in quanto i pixel non rappresentano aree perfettamente quadrate. I valori f_x e f_y sono dati dal prodotto tra la lunghezza focale F e le dimensioni s_x e s_y degli elementi fisici: $f_x = Fs_x$ e $f_y = Fs_y$.

La relazione che mappa i punti dello spazio sull'immagine è detta trasformazione proiettiva (*projective transform*). Per esprimerla, è necessario introdurre il concetto di coordinate omogenee, descritte da un vettore di $(n + 1)$ elementi, dove n rappresenta il numero di dimensioni. Un vettore di coordinate bidimensionali (x, y) , ad esempio, è rappresentabile tramite coordinate omogenee in un vettore di 3 elementi $q = (q_1, q_2, q_3)$:

$$q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad . \quad (2.4)$$

Le coordinate effettive sono ricavabili tramite:

$$p = (x, y) = \left(\frac{q_1}{q_3}, \frac{q_2}{q_3} \right) \quad . \quad (2.5)$$

I parametri finora trattati consentono di definire la matrice descrittiva dei parametri propri della fotocamera (matrice intrinseca M):

$$M = \begin{bmatrix} f_x & \alpha f_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} , \quad (2.6)$$

in cui α definisce l'angolo tra gli assi x e y dell'immagine. Tramite questa matrice, i punti bidimensionali vengono calcolati secondo l'equazione:

$$q = MQ , \quad (2.7)$$

dove q rappresenta le coordinate omogenee del punto nel piano bidimensionale e Q rappresenta le coordinate del punto nello spazio tridimensionale.

Nella realtà, la luce che attraversa il pinhole non è sufficiente all'acquisizione di immagini di buona qualità. Al fine di ovviare al problema e concentrare un maggiore quantitativo di luce su un singolo punto, le fotocamere utilizzano delle lenti. Le lenti, però, introducono distorsione nelle immagini. È possibile distinguere due categorie di distorsione: radiale e tangenziale.

La distorsione radiale è dovuta al fatto che le lenti filtrano maggiormente la luce nelle zone ai bordi dell'area di acquisizione (Figura 2.3).

Questa distorsione è generalmente di piccola entità e può essere modellata tramite un'espansione in serie di Taylor nell'intorno di $r = 0$, dove r è la distanza dal centro dell'area di acquisizione. È quindi possibile correggere la posizione del punto sul piano dell'immagine tramite la formula:

$$\begin{aligned} x_{\text{rad}} &= x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) , \\ y_{\text{rad}} &= y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) , \end{aligned} \quad (2.8)$$

dove k_1 , k_2 e k_3 rappresentano i coefficienti di distorsione, che variano a seconda della fotocamera.

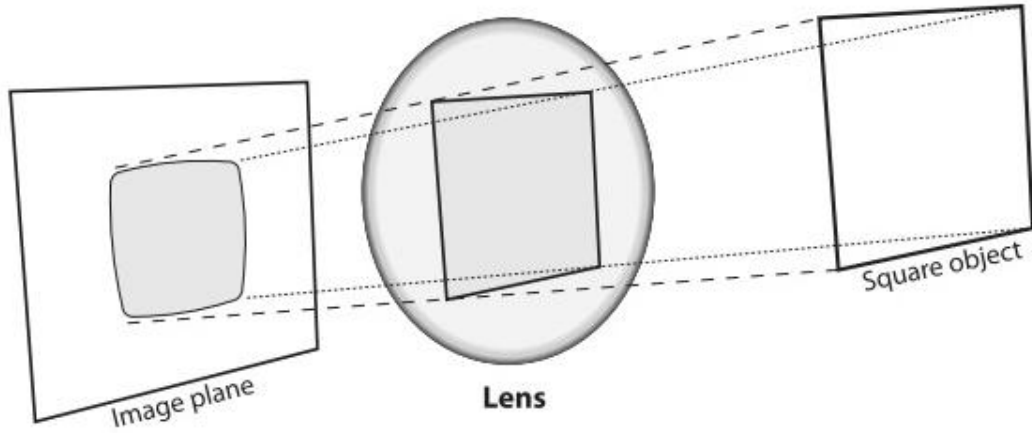


Figura 2.3: *Distorsione radiale*

La distorsione tangenziale dipende invece dal fatto che la lente delle fotocamere non è perfettamente parallela al piano dell'immagine, come mostrato in Figura 2.4. In modo analogo alla distorsione radiale, è possibile descrivere la distorsione tangenziale tramite i parametri p_1 e p_2 , che consentono di correggere le coordinate dei pixel dell'immagine secondo la formula:

$$\begin{aligned} x_{\text{tan}} &= x + [2p_1y + p_2(r^2 + 2x^2)] \quad , \\ y_{\text{tan}} &= y + [p_1(r^2 + 2y^2) + 2p_2x] \quad , \end{aligned} \quad (2.9)$$

I parametri $(k_1, k_2, k_3, p_1, p_2)$ utilizzati per descrivere la distorsione formano il vettore di distorsione (*distortion vector*).

2.1.1 Calibrazione di una fotocamera

La calibrazione è la procedura mediante la quale si calcolano i parametri intrinseci e di distorsione.

La calibrazione della fotocamera viene eseguita prima dell'utilizzo vero e proprio del sistema di acquisizione. Esistono diverse varianti di questa procedura, ma tipicamente questa operazione viene eseguita basandosi sull'utilizzo di pattern geometrici dei quali sono note con precisione le

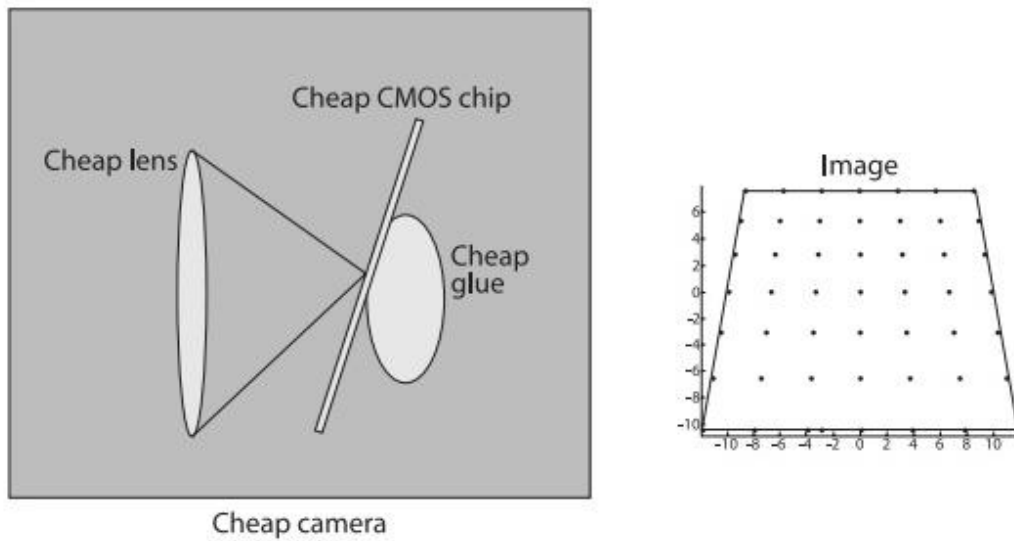


Figura 2.4: *Distorsione tangenziale*

caratteristiche. Mediante l'acquisizione di tali pattern è possibile stimare i parametri che caratterizzano il sistema.

È possibile effettuare la calibrazione di un sistema tramite oggetti tridimensionali oppure pattern stampati in immagini bidimensionali. L'utilizzo di immagini bidimensionali, però, richiede un determinato numero di acquisizioni del pattern stesso in differenti posizioni ed angolazioni.

Gli algoritmi di calibrazione più utilizzati sono il Direct Linear Transformation (DLT), il metodo proposto da Tsai in [3] e il metodo proposto da Zhang in [4]. In particolare, il metodo proposto da Zhang è il più recente ed effettua la calibrazione sfruttando immagini di scacchiere, e si compone di tre operazioni preliminari:

1. costruzione di una scacchiera perfettamente planare e rettangolare, con un numero di spigoli (*corner*) diverso nelle due dimensioni (in modo che l'orientamento sia univoco), in numero uguale o superiore a 6×5 (come mostrato in Figura 2.5);
2. acquisizione della scacchiera in almeno 10 posizioni diverse, con illuminazione uniforme;



Figura 2.5: *Esempio di pattern geometrico di calibrazione*

3. estrazione della posizione dei corner.

Successivamente viene eseguito un algoritmo divisibile nei seguenti passi:

1. calcolo dei parametri intrinseci ed estrinseci;
2. calcolo dei parametri di distorsione;
3. raffinamento della soluzione.

I metodi di Tsai e DLT, pur essendo parzialmente superati, possono vantare un elevato numero di implementazioni in svariati linguaggi, il che li rende tuttora molto usati. In particolare, il metodo di Tsai ha prestazioni superiori al metodo di Zhang nel caso di calibrazione di una fotocamera singola con forte distorsione radiale, come descritto in [5].

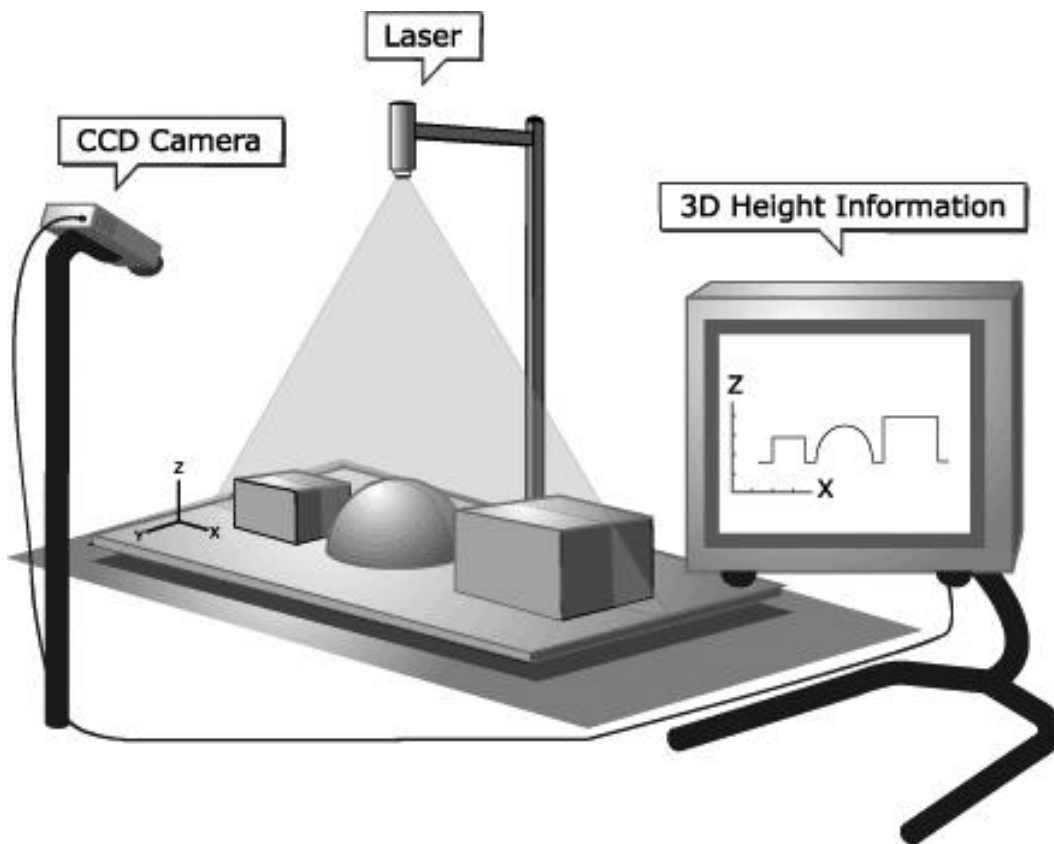


Figura 2.6: Funzionamento della luce strutturata

2.2 Sistemi basati su luce strutturata

Le tecniche basate su luce strutturata sfruttano la proiezione di un pattern di luce su di un oggetto; tale pattern può consistere in una lama laser (Figura 2.6) oppure in geometrie più complesse (Figura 2.7). Proiettando il pattern di luce su un oggetto ed analizzandone la deformazione, è possibile ricavarne la forma. Le deformazioni, infatti, sono riconducibili analiticamente alla geometria dell'oggetto.

Il setup più semplice che viene utilizzato nelle applicazioni reali consiste in un proiettore ed una fotocamera (Figura 2.8). Il processo di calcolo del modello tridimensionale si compone di quattro passi principali:

1. **Proiezione:** il proiettore imprime un pattern di luce sull'oggetto. Uno



Figura 2.7: *Utilizzo di un pattern più complesso per l'analisi di superfici*

dei pattern più usati consiste in una serie di strisce parallele, che rendono possibile l'analisi di più porzioni dell'oggetto in una sola immagine;

2. **Acquisizione:** l'oggetto ed il pattern proiettato vengono acquisiti da una fotocamera. La distanza della fotocamera dal proiettore influenza direttamente la deformazione del pattern;
3. **Decodifica:** le strisce deformate sono decodificate per essere distinte le une dalle altre;
4. **Ricostruzione:** le strisce decodificate sono analizzate singolarmente in modo da estrarne le informazioni sulla profondità dell'oggetto.

Questo approccio trova la sua utilità sia nel determinare la forma di un oggetto che nel rilevamento di imperfezioni (le asperità di una superficie risultano facilmente visibili con questo metodo). Può, inoltre, essere utilizzato per localizzare un certo oggetto (con geometria nota) nell'ambiente.

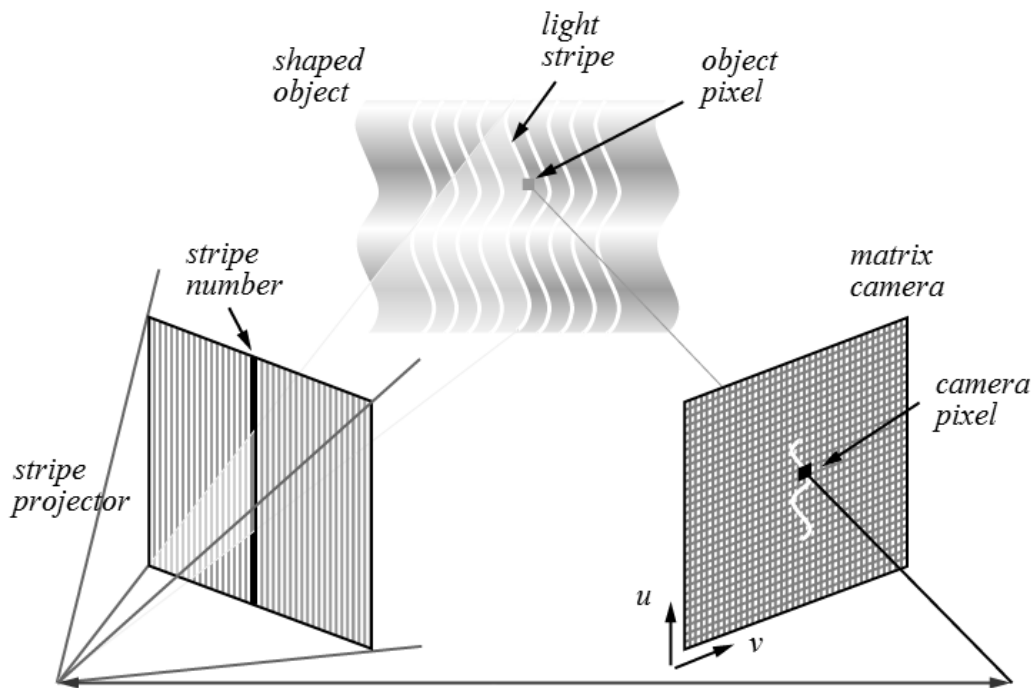


Figura 2.8: Setup di base per l'utilizzo della luce strutturata

Lo svantaggio della luce strutturata consiste nel fatto che, a differenza di altri approcci, necessita di apparecchiature di illuminazione dedicate.

2.3 Sistemi basati su lama laser

I sistemi basati su lama laser (detti anche *laser scanner*), utilizzano una luce laser per effettuare la scansione dell'oggetto e ricostruirne la struttura tridimensionale. Un laser è un'apparecchiatura che trasforma energia da una forma primaria (elettrica, ottica, chimica, termica o nucleare) in un fascio di radiazioni elettromagnetiche di intensità elevata. Tale fascio deve essere monocromatico, cioè contenere luce ad una sola frequenza, e coerente, cioè propagarsi in una sola direzione. Un fascio di luce monocromatico e coerente è in grado di propagarsi a grandi distanze.

Il dispositivo più semplice che sfrutta la luce laser per acquisire infor-

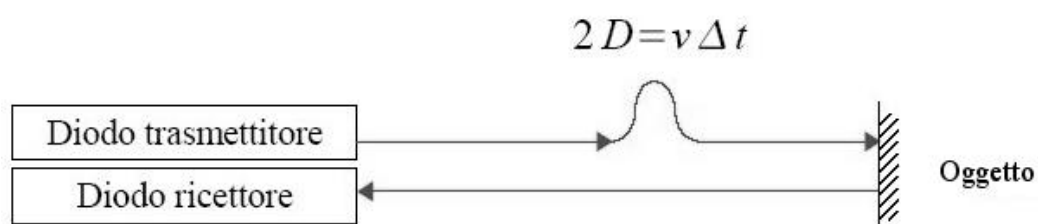


Figura 2.9: *Un distanziometro ad impulsi. La distanza D dell'oggetto dal trasmettitore è proporzionale al tempo di percorrenza degli impulsi*

mazioni su un punto è detto distanziometro. Come il nome suggerisce, tale dispositivo emette una luce laser e analizza la porzione di luce riflessa dall'oggetto per determinarne la distanza dalla sorgente. I distanziometri comunemente utilizzati ricavano questa informazione misurando il tempo trascorso tra l'invio di un impulso e la ricezione dell'impulso riflesso (Figura 2.9).

I laser scanner abbinano i distanziometri laser ad altri apparati meccanici come emettitori (che regolano la direzione della luce laser) e specchi (necessari per deviare la luce). Regolando l'emissione del fascio di luce laser su ogni punto della superficie dell'oggetto e analizzando il tempo trascorso per la scansione di ogni punto, si ottiene una nuvola di punti che rappresenta il primo passo per la costruzione di un modello tridimensionale.

In particolare, in [6] il procedimento di ricostruzione tridimensionale basato su lama laser è riassunto nei seguenti passi:

1. posizionamento dello scanner laser nella posizione desiderata;
2. acquisizione della nuvola di punti;
3. eliminazione degli errori di acquisizione;
4. fusione delle scansioni (nel caso siano state necessarie più acquisizioni separate);

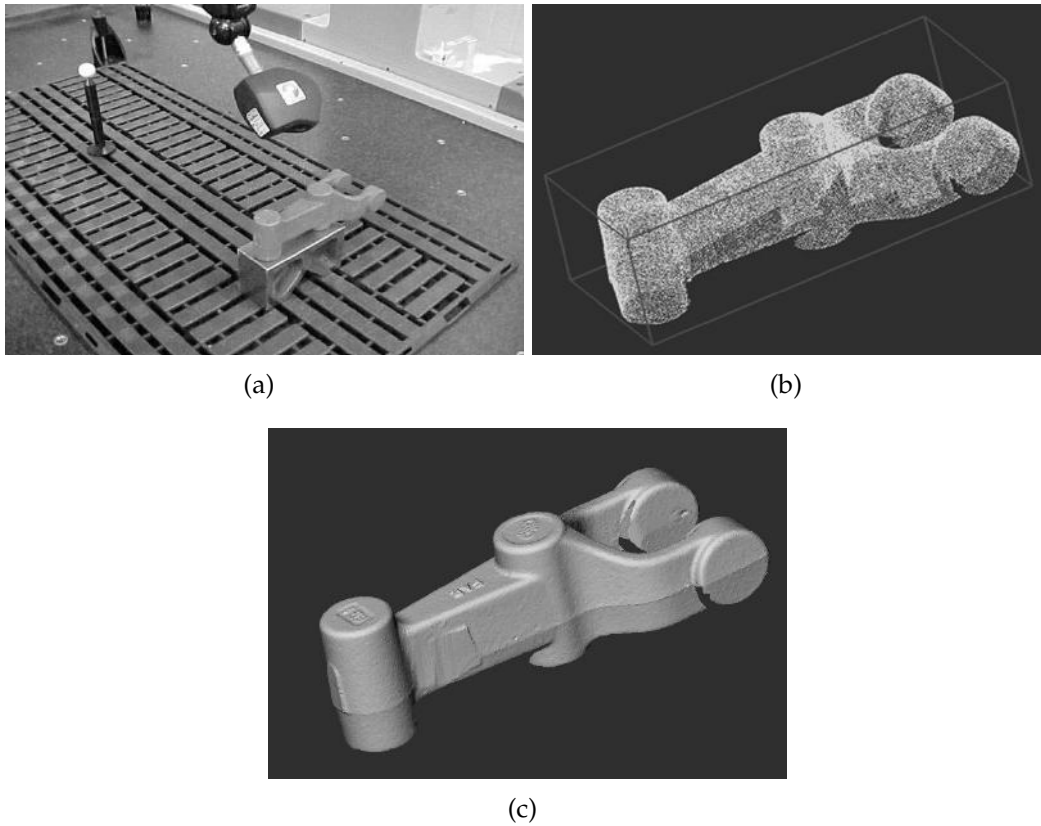


Figura 2.10: *Esempio di scanner laser: (a) scanner laser; (b) nuvola di punti; (c) modello tridimensionale*

5. costruzione del modello tridimensionale tramite creazione di una *mesh*;
6. colorazione del modello tridimensionale (opzionale).

In Figura 2.10 sono mostrati un esempio di scanner laser, la nuvola di punti ottenuta e il modello tridimensionale ricostruito.

2.4 Sistemi basati su viste multiple

2.4.1 Shape from silhouette

La tecnica di Shape from silhouette, come descritto in [7] e [8], consiste nell'utilizzo della proiezione dei contorni di un oggetto (*silhouette*) per inferire la geometria tridimensionale della scena. Le silhouette contengono l'informazione secondo cui si può conoscere se un particolare punto dell'oggetto è visibile dall'angolazione con cui è stata acquisita l'immagine.

Con il termine silhouette si intende un'immagine di un oggetto che ne visualizza solamente i contorni. Questa immagine è tipicamente un'immagine in bianco e nero: le zone di colore bianco corrispondono alle porzioni dell'immagine occupate dall'oggetto, mentre le zone di colore nero corrispondono allo sfondo.

Le silhouette, ottenute tramite metodi di segmentazione dell'immagine, sono proiettate in uno spazio tridimensionale virtuale secondo la posizione delle fotocamere con cui sono state acquisite. Successivamente sono intersecate al fine di ottenere l'approssimazione del volume tridimensionale dell'oggetto. Tale approssimazione, detta *Visual Hull* (Figura 2.12), è creata dall'intersezione di coni generalizzati (Figura 2.11). Un cono generalizzato è la superficie generata da un insieme di linee passanti per un unico vertice (la fotocamera) e tutti i punti di un contorno (la silhouette).

Intuitivamente, la precisione della ricostruzione aumenta con il numero di punti di vista. Risulta infatti valida la formula:

$$\bigcap_{i=1}^N B_i \subseteq \bigcap_{j=1}^M B_j \quad \text{con } N > M \quad , \quad (2.10)$$

dove B_i e B_j rappresentano due insiemi di silhouette e N e M rispettivamente il numero di silhouette disponibili in ogni insieme.

La definizione di Visual Hull [7] dice che, avendo a disposizione un infinito numero di punti di vista, esso rappresenta la migliore approssimazione possibile. Dal momento che è impossibile disporre di infinite immagini, l'approssimazione ottenibile con un numero finito di immagini è

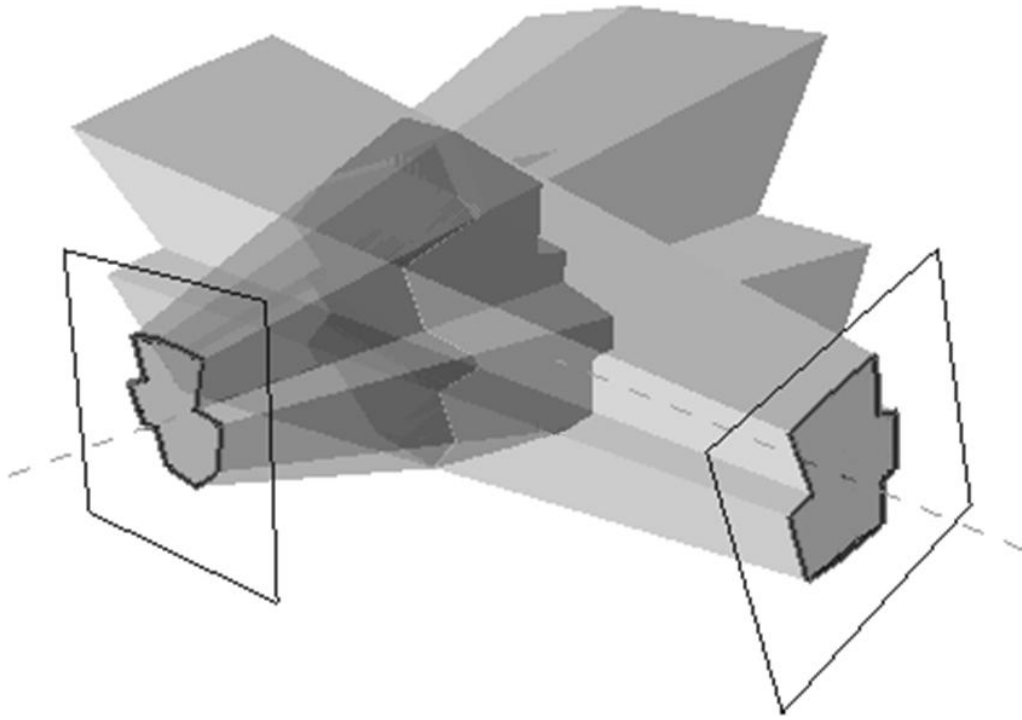


Figura 2.11: *Intersezione di coni generalizzati*

chiamata *Inferred Visual Hull*. In letteratura ([9]) esistono tecniche per determinare il migliore insieme di immagini da cui effettuare la ricostruzione, senza conoscenze pregresse sulla forma dell'oggetto. In [10], inoltre, si giunge alla conclusione che anche per ricostruire esattamente oggetti con un numero finito di superfici piane potrebbe essere necessario utilizzare un infinito numero di punti di vista.

La tecnica di Shape from silhouette è relativamente semplice e robusta, ma presenta alcuni difetti fondamentali:

- l'intera silhouette dell'oggetto deve essere contenuta in ogni immagine. In caso contrario, per le proprietà dell'operazione di intersezione, le parti mancanti vengono escluse e non sono ricostruite;
- l'intersezione di coni generalizzati introduce sempre degli spigoli. È

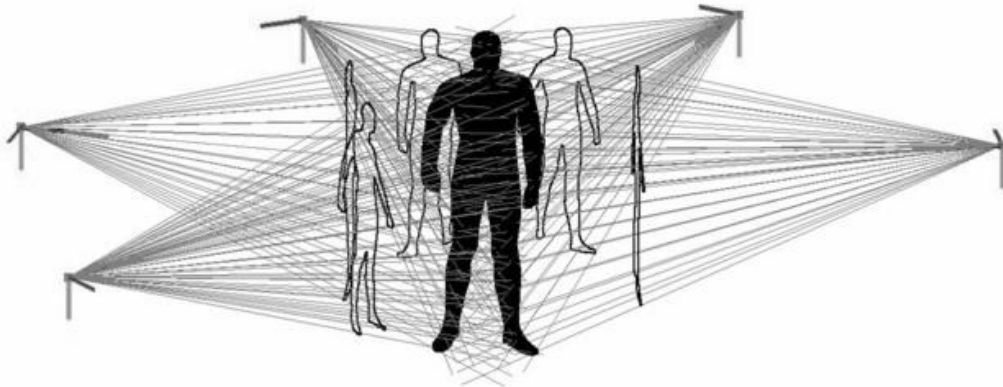


Figura 2.12: *Visual Hull*

possibile dunque ricostruire un modello tridimensionale con una precisione arbitrariamente alta (supponendo di usare infinite immagini) ma non in modo esatto;

- le forme concave non sono rappresentabili poichè la porzione concava degli oggetti non proietta ombra e quindi non influisce sulla silhouette. Non è quindi possibile ricostruire qualsiasi forma, anche assumendo infiniti punti di vista.

Risulta comunque possibile utilizzare altre tecniche per ovviare ai problemi tipici di quest'approccio. Per esempio si può usare la tecnica di Shape from silhouette per ottenere il Visual Hull dell'oggetto, e successivamente raffinare questa rappresentazione con tecniche stereoscopiche, come descritto in [11], oppure con tecniche di Shape from photo-consistency, come proposto in [12]

2.4.2 Shape from photo-consistency

La tecnica di Shape from photo-consistency, rispetto alla tecnica di Shape from silhouette, rappresenta un metodo più generale per la ricostruzione tridimensionale di un oggetto partendo da un insieme di immagini a colori o in scala di grigi.

Con la tecnica di Shape from photo-consistency è possibile sfruttare, oltre ai contorni, anche i valori dei pixel in ogni immagine per ottenere un modello tridimensionale migliore. Mentre le silhouette realizzano un vincolo binario sul modello tridimensionale, le immagini in scala di grigi (o a colori) permettono di definire vincoli più complessi.

Partendo dal modello tridimensionale costruito con la tecnica di Shape from silhouette, è possibile raffinare questo modello mantenendo solo i punti della superficie che rispettano un particolare tipo di vincoli. Si può dire che un punto sulla superficie di un oggetto rispetta il vincolo di photo-consistency con un insieme di immagini se è proiettato con lo stesso colore in ogni immagine in cui quel punto è visibile (a patto che l'illuminazione dell'oggetto sia uniforme in tutte le immagini).

L'insieme di punti che rispettano il vincolo di photo-consistency è detto *Photo Hull*.

2.4.3 Structure from motion

La tecnica di Structure from motion, descritta in [13] e [14], consiste nel ricostruire la struttura tridimensionale dell'oggetto analizzandone il movimento nel tempo (per esempio, è applicabile nei filmati). Gli esseri umani, infatti, percepiscono molte informazioni riguardo alla geometria di un ambiente muovendosi attraverso di esso. Questo principio, detto *Kinetic Depth Effect* ([15]), afferma che la forma tridimensionale di un oggetto visto in prospettiva è percepibile solo se l'oggetto è in rotazione.

La procedura per la ricostruzione tramite Structure from Motion, descritta in [16], è suddivisibile nei seguenti passi:

1. **Estrazione delle feature dalle immagini:** per facilitare la ricerca delle corrispondenze, le feature rappresentano punti facilmente riconoscibili (come bordi o angoli). I metodi di estrazione delle feature che si basano sul calcolo del gradiente sono suddivisibili nei seguenti passi:
 - prefiltraggio dell'immagine;

- calcolo del gradiente dell'immagine;
 - localizzazione e descrizione dei punti;
2. **Correlazione delle feature tra le immagini:** esistono numerose varianti degli algoritmi di ricerca delle correlazioni; essi hanno in comune le seguenti caratteristiche:
- effettuano la ricerca su coppie di immagini;
 - la ricerca avviene solo in un intorno delle coordinate del primo punto;
 - i punti devono essere vicendevolmente correlati;
3. **Calibrazione delle fotocamere;**
4. **Ricostruzione iniziale basata sui punti estratti:**
- triangolazione dei punti correlati;
 - unificazione delle triangolazioni;
5. **Creazione di una rappresentazione densa:** conoscendo la posizione dei punti correlati, è possibile stimare anche le traiettorie dei restanti punti, ottenendo una ricostruzione densa. Il procedimento di creazione di una rappresentazione densa si divide nei seguenti passi principali:
- correzione delle immagini basata sui dati di calibrazione;
 - ricerca delle correlazioni su tutta l'immagine;
 - triangolazione di tutti i punti correlati;
6. **Correzione del modello:**
- applicazione delle informazioni sulla geometria della scena;
 - applicazione delle *texture*;
 - applicazione dell'illuminazione.

I setup possibili per quest'approccio sono molteplici: è possibile muovere la telecamera, usare telecamere multiple o muovere il soggetto.

Il problema principale di quest'approccio è la ricerca delle corrispondenze tra le immagini, ovvero la conoscenza della posizione di uno stesso punto nell'insieme di immagini che si hanno a disposizione. ,

2.4.4 Sistemi basati su immagini stereoscopiche

La tecnica di ricostruzione tridimensionale stereoscopica si ispira direttamente al sistema di visione umano e permette di elaborare le proprietà geometriche della scena fondendo le informazioni ottenute da due punti di vista leggermente separati.

La visione stereoscopica non impone vincoli sulle caratteristiche degli oggetti della scena: non devono avere particolari caratteristiche di forma (Shape from silhouette), di illuminazione (Shape from photo-consistency) o di movimento (Structure from motion).

Nelle successive sezioni, come descritto in [1] e in [2], vengono estesi i concetti geometrici visti nella sezione 2.1 ed adattati ad un sistema stereoscopico. Infine sono descritte in dettaglio le varie fasi che compongono una procedura di ricostruzione stereoscopica, brevemente riassumibili in quattro passaggi:

1. **Calibrazione:** la procedura mediante la quale si determinano i parametri della fotocamera e la posizione dell'oggetto rispetto alle telecamere.
2. **Rettificazione:** procedura che utilizza i dati ottenuti tramite calibrazione per correggere le immagini in modo che siano e adatte per la ricostruzione tridimensionale.
3. **Corrispondenza:** consiste nella ricerca delle correlazioni tra le immagini.

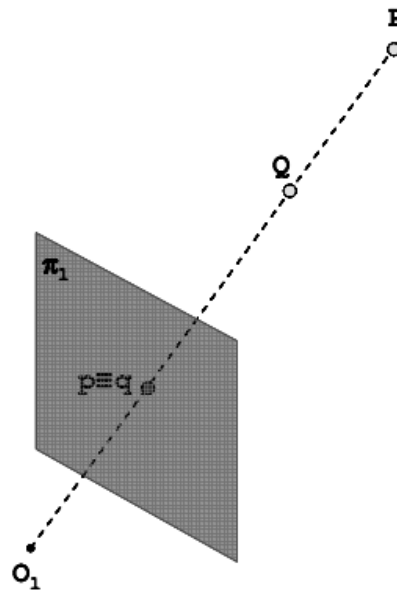


Figura 2.13: *Perdita di informazione relativa alla profondità*

4. **Triangolazione:** permette di determinare la profondità di un punto nello spazio tridimensionale basandosi sulla sua posizione nelle due immagini acquisite dal sistema stereoscopico.

2.4.4.1 Geometria di un sistema stereoscopico

La trasformazione prospettica che mappa i punti dello spazio tridimensionale in punti sull'immagine comporta necessariamente la perdita dell'informazione relativa alla distanza dell'oggetto dalla fotocamera. In Figura 2.13 viene mostrato come i punti P e Q abbiano la stessa proiezione $p \equiv q$ sul piano dell'immagine π_1 corrispondente al centro ottico O_1 .

Una tecnica per determinare a quale punto dello spazio tridimensionale corrisponda la proiezione di un punto sul piano dell'immagine consiste nell'applicare il metodo di triangolazione dei punti corrispondenti in immagini acquisite da telecamere poste a differenti coordinate spaziali. Come si vede in Figura 2.14, nel caso di un sistema composto da due fotocamere, i punti P e Q che giacciono sul raggio che passa per il centro ottico O_l e

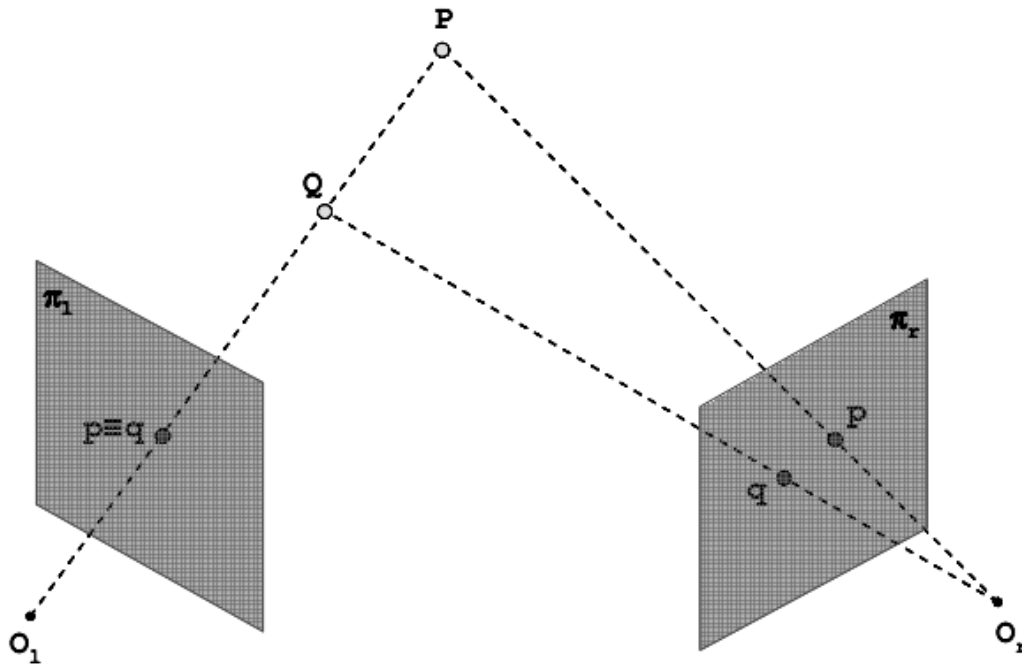


Figura 2.14: Usando due fotocamere l'informazione relativa alla profondità non viene persa

il punto Q , hanno la stessa proiezione $p \equiv q$ sul piano immagine π_l , ma proiezioni distinte sul piano immagine π_r . La determinazione dei punti omologhi nei due piani consente di mettere in relazione le proiezioni dello stesso punto sui due piani immagine e di risalire, mediante triangolazione, alle coordinate dei punti dello spazio tridimensionale rispetto ad un sistema di riferimento opportuno.

Uno dei principali problemi consiste, quindi, nel trovare il punto corrispondente q' nel piano π_r , dato un punto q nel piano π_l . Questa ricerca può essere ridotta ad un'analisi monodimensionale sfruttando un principio del sistema stereoscopico. Come mostrato nella Figura 2.15, gli omologhi di tutti i punti dello spazio che sono possibili proiezioni dello stesso punto q del piano immagine π_l giacciono sulla linea epipolare (*epipolar line*) ottenuta dall'intersezione tra il piano immagine e il piano epipolare, definito dal centro di proiezione P e dai due centri ottici O_l e O_r . L'intersezione

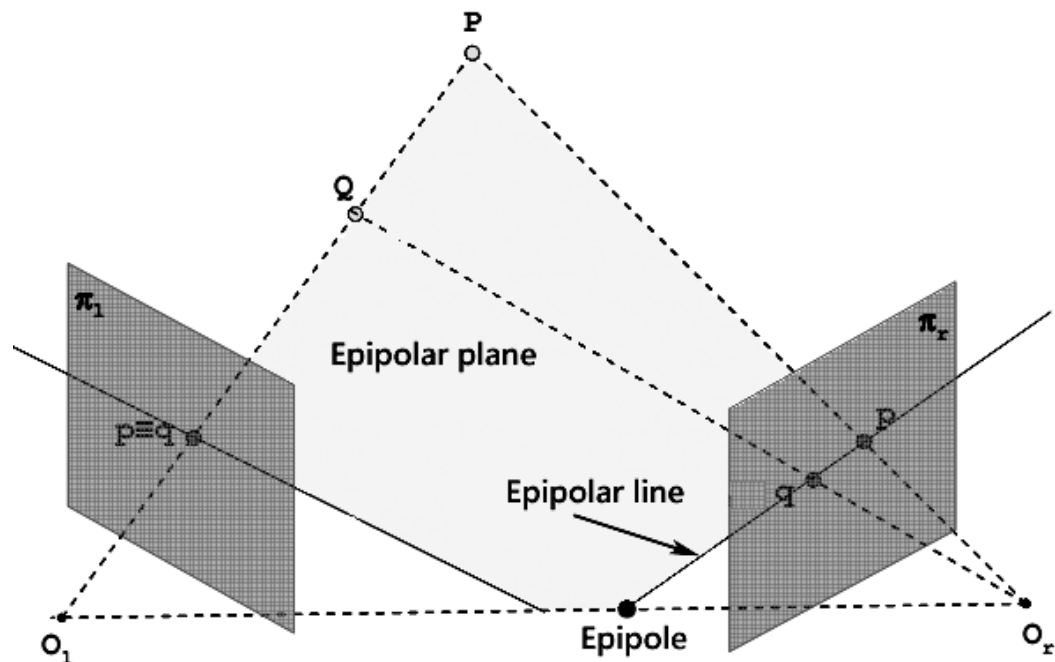


Figura 2.15: *Vincolo epipolare*

della linea epipolare con la linea che unisce i centri ottici è detto epipolo.

Tale vincolo, denominato vincolo epipolare, consente di limitare lo spazio di ricerca dei punti omologhi ad un segmento, semplificando il problema delle corrispondenze sia dal punto di vista della complessità algoritmica che dal punto di vista della correttezza della soluzione.

Il problema della ricerca delle corrispondenze non ha una soluzione per tutti i punti delle immagini. A causa della diversa posizione delle telecamere che compongono un sistema di visione stereoscopico nello spazio è possibile che un punto non risulti proiettato su tutti i piani immagini delle telecamere (problema delle occlusioni).

2.4.4.2 Acquisizione delle immagini

La disposizione fisica delle telecamere di un sistema stereoscopico può essere effettuata in differenti modalità: due telecamere allineate verticalmente, orizzontalmente, più telecamere (con diverse angolazioni) che cir-



Figura 2.16: Esempio di sistema per l'acquisizione di immagini stereoscopiche

condano l'oggetto, etc. Nella sua forma più semplice, comunque, la visione stereoscopica riproduce la conformazione degli occhi umani: due fotocamere allineate orizzontalmente e separate da una distanza nota (*baseline*), come si vede nell'esempio in Figura 2.16, tratto da [17].

In molti sistemi, infatti, la baseline corrisponde alla distanza media tra gli occhi umani (ca. 6.35 cm). La lunghezza della baseline risulta determinante per la qualità dei dettagli ricostruiti in quanto la differenza della distanza di un punto nelle immagini stereoscopiche (chiamata *disparity*) è strettamente legata alla lunghezza della baseline. In 2.17 viene mostrato come la disparity diminuisca progressivamente con la distanza dell'oggetto dalla telecamera. Quando la disparity si avvicina a zero, le due immagini stereoscopiche risultano molto simili, e l'informazione relativa alla profondità viene persa. Per questo motivo, i sistemi stereoscopici hanno una buona precisione solo per oggetti relativamente vicini alla fotocamera.

Una baseline maggiore del necessario può essere usata per produrre effetti *iper-stereo*, cioè ricostruzioni tridimensionali che danno una percezione della profondità non realistica ma aumentata rispetto alla normalità. La baseline, quindi, deve essere aumentata all'aumentare della distanza dell'oggetto dalla telecamera, e diminuita in caso contrario.

Un compromesso spesso usato nella progettazione di sistemi di acqui-

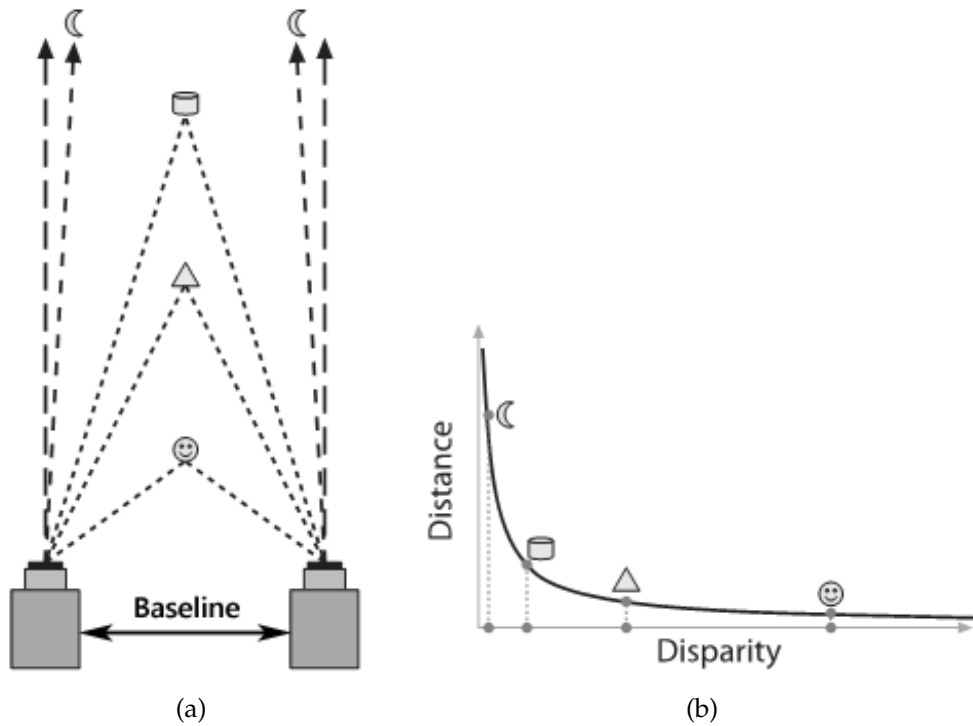


Figura 2.17: La relazione tra la distanza dell'oggetto e la disparity: (a) oggetti posizionati a varie distanze dalle fotocamere; (b) relazione tra distanza e disparity

sizione stereoscopici è chiamato regola del 1/30: nel caso in cui l'oggetto sia ad una distanza di almeno 1.8 m dalla fotocamera, la baseline deve essere 1/30 della distanza dell'oggetto dalla telecamera.

A seconda delle telecamere utilizzate e della distanza tra di esse, inoltre, cambia l'ampiezza dell'area di acquisizione utile. Viene considerato il campo di visione (*Field of View*), definito come l'intersezione, in senso verticale ed orizzontale, del campo di visione delle due fotocamere. Il campo di visione verticale e il campo di visione orizzontale sono calcolabili secondo le formule:

$$\begin{aligned}
 V_{FOV} &= \arctan \frac{c_y}{f_y} + \arctan \frac{H - 1 - c_y}{f_y} , \\
 H_{FOV} &= \arctan \frac{c_x}{f_x} + \arctan \frac{H - 1 - c_x}{f_x} ,
 \end{aligned} \tag{2.11}$$

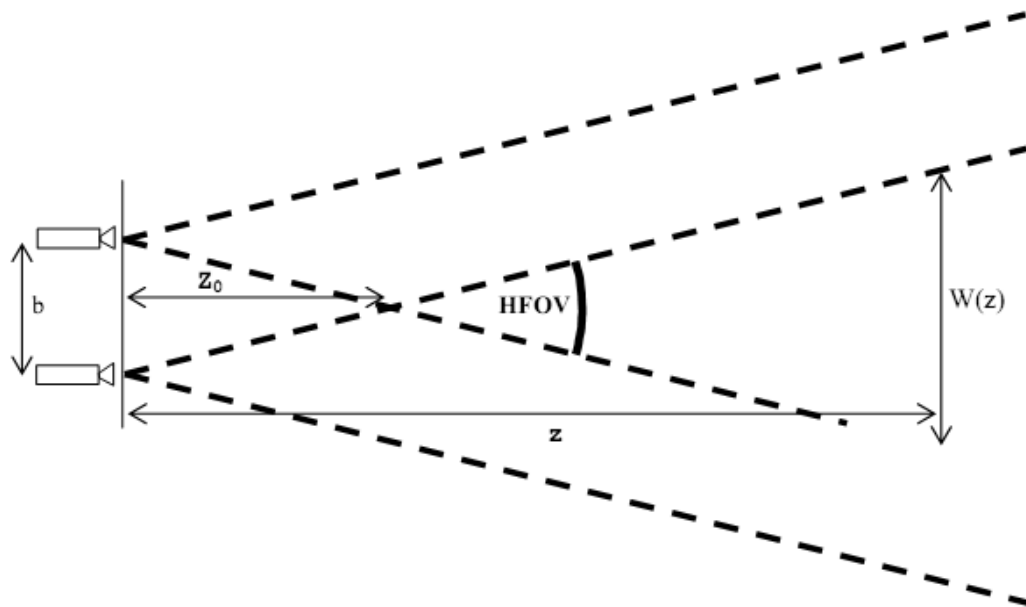


Figura 2.18: *Field of View.* La forma dei campi visivi dipende dalla lunghezza focale

dove V_{FOV} e H_{FOV} rappresentano rispettivamente il campo di visione verticale (*Vertical Field of View*) ed il campo di visione orizzontale (*Horizontal Field of View*), W e H sono rispettivamente la larghezza e l'altezza dell'immagine.

Infine, l'acquisizione delle immagini deve essere simultanea. Questo vincolo può essere eliminato nel caso di scene statiche. L'acquisizione simultanea di più immagini provenienti da una sorgente video analogica o digitale può avvenire mediante diverse tecnologie. Nel caso di segnali video analogici si utilizzano *frame-grabbers*, dispositivi capaci di acquisire immagini da più sorgenti video.

2.4.4.3 Calibrazione di un sistema stereoscopico

La calibrazione di un sistema stereoscopico consiste nella calibrazione indipendente delle due fotocamere con il procedimento presentato nella Sezione 2.1.1. Risulta inoltre necessario trattare i parametri descrittivi della po-

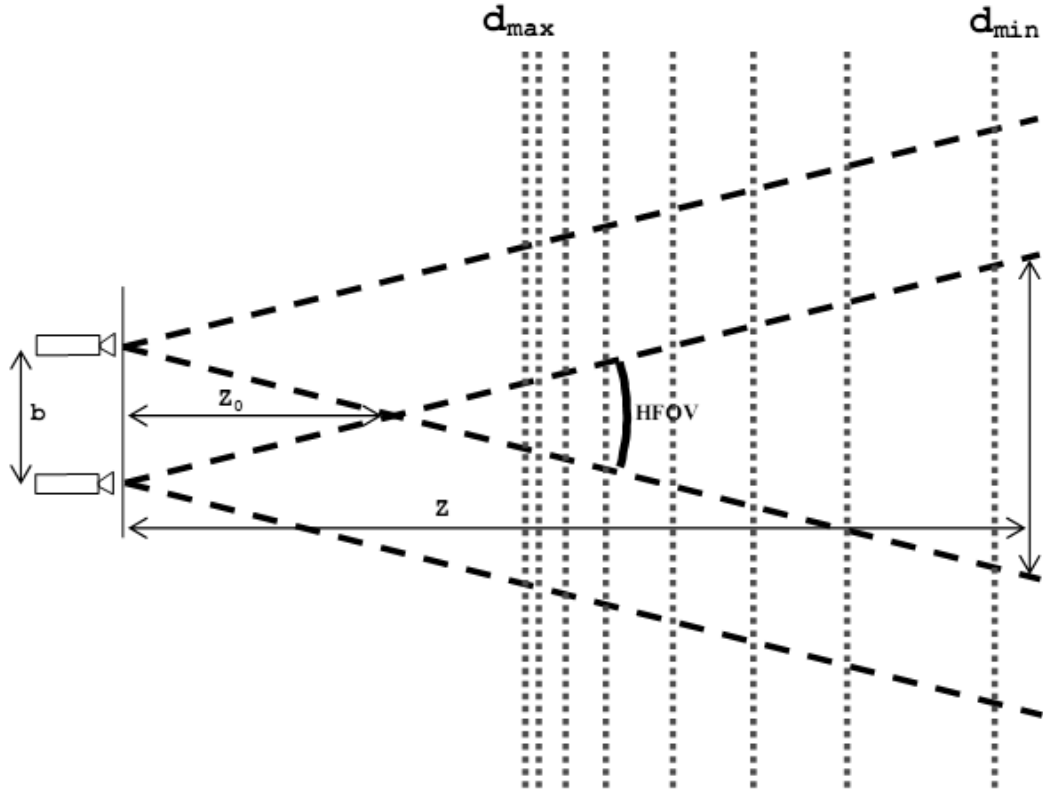


Figura 2.19: Area utile per il calcolo delle disparità

sizione dell'oggetto rispetto alla fotocamera. Il sistema di riferimento dell'oggetto, come mostrato in Figura 2.20, è ruotato e traslato rispetto al sistema di riferimento della fotocamera. I parametri che descrivono la rotazione e la traslazione dell'oggetto sono chiamati parametri estrinseci.

Una rotazione in un qualsiasi numero di dimensioni può essere espressa come la moltiplicazione di un vettore di coordinate per una matrice quadrata. Descrivere la rotazione dell'oggetto equivale a descrivere la rotazione del suo sistema di riferimento rispetto al sistema di riferimento della fotocamera.

Matematicamente, una rotazione completa R è la combinazione delle rotazioni lungo ciascun asse. Vale quindi:

$$R = R_x(\psi)R_y(\phi)R_z(\theta) \quad , \quad (2.12)$$

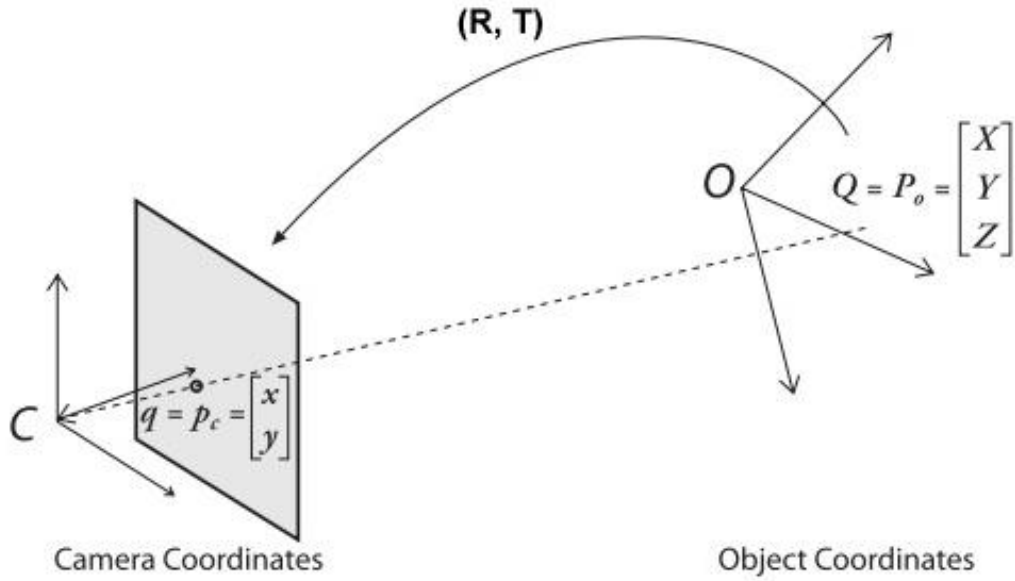


Figura 2.20: Rotazione R e traslazione T del sistema di riferimento dell'oggetto (con origine O) rispetto al sistema di riferimento della fotocamera (con origine C)

dove ψ , ϕ e θ sono i tre angoli di rotazione e $R_x(\psi)$, $R_y(\phi)$ e $R_z(\theta)$ le rispettive matrici di rotazione:

$$R_x(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & \sin \psi \\ 0 & -\sin \psi & \cos \psi \end{bmatrix} ,$$

$$R_y(\phi) = \begin{bmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{bmatrix} ,$$

$$R_z(\theta) = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} . \quad (2.13)$$

Una traslazione è rappresentata come lo spostamento delle origini dei due sistemi di riferimento:

$$T = O_{\text{obj}} - O_{\text{cam}} \quad , \quad (2.14)$$

dove O_{obj} rappresenta le coordinate dell'origine del sistema di riferimento dell'oggetto e O_{cam} rappresenta le coordinate dell'origine del sistema di riferimento della fotocamera.

Un oggetto di coordinate reali P_o ha coordinate P_c nel sistema di riferimento della fotocamera, calcolate come:

$$P_c = R(P_o - T) \quad . \quad (2.15)$$

Risulta quindi necessario determinare la rotazione di una fotocamera rispetto all'altra, ottenibile tramite la formula:

$$R = R_r(R_l)^T \quad , \quad (2.16)$$

dove R_l , T_l , R_r , e T_r rappresentano le matrici di rotazione ed i vettori di traslazione dell'oggetto relativamente alla fotocamera sinistra ed alla fotocamera destra. In questo caso R e T rappresentano la rotazione e la traslazione della fotocamera destra rispetto a quella sinistra. Poichè le matrici di rotazione hanno la proprietà $A^T = A^{-1}$, risulta che la matrice R_r è ruotata in senso contrario alla rotazione R_l , con il risultato che la matrice R rappresenta la differenza tra le due rotazioni. Se le matrici R_l e R_r sono equivalenti, infatti, vale:

$$\begin{aligned} R_l &= R_r \\ R &= R_r R_r^{-1} = I \quad , \end{aligned} \quad (2.17)$$

dove I è la matrice identità che corrisponde ad una rotazione nulla.

La traslazione della fotocamera destra rispetto alla fotocamera sinistra è ottenibile tramite la formula:

$$T = T_r - RT_l \quad , \quad (2.18)$$

dove RT_l è il vettore di traslazione della fotocamera sinistra ruotato in modo da essere adattato al sistema di riferimento della fotocamera destra.

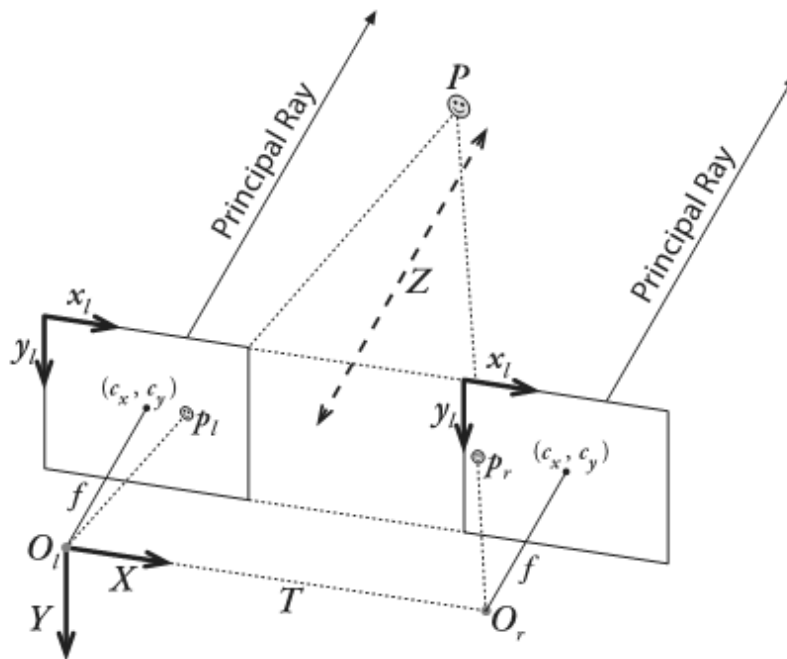


Figura 2.21: Sistema stereo (virtuale) dopo la rettificazione

2.4.4.4 Rettificazione delle immagini

La rettificazione è una procedura che trasforma le immagini stereoscopiche in modo che risultino come acquisite da due fotocamere perfettamente coplanari ed allineate, come mostrato in Figura 2.21, sfruttando i parametri intrinseci ed estrinseci ottenuti mediante la calibrazione.

In un sistema così trasformato vale il vincolo epipolare descritto in 2.4.4.1. La rettificazione è infatti il procedimento tramite il quale vengono compensati i problemi derivanti dalle non-idealità delle fotocamere e dalle distorsioni introdotte dalle ottiche. In un sistema stereoscopico rettificato i raggi principali (definiti come le rette che collegano i centri ottici delle due fotocamere con i rispettivi punti principali) sono paralleli. A titolo di esempio, in Figura 2.22 sono mostrate due immagini prima e dopo la procedura di rettificazione.

L'algoritmo più utilizzato per calcolare i parametri di rettificazione di un'immagine è l'algoritmo di Bouguet ([18]), derivante da una raffinazione

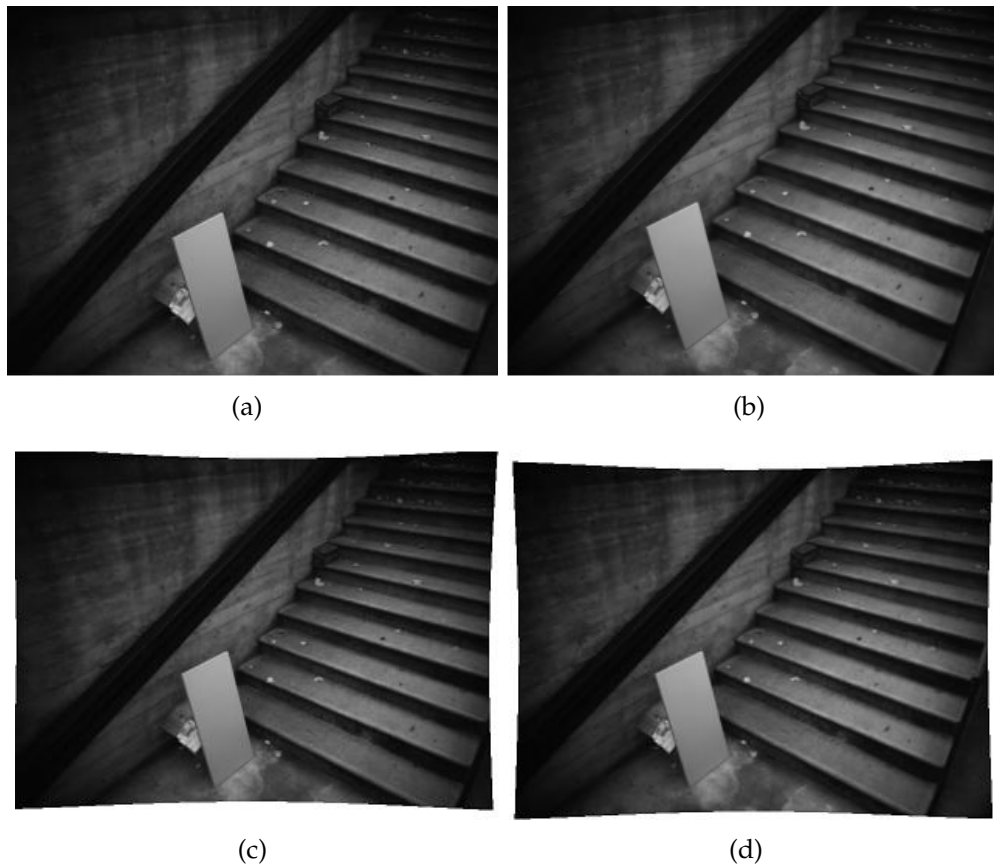


Figura 2.22: *Cambiamenti apportati dalla rettificazione: (a),(b) immagini prima della rettificazione; (c),(d) immagini dopo la rettificazione*

dei metodi proposti da Tsai in [3] e da Zhang in [4]. L'algoritmo di Bouguet si basa sulla matrice di rotazione R e sul vettore di traslazione T calcolati tramite la calibrazione stereoscopica e consiste di tre passi principali:

1. rotazione delle due fotocamere in modo da renderle coplanari;
2. allineamento orizzontale delle fotocamere;
3. massimizzazione dell'area comune tra le immagini.

2.4.4.5 Corrispondenza stereoscopica

La letteratura relativa agli algoritmi di correlazione è estremamente vasta. È possibile, in prima approssimazione, classificare gli algoritmi in due categorie principali:

- algoritmi basati su feature;
- algoritmi di tipo denso.

Gli algoritmi appartenenti a ciascuna delle due categorie hanno come obiettivo la creazione di una mappa di disparità (*disparity map*), composta dall'insieme delle distanze tra i punti correlati.

Gli algoritmi appartenenti alla prima categoria consentono di ottenere la mappa di disparità tra i punti per un numero limitato di punti delle immagini, corrispondenti a feature particolari e facilmente identificabili (come linee, segmenti o angoli). Questi algoritmi, però, sono poco utilizzati in quanto le mappe di disparità risultano limitate e contengono solo pochi particolari punti.

Gli algoritmi di tipo denso, invece, generano mappe di disparità per ogni punto dell'immagine. Secondo una recente tassonomia, descritta in [19], è possibile classificare gli algoritmi di *dense matching* in due tipologie:

- algoritmi di tipo locale;
- algoritmi di tipo globale.

Gli algoritmi di tipo locale ricercano i punti omologhi singolarmente ed indipendentemente gli uni dagli altri, analizzando l'intorno di ogni pixel (supporto locale). Questo metodo, detto anche della finestra di correlazione, ha lo scopo di aumentare il rapporto segnale/rumore. Gli algoritmi di tipo locale assumono che all'interno di ogni supporto locale la disparità sia costante. Questa ipotesi, però, non è sempre verificata e conduce ad una discrepanza tra gli oggetti della scena e le mappe di disparità in prossimità di punti posti a differente distanza dalle fotocamere.

È importante osservare che esistono anche algoritmi locali che adattano il proprio supporto in base alle caratteristiche locali delle immagini ([20]). In questo modo è possibile ottenere risultati migliori ad un costo computazionale più alto rispetto agli algoritmi locali standard.

Gli algoritmi locali, inoltre, risultano implementabili in modo relativamente semplice anche a livello hardware, utilizzabili applicazioni real-time.

Questi algoritmi, comunque, hanno un limite legato al fatto che non sempre sono in grado di produrre mappe di disparità completamente dense. Ciò è dovuto alla presenza di regioni uniformi e di occlusioni. Nel primo caso non è possibile determinare informazioni di disparità in regioni che presentano intensità costante (o comunque pochi elementi distintivi) a causa del fatto che risulta difficile identificare con precisione un punto all'interno di tali regioni. Nel secondo caso, alcuni punti possono risultare visibili in una sola immagine, rendendo impossibile risolvere la corrispondenza tra i pixel delle due immagini stereoscopiche.

Diversamente dagli algoritmi densi di tipo locale, gli algoritmi densi di tipo globale non utilizzano alcun supporto locale, ma si basano su alcune ipotesi sulla natura degli oggetti (*smoothness assumption*) e ricercano un valore di disparità per i punti in modo da minimizzare l'errore globale. Gli algoritmi densi di tipo globale consentono spesso di ottenere mappe di disparità migliori rispetto agli algoritmi di tipo locale, con lo svantaggio di avere un costo computazionale più elevato (il che li rende attualmente non utilizzabili in applicazioni real-time).

Gli algoritmi densi di tipo globale non producono necessariamente mappe di disparità completamente dense.

2.4.4.6 Triangolazione

La triangolazione permette di ricavare la profondità di un punto nello spazio tridimensionale conoscendo le posizioni della proiezione del punto nelle due immagini.

Il metodo, come mostrato in Figura 2.23, sfrutta il principio dei triangoli

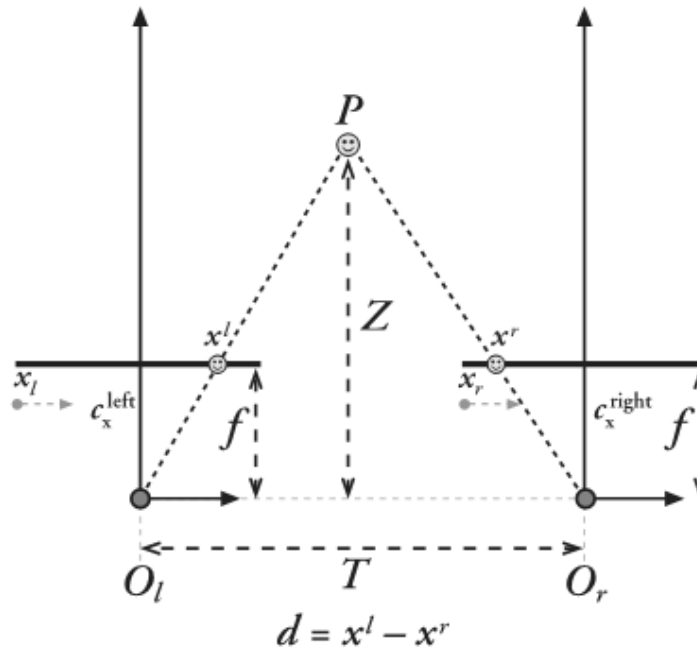


Figura 2.23: Triangolazione a partire dal valore di disparità

simili. La profondità z è data dalla formula:

$$z = \frac{fT}{x_l - x_r} \quad , \quad (2.19)$$

dove T è la distanza tra i centri ottici delle due fotocamere, x_l e x_r sono le posizioni delle proiezioni del punto P nelle due immagini ed f è la lunghezza focale.

Dalla formula 2.19 risulta che la distanza tra i due punti $d = x_l - x_r$ è inversamente proporzionale alla profondità Z .

2.5 Rappresentazione della scena nello spazio tridimensionale

Per rappresentazione della scena si intende lo spazio virtuale che descrive l'oggetto tridimensionale ricostruito. La maggioranza delle attuali appli-

cazioni utilizza una griglia tridimensionale, i cui punti sono descritti da una terna di coordinate.

Lo spazio occupato all'interno di questa griglia può essere rappresentato tramite *voxel* oppure da *level-sets*. I *voxel* (*volumetric pixel*) sono l'equivalente tridimensionale dei pixel. Come i pixel, la posizione dei voxel non è memorizzata esplicitamente ma è definita in relazione alla posizione degli altri elementi.

I *level-sets* sfruttano una funzione che codifica una distanza e sono adatti per rappresentare forme che variano nel tempo. Le immagini (a),(b),(c) della Figura 2.24 mostrano l'evolversi nel tempo di una scena: questa scena può essere intesa come generata dall'intersezione tra un piano e un volume tridimensionale di sezione crescente. Le immagini (d),(e),(f) della Figura 2.24 mostrano come sia sufficiente modificare il valore dell'altezza del piano rispetto al volume per tenere conto del cambiamento.

Un altro metodo di rappresentazione consiste nell'utilizzo dei *polygon-mesh*, un insieme di facce planari connesse tra di loro. La scena è descritta da un numero discreto di punti, connessi tra di loro formando un insieme di poligoni (che possono essere triangoli, quadrilateri, ecc). Uno dei metodi più utilizzati per il calcolo di un mesh di triangoli (Figura 2.25) è rappresentato dalla triangolazione di Delaunay, che definisce un insieme di connessioni tra i punti in modo da massimizzare l'angolo minimo dei triangoli risultanti.

Il numero di poligoni in relazione alla superficie dell'oggetto determina la precisione del modello: maggiore è il numero di punti e di poligoni, più il modello descritto dal mesh è simile all'oggetto reale. Questo metodo risulta efficiente in termini di memorizzazione e visualizzazione, in quanto anche per un numero di punti non elevato può garantire una buona approssimazione.

Risulta infine possibile codificare la struttura tridimensionale della scena tramite *disparity map*, che ha il notevole vantaggio di memorizzare le informazioni di un oggetto in tre dimensioni tramite immagini in due

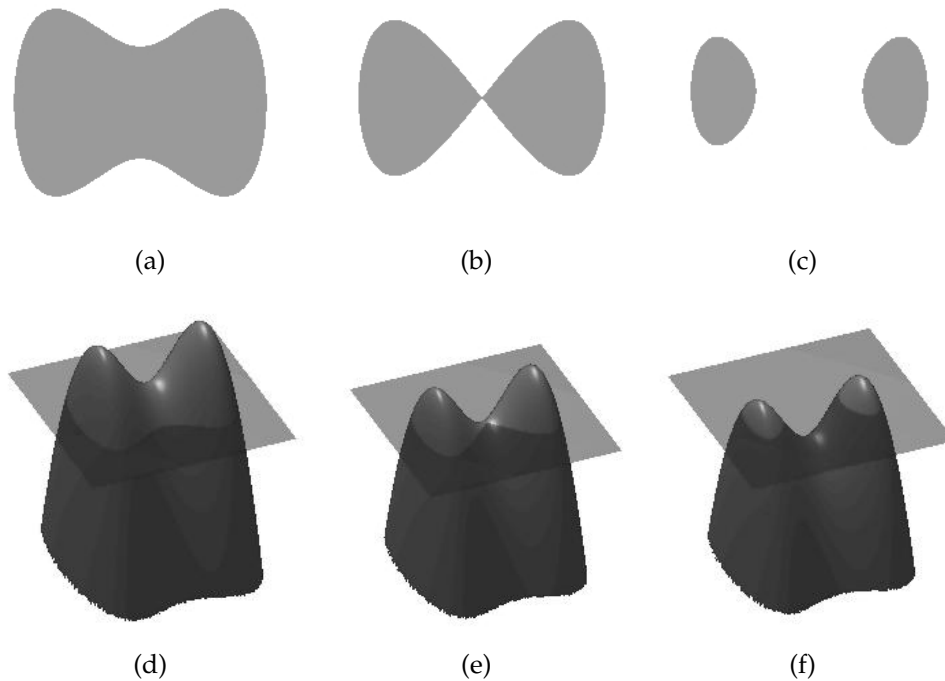


Figura 2.24: Esempio di rappresentazione di una scena tramite level-set: (a),(b),(c) scena che varia nel tempo; (d),(e),(f) variazione dell'altezza del piano intersecante il volume tridimensionale

dimensioni. Le disparity map sono create a partire da una coppia di immagini stereoscopiche (Figura 2.26a e Figura 2.26b): ogni punto dell'oggetto tridimensionale è acquisito nelle due immagini in una posizione differente. La disparity map contiene la distanza tra le posizioni dei due pixel corrispondenti nelle due immagini. Questa distanza è inversamente proporzionale alla profondità del punto dell'oggetto (Figura 2.26c).

Una disparity map permette di ricostruire la posizione del punto in una griglia tridimensionale solo nel caso in cui siano noti i parametri di calibrazione delle telecamere.

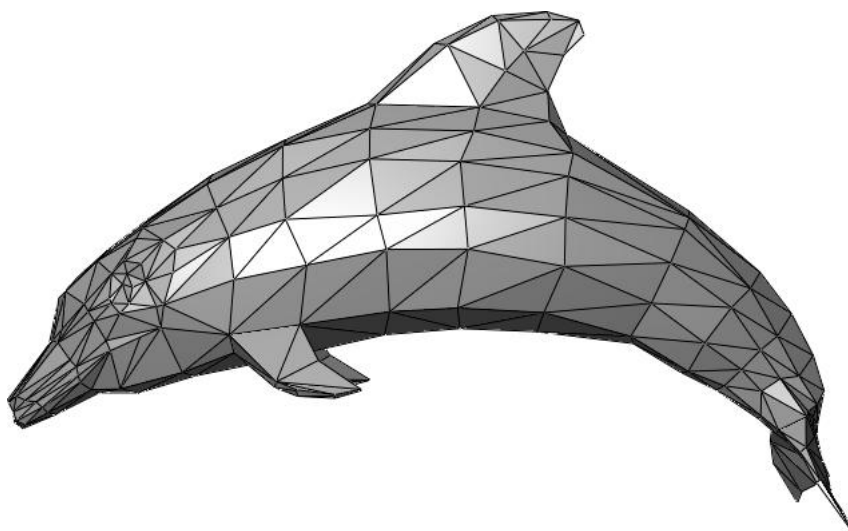


Figura 2.25: Mesh di triangoli

2.6 Ricostruzione tridimensionale con OpenCV

OpenCV [1] è una libreria open source dedicata alla *computer vision*. È utilizzabile tramite diversi linguaggi di programmazione ed è orientata all'elaborazione delle immagini e ad applicazioni in tempo reale. Uno degli scopi di OpenCV è quello di fornire funzioni utili allo sviluppo di software per controllo di qualità, controlli medici, controlli di sicurezza, oltre a calibrazione delle fotocamere/telecamere, visione stereoscopica, ricostruzione tridimensionale, ecc.

La procedura di calibrazione di OpenCV implementa l'algoritmo di Zhang ([4]) per il calcolo dei parametri intrinseci ed estrinseci, l'algoritmo di Brown ([21]) per il calcolo dei parametri di distorsione e il procedimento di Bouguet ([18]) per la rettificazione delle immagini. Inoltre mette a disposizione le funzioni necessarie ad effettuare una ricostruzione tridimensionale partendo da immagini stereoscopiche. Il procedimento è riassumibile nei passi:

1. estrazione dei corner;
2. calcolo dell'omografia;

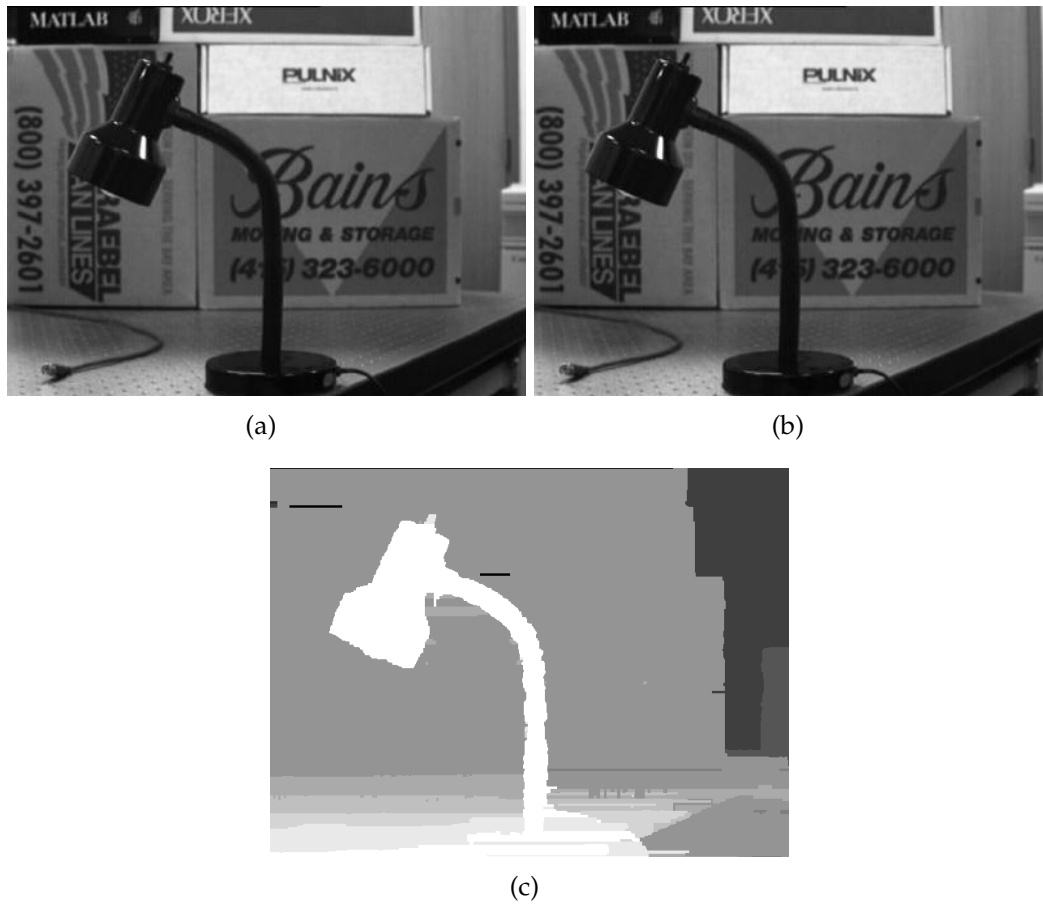


Figura 2.26: Esempio di calcolo di una disparity map partendo da immagini acquisite da un sistema stereoscopico: (a) immagine sinistra; (b) immagine destra; (c) disparity map: i pixel più chiari corrispondono a punti più vicini

3. calibrazione della fotocamera;
4. calcolo dei parametri di distorsione;
5. calibrazione stereoscopica;
6. correzione delle distorsioni;
7. rettificazione;
8. ricerca delle corrispondenze e costruzione della disparity map;

9. calcolo del modello tridimensionale.

2.6.1 Numero di immagini di calibrazione

Il numero di immagini necessarie alla calibrazione dipende dal numero di parametri del modello di calibrazione; teoricamente, la conoscenza della posizione di 3 corner della scacchiera è sufficiente per determinare i cinque parametri di distorsione del modello. L'algoritmo di OpenCV si basa su un numero maggiore di corner al fine di ottenere una maggiore robustezza del risultato.

Considerando N angoli per immagine e K immagini, le K immagini forniscono $2NK$ vincoli (ogni angolo ha una coordinata x e y). Da questi vincoli è necessario calcolare 4 parametri intrinseci e $6K$ parametri estrinseci. Quindi:

$$2NK \geq 6K + 4 \quad (2.20)$$

o, più semplicemente:

$$(N - 3)K \geq 2 \quad (2.21)$$

Ponendo $N = 5$, l'equazione 2.21 è verificata per una sola immagine ($K = 1$). Non è possibile, però, usare una sola immagine perché i parametri intrinseci ed estrinseci del sistema sono descritti da un'omografia, che è completamente descritta da 4 punti.

Un'omografia, cioè una relazione biunivoca tra i punti dello spazio tridimensionale ed i punti nell'immagine, è una mappatura dall'oggetto tridimensionale alla superficie dell'immagine ed esprime la trasformazione di un quadrato in un qualsiasi quadrilatero generico (Figura 2.27).

In ogni immagine è quindi possibile considerare un massimo di quattro punti. Sostituendo $N = 4$ nell'equazione 2.21 risulta $K \geq 2$. Il numero minimo di immagini è quindi 2, ognuna con 3×3 angoli.

Per ottenere un'accuratezza accettabile della procedura di calibrazione, comunque, in [1] viene consigliato di utilizzare un minimo di 10 immagini, raffiguranti almeno 7×8 corner di una scacchiera.

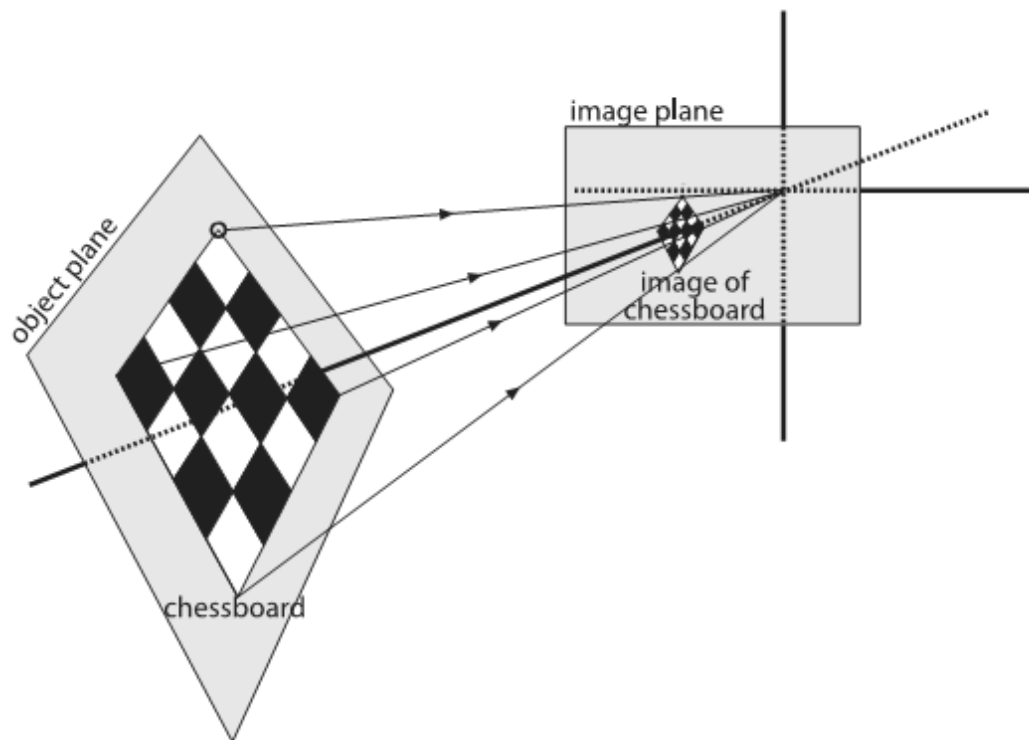


Figura 2.27: Mapping come descritto da un'omografia

2.6.2 Estrazione dei corner

La funzione di riconoscimento dei corner accetta in ingresso un'immagine in scala di grigi, il numero di corner della scacchiera ed il tipo di prefiltraggio da effettuare. L'algoritmo si compone di tre passi fondamentali:

1. prefiltraggio;
2. calcolo del gradiente dell'immagine;
3. estrazione dei corner;

I tipi di prefiltraggio possibile sono quattro:

- sogliatura in base a luminosità media;
- sogliatura adattativa;

- equalizzazione dell'istogramma;
- filtraggio dei quadrangoli.

In particolare, il filtraggio dei quadrangoli elimina dall'immagini i quadrati che non fanno parte della scacchiera. I filtri, inoltre, possono essere usati in modo combinato.

Nella fase di estrazione dei corner l'algoritmo utilizzato da OpenCV ricerca i corner in tutte le immagini disponibili e restituisce il numero di corner trovati in ogni immagine. In questo modo è possibile scartare le immagini in cui l'algoritmo non riesce a riconoscere tutti i corner.

La definizione di corner usata da OpenCV è quella fornita da Harris in [22], e comprende il calcolo della matrice Hessiana dell'immagine, cioè la matrice che contiene le derivate del secondo ordine dell'immagine I :

$$H(I) = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} , \quad (2.22)$$

Dalla matrice Hessiana vengono calcolate le autocorrelazioni delle derivate del secondo ordine dell'immagine, definite nell'intorno di ogni punto (x, y) :

$$\begin{aligned} & \sum_{-K \leq i, j \leq K} w_{i,j} \frac{\partial^2 I}{\partial x^2}(x+i, y+1) \quad ; \\ & \sum_{-K \leq i, j \leq K} w_{i,j} \frac{\partial I}{\partial x}(x+i, y+1) \frac{\partial I}{\partial y}(x+i, y+1) \quad ; \\ & \sum_{-K \leq i, j \leq K} w_{i,j} \frac{\partial I}{\partial x}(x+i, y+1) \frac{\partial I}{\partial y}(x+i, y+1) \quad ; \\ & \sum_{-K \leq i, j \leq K} w_{i,j} \frac{\partial^2 I}{\partial y^2}(x+i, y+1) \quad , \end{aligned} \quad (2.23)$$

dove K è la dimensione dell'intorno e $w_{i,j}$ è un termine di peso, che può essere uniforme o distribuito in modo Gaussiano. I corner sono definiti come i pixel dell'immagine per cui la matrice composta dalle autocorrelazioni 2.23 ha due autovalori maggiori di una soglia fissata.

2.6.3 Calcolo dell'omografia

Il procedimento di calibrazione sfrutta la matrice di omografia, che ha la proprietà di essere calcolabile conoscendo la posizione dei corner. Per la matrice di omografia H vale la seguente equazione:

$$q = sHQ \quad , \quad (2.24)$$

dove s è un fattore di scala e q e Q sono rispettivamente la posizione del punto sul piano bidimensionale e le coordinate del punto nello spazio tridimensionale, espresse in coordinate omogenee:

$$q = \begin{bmatrix} x & y & 1 \end{bmatrix}^T \quad ; \quad Q = \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T \quad . \quad (2.25)$$

La matrice di omografia H è composta da una matrice W , che descrive la rotazione e la traslazione dell'oggetto, e da una matrice di proiezione M , che descrive i parametri intrinseci della fotocamera:

$$H = WM \quad , \quad (2.26)$$

dove:

$$W = \begin{bmatrix} R & T \end{bmatrix} \quad ; \quad M = \begin{bmatrix} f_x & \alpha f_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad . \quad (2.27)$$

R e T sono rispettivamente la matrice di rotazione e il vettore di traslazione. L'equazione 2.24 può quindi essere riscritta come:

$$q = sMWQ \quad . \quad (2.28)$$

Se si considerano le coordinate del punto nello spazio tridimensionale riferite solamente al piano (scacchiera) di cui fa parte, la profondità del punto si può considerare uniforme e quindi si può assumere $Z = 0$. La terza colonna della matrice di rotazione R è combinazione lineare delle prime due e si può eliminare:

$$H' = sM \begin{bmatrix} r_1 & r_2 & T \end{bmatrix} , \quad (2.29)$$

dove r_1, r_2 sono le prime due colonne della matrice di rotazione e T è il vettore di traslazione. La relazione che mappa i punti nello spazio in punti nell'immagine diventa:

$$q = sH'Q' , \quad (2.30)$$

dove:

$$Q' = \begin{bmatrix} X & Y & 1 \end{bmatrix}^T . \quad (2.31)$$

Un'omografia è descritta dalla corrispondenza tra quattro punti dell'oggetto reale e quattro punti dell'immagine. Sostituendo nell'equazione 2.30 le coordinate di quattro punti dell'oggetto reale e di quattro corner estratti dalla scacchiera e risolvendo il sistema di otto equazioni per almeno due immagini, è possibile calcolare la matrice di omografia.

2.6.4 Calibrazione delle fotocamere

Il procedimento di calibrazione è inizialmente effettuato ipotizzando che non vi siano distorsioni. Dopo aver calcolato i parametri intrinseci ed estrinseci, si determinano i parametri di distorsione e si ripete il procedimento di calibrazione.

La matrice di omografia, precedentemente calcolata, può essere scomposta in tre vettori h_1, h_2, h_3 :

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = sM \begin{bmatrix} r_1 & r_2 & T \end{bmatrix} , \quad (2.32)$$

dove r_1, r_2 sono le prime due colonne della matrice di rotazione, s è il fattore di scala, M è la matrice intrinseca e T è il vettore di traslazione.

L'equazione 2.32 può essere scritta come un sistema di equazioni:

$$\begin{aligned} h_1 &= sMr_1 \\ h_2 &= sMr_2 \\ h_3 &= sMT \end{aligned} . \quad (2.33)$$

I vettori r_1, r_2 che compongono la matrice di rotazione R sono ortogonali per costruzione (rappresentano rotazioni lungo assi ortogonali) e di ampiezza uguale (il fattore di scala s è esterno). r_1 e r_2 sono quindi ortonormali (hanno la stessa norma e il loro prodotto scalare è uguale a 0):

$$\begin{aligned} r_1^T r_2 &= 0 \\ ||r_1|| &= ||r_2|| \end{aligned} \quad . \quad (2.34)$$

Risolvendo le equazioni 2.33 per r_1 e r_2 e sostituendo in 2.34 si ottiene:

$$\begin{aligned} h_1^T M^{-T} M^{-1} h_2 &= 0 \\ h_1^T M^{-T} M^{-1} h_1 &= h_2^T M^{-T} M^{-1} h_2 \end{aligned} \quad , \quad (2.35)$$

Per semplificare la scrittura delle equazioni, è possibile definire una matrice B nel modo seguente:

$$B = M^{-T} M^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} \quad . \quad (2.36)$$

La matrice B ha una soluzione in forma chiusa:

$$B = \begin{bmatrix} \frac{1}{f_x^2} & 0 & \frac{-c_x}{f_x^2} \\ 0 & \frac{1}{f_y^2} & \frac{-c_y}{f_y^2} \\ \frac{-c_x}{f_x^2} & \frac{-c_y}{f_y^2} & \frac{c_x^2}{f_x^2} + \frac{c_y^2}{f_y^2} + 1 \end{bmatrix} \quad , \quad (2.37)$$

dove c_x, c_y, f_x, f_y sono i parametri intrinseci della fotocamera.

Sostituendo $B = M^{-T} M^{-1}$ nelle equazioni 2.35 si ottiene:

$$\begin{aligned} h_1^T B h_2 &= 0 \\ h_1^T B h_1 &= h_2^T B h_2 \end{aligned} \quad . \quad (2.38)$$

Usando K immagini è possibile costruire un sistema di $2K$ equazioni con due equazioni 2.38 per ogni immagine e calcolare gli elementi della matrice B . I parametri intrinseci sono calcolati tramite le formule:

$$\begin{aligned} f_x &= \sqrt{\frac{\lambda}{B_{11}}} \\ f_y &= \sqrt{\frac{\lambda B_{11}}{B_{11}B_{22} - B_{12}^2}} \\ c_x &= -\frac{B_{13}f_x^2}{\lambda} \\ c_y &= \frac{B_{12}B_{13} - B_{11}B_{23}}{B_{11}B_{22} - B_{12}^2} , \end{aligned} \quad (2.39)$$

dove:

$$\lambda = B_{33} - \frac{(B_{33} + c_y(B_{12}B_{13} - B_{11}B_{23}))}{B_{11}} . \quad (2.40)$$

I parametri estrinseci sono ricavati tramite le formule:

$$\begin{aligned} r_1 &= \lambda M^{-1}h_1 \\ r_2 &= \lambda M^{-1}h_2 \\ r_3 &= r_1 \times r_2 \\ t &= \lambda M^{-1}h_3 , \end{aligned} \quad (2.41)$$

dove:

$$\lambda = \frac{1}{||M^{-1}h_1||} . \quad (2.42)$$

2.6.5 Calcolo dei parametri di distorsione

Dalla conoscenza dei parametri intrinseci è possibile calcolare la proiezione ideale di un punto tridimensionale su un piano bidimensionale, secondo il modello pinhole:

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = \begin{bmatrix} f_x \frac{X}{Z} + c_x \\ f_y \frac{Y}{Z} + c_y \end{bmatrix} , \quad (2.43)$$

dove (x_p, y_p) sono le coordinate ideali del punto sul piano bidimensionale. Combinando le equazioni 2.8 e 2.9 in unico sistema si ottiene:

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \begin{bmatrix} x_d \\ y_d \end{bmatrix} + \begin{bmatrix} 2p_1 x_d y_d + p_2 (r^2 + 2x_d^2) \\ p_1 (r^2 + 2y_d^2) + 2p_2 x_d y_d \end{bmatrix}, \quad (2.44)$$

dove (x_d, y_d) sono le coordinate del punto soggetto a distorsione e k_1, k_2, k_3, p_1, p_2 sono i parametri di distorsione. Risolvendo le equazioni 2.44 per un numero sufficientemente alto di punti si ottengono, con buona precisione, i parametri di distorsione del sistema. Successivamente, l'algoritmo di calibrazione ripete il procedimento di calibrazione considerando i parametri di distorsione. La fase di calibrazione e di calcolo dei parametri di distorsione sono ripetute fino alla convergenza dei valori o al raggiungimento di un numero limite di iterazioni.

Dopo aver calcolato i parametri di distorsione definitivi, si rimuovono gli effetti di distorsione introdotti dalle ottiche e si correggono le posizioni dei pixel tramite le equazioni 2.8 e 2.9.

2.6.6 Calibrazione stereoscopica

La matrice di rotazione R e il vettore di traslazione T di una fotocamera rispetto all'altra sono ricavate dalle matrici di rotazione e traslazione delle singole immagini, secondo le formule:

$$R = R_r(R_l)^T \quad ; \quad T = T_r - RT_l, \quad (2.45)$$

dove R_l, R_r, T_l e T_r sono le matrici di rotazione e i vettori di traslazione dell'immagine sinistra e dell'immagine destra.

L'algoritmo usato dalla libreria OpenCV effettua questo calcolo su tutte le immagini disponibili e determina i valori medi di R e T . Viene successivamente utilizzato l'algoritmo di Levenberg-Marquardt per raffinare i valori di R e T fino a trovare il minimo locale dell'errore di riproiezione.

L'errore di riproiezione è definito come la somma totale delle distanze al quadrato tra la posizione dei corner rilevati e la posizione dei corner calcolata in base alla stima corrente dei valori di R e T .

Dopo la fase di calcolo dei parametri di rettificazione (vedi 2.6.7), il sistema rettificato è calibrato nuovamente, ottenendo le matrici intrinseche del sistema rettificato M_{rectL} e M_{rectR} .

2.6.7 Rettificazione

L'algoritmo utilizzato da OpenCV per la rettificazione delle immagini minimizza il cambiamento da applicare alle immagini affinché siano equivalenti ad immagini acquisite da un sistema stereoscopico con telecamere allineate e coplanari, mentre massimizza l'area in comune tra le due immagini.

La rotazione definita dalla matrice R è divisa in due rotazioni r_l e r_r ed alle due telecamere è applicata metà della rotazione. In questo modo le telecamere sono coplanari (i raggi principali sono paralleli) ma non allineate.

La matrice di rotazione che allinea le due telecamere è costituita da 3 vettori e_1 , e_2 , e_3 . Il vettore e_1 è dato dal vettore di traslazione tra le due telecamere, normalizzato:

$$e_1 = \frac{T}{\|T\|} \quad . \quad (2.46)$$

Il vettore e_2 , ortogonale a e_1 , si calcola tramite il prodotto vettoriale tra e_1 e la direzione del raggio principale:

$$e_2 = \frac{[-T_y \quad T_x \quad 0]^T}{\sqrt{T_x^2 + T_y^2}} \quad , \quad (2.47)$$

dove il termine al denominatore normalizza il vettore.

Il vettore e_3 , ortogonale ai vettori e_1 e e_2 , è calcolato tramite il prodotto vettoriale di e_1 e e_2 :

$$e_3 = e_1 \times e_2 \quad . \quad (2.48)$$

La matrice di allineamento delle telecamere è definita come:

$$R_{\text{rect}} = \begin{bmatrix} e_1^T \\ e_2^T \\ e_3^T \end{bmatrix} . \quad (2.49)$$

Le matrici utilizzate per la correzione della rotazione e l'allineamento delle due telecamere sono calcolate tramite le formule:

$$R_l = R_{\text{rect}} r_l \quad ; \quad R_r = R_{\text{rect}} r_r . \quad (2.50)$$

Per modificare le immagini stereoscopiche in modo che siano compatibili con il sistema rettificato, è necessario calcolare, dalle posizioni dei pixel delle immagini di partenza, le posizioni dei corrispondenti pixel nelle immagini rettificate (*mapping*). Il mapping diretto, cioè il calcolo delle nuove posizioni partendo da quelle originali, non è realizzabile in quanto le posizioni finali sarebbero numeri reali e non corrisponderebbero ad ogni pixel dell'immagine di destinazione. Si effettua invece un mapping inverso: per ogni coordinata intera del pixel nell'immagine di destinazione, si cerca da quale coordinata reale nell'immagine di partenza è stata generata. Viene quindi effettuata un'interpolazione bilineare dei pixel circostanti per ottenere dei valori interi per le coordinate di destinazione.

Infine, l'algoritmo calcola nuovi centri delle immagini e nuovi bordi in modo da massimizzare l'area in comune tra le due immagini.

2.6.8 Ricerca delle corrispondenze e costruzione della disparity map

L'algoritmo di ricerca delle corrispondenze utilizzato da OpenCV è una variante dell'algoritmo basato su Block Matching proposto da Konolige in [23].

L'algoritmo necessita di immagini rettificate e corrette da distorsione e si compone di tre passi principali:

1. normalizzazione della luminosità e miglioramento del contrasto;
2. ricerca delle corrispondenze lungo le linee epipolari;
3. postprocessing per eliminare le corrispondenze errate.

Le corrispondenze sono calcolate con una finestra SAD (*Sum of Absolute Difference*), che calcola la somma dei valori assoluti delle differenze dei valori di intensità dei pixel in aree locali delle due immagini. Per ogni feature nell'immagine sinistra, viene ricercata la feature corrispondente nell'immagine destra. Poichè le immagini sono rettificate, il match deve avvenire sulla stessa riga: se la prima feature ha coordinate (x_0, y_0) nell'immagini sinistra, deve avere coordinata (x_1, y_0) sull'immagine destra, con x_1 uguale a x_0 o alla sinistra di x_0 . In Figura 2.28 è mostrato come la feature nell'immagine sinistra alla posizione (x_0, y_0) sia cercata nell'immagine destra, a partire dalla coordinata associata alla disparità minima fino ad arrivare alla coordinata associata alla disparità massima. Il numero di pixel valutato in questo intervallo è detto numero di disparità. La funzione di matching in questo intervallo ha generalmente un punto di massimo globale circondato da punti di massimo locali di altezza inferiore.

Esiste un vincolo di ordinamento delle feature rilevate nelle due immagini: l'ordine di queste feature deve rimanere invariato tra l'immagine sinistra e l'immagine destra. È possibile che alcune feature dell'immagine sinistra non abbiano corrispondenze nell'immagine destra (occlusioni) oppure che nell'immagine destra siano presenti feature non associate ad alcun punto nell'immagine sinistra (inserzioni). Il vincolo di ordinamento deve quindi essere applicato alle feature che presentano una corrispondenza biunivoca.

Viene, inoltre, effettuata una fase di postprocessing al fine di eliminare corrispondenze non valide. L'algoritmo di ricerca delle corrispondenze può fornire infatti risultati errati in zone dell'immagine vicine ai bordi, poichè la finestra di matching estrae l'oggetto (*foreground*) in un'immagine e lo sfondo (*background*) nell'altra. Questo produce una funzione di match-

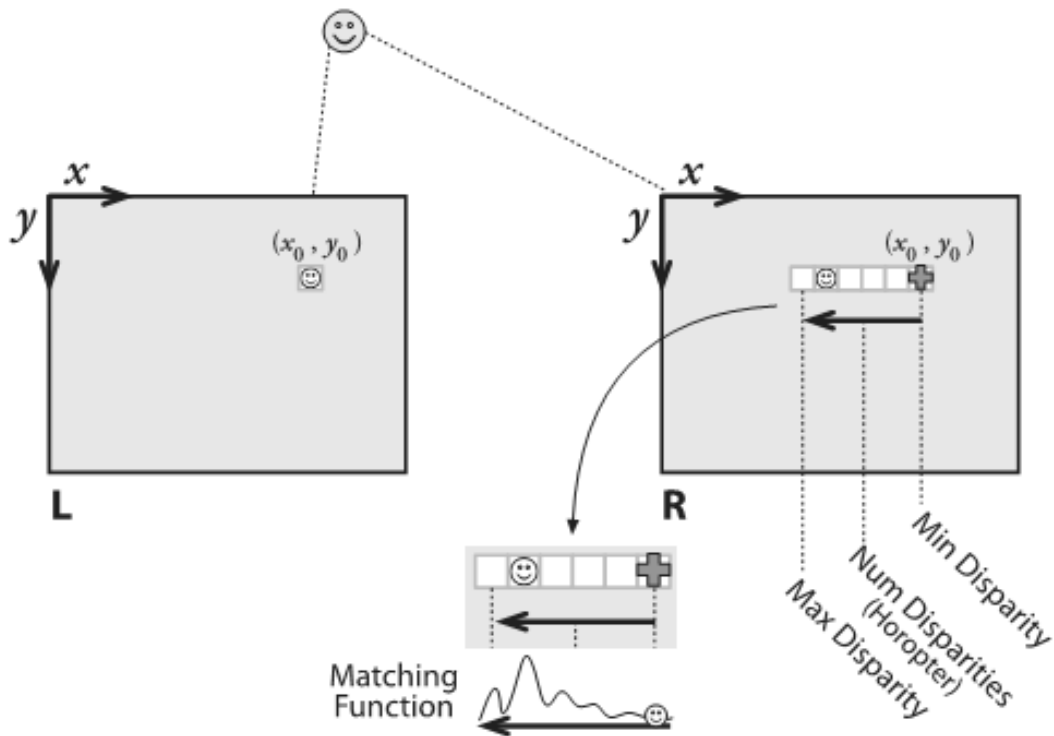


Figura 2.28: Ricerca delle corrispondenze

ing rumorosa, con valori di disparità molto grandi e molto piccoli (*speckle*). Per evitare questo problema l'algoritmo usa un rilevatore di speckle locale: se la disparità massima e quella minima sono nel range accettato dal rilevatore di speckle, la corrispondenza viene considerata valida.

La disparity map viene quindi ottenuta calcolando la distanza tra le feature nelle due immagini.

2.6.9 Costruzione del modello tridimensionale

Tramite la matrice intrinseca M_{rectL} della telecamera rettificata ed il vettore di traslazione T è possibile costruire la matrice di riproiezione (*reprojection*

matrix):

$$Q = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 0 & f \\ 0 & 0 & -\frac{1}{T_x} & 0 \end{bmatrix} , \quad (2.51)$$

in cui i parametri c_x , c_y e f sono le coordinate del punto principale e la lunghezza focale delle telecamere nel sistema stereoscopico rettificato e T_x rappresenta la coordinata x del vettore di traslazione. Dato un punto della disparity map con coordinate (x, y) e valore di disparità d , la posizione del punto tridimensionale ricostruito (in coordinate omogenee) è data da:

$$Q \begin{bmatrix} x \\ y \\ d \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} , \quad (2.52)$$

da cui è possibile ricavare le coordinate effettive nello spazio tridimensionale come: $(X/W, Y/W, Z/W)$.

IMPRONTE DIGITALI ACQUISITE SENZA CONTATTO

In questo capitolo vengono trattate le tecnologie per il riconoscimento biometrico tramite impronte digitali che non prevedono il contatto del dito umano con un sensore. Per prima cosa sono introdotte le differenze tra questo approccio innovativo e quello classico basato sul contatto. Successivamente, vengono presentati esempi di sistemi di riconoscimento basati su scansione senza contatto dell'impronta digitale. Infine, sono introdotte alcune tecnologie in grado di ottenere una rappresentazione tridimensionale del dato biometrico e le loro compatibilità con i sistemi attualmente in uso.

3.1 Impronte digitali a contatto e senza contatto

Le tecniche di acquisizione di impronte digitali senza contatto consentono di acquisire immagini del dermatoglypho del dito senza che il dito stesso venga a contatto con alcun sensore.

I motivi che hanno spinto la ricerca verso questo nuovo tipo di scansione dell'impronta digitale sono molteplici:

- l'identificazione basata su contatto richiede la pressione del dito sul sensore e comporta deformazioni dell'acquisizione (come l'unione delle creste o la non visibilità delle creste esistenti), che rendono ogni scansione diversa dalla precedente;
- il mapping tra la tridimensionalità del dito e la superficie bidimensionale del sensore può modificare la struttura delle minuzie nell'immagine acquisita, soprattutto se il dito viene mosso durante il contatto con la superficie di acquisizione;
- le condizioni ambientali (umidità, secchezza) e del dito (sudore, sporco) influenzano il procedimento, come mostrato in Figura 3.1;
- dopo ogni scansione rimangono tracce sul sensore (impronte latenti) che possono deteriorare la qualità delle scansioni successive o essere utilizzate al fine di ricostruire modelli dell'impronta digitale ([24]);
- il contatto con un sensore potrebbe non essere gradito dagli utenti a causa del timore di contagi oppure per retaggio culturale;
- lo spessore delle creste del dito può cambiare durante gli anni, a causa dell'usura delle dita. Sebbene questo rappresenti un problema anche nel caso di sistemi di riconoscimento basati su immagini di impronte digitali acquisite senza contatto, lo è maggiormente con le tecniche classiche;
- non è necessario impiegare sensori dedicati ma è possibile usare fotocamere o telecamere;

A questi benefici bisogna aggiungere anche alcuni svantaggi:

- sono necessarie nuove tecniche di controllo della qualità dell'immagine;
- le tecniche anti frode classiche, ad esempio i controlli della vitalità del dito, non sono utilizzabili;

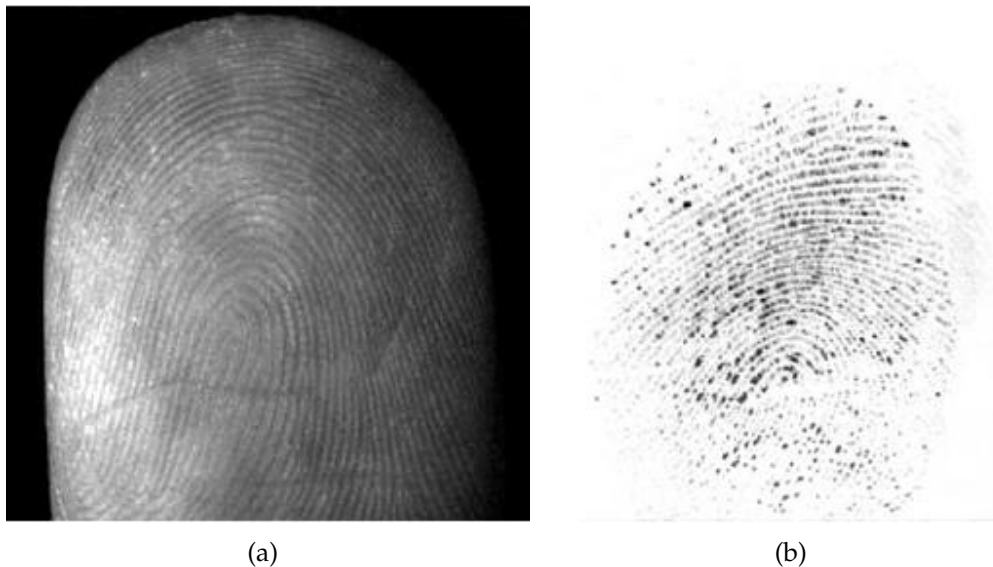


Figura 3.1: Confronto tra fingerprint in caso di dito secco: (a) fingerprint acquisita senza contatto; (b) fingerprint acquisita a contatto

- risulta necessario assicurare la compatibilità i database biometrici esistenti, specialmente con i sistemi adottati dalle forze dell'ordine;

Le immagini prodotte con dispositivi senza contatto, come mostrato in Figura 3.2, sono molto diverse da quelle ottenibili con sensori dedicati e, in prima analisi, non utilizzabili per il riconoscimento biometrico. Attualmente, i sistemi di riconoscimento biometrico basati su immagini di impronte digitali acquisite senza contatto si basano su tecniche di miglioramento delle immagini allo scopo di poter elaborare immagini molto simili a quelle acquisite tramite sensori tradizionali

3.2 Sistemi basati su immagini bidimensionali

La struttura generale di un sistema biometrico basato su impronte digitali a contatto è mostrata in Figura 3.3. Questo tipo di sistemi usa un sensore a contatto, generalmente ottico o capacitivo, come modulo di acqui-

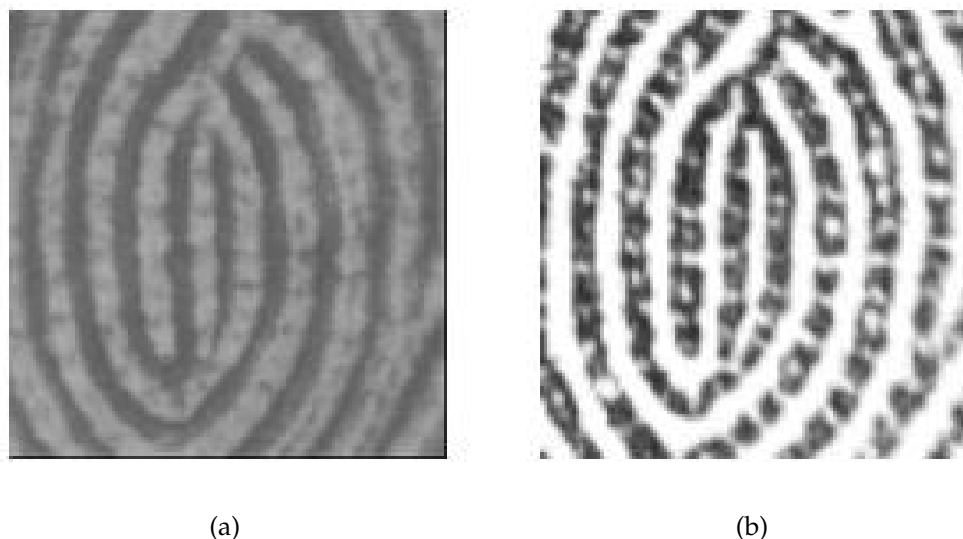


Figura 3.2: Immagini della medesima impronta: (a) immagine acquisita a distanza (b) Immagine acquisita tramite sensore a contatto

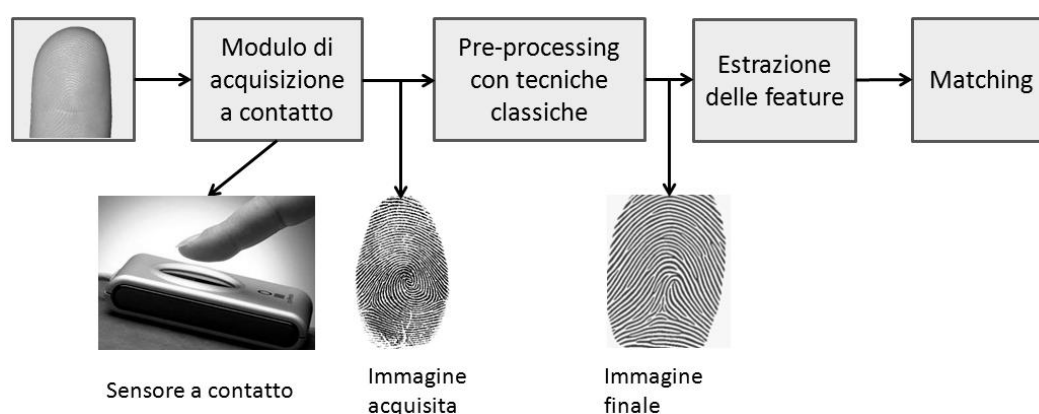


Figura 3.3: Struttura generale di un sistema biometrico basato su impronte digitali acquisite tramite contatto

sizione e tecniche di miglioramento dell'immagine adatte all'elaborazione di immagini ad alto contrasto.

La struttura dei sistemi di riconoscimento biometrico dell'impronta digitale che non prevedono il contatto del dito con un sensore è mostrata in Figura 3.4. Essi hanno in comune con le tecniche classiche i moduli di es-

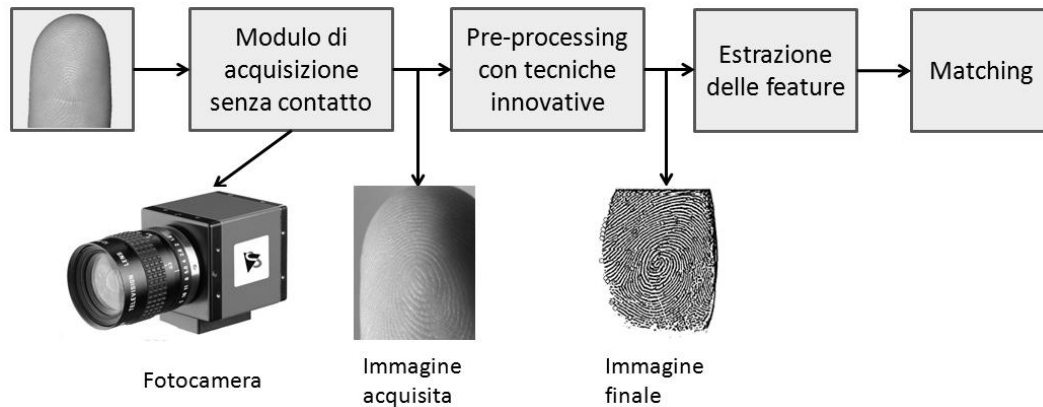


Figura 3.4: *Struttura generale di un sistema biometrico basato su impronte digitali acquisite senza contatto*

trazione delle feature e gli algoritmi di matching. Usano, però, fotocamere per acquisire l'immagine del dito e sfruttano tecniche di miglioramento delle immagini studiate appositamente per immagini a basso contrasto.

3.2.1 Sistemi di acquisizione

Secondo [25], gli attuali sistemi di riconoscimento delle impronte digitali si dividono in due grandi categorie: quelli basati sulla riflessione (i più usati) e quelli basati sulla trasmittanza.

I metodi basati su riflessione (schematizzati in Figura 3.5) illuminano il dito e ne acquisiscono la luce riflessa. La posizione delle fonti di luce è determinante: per ottenere una buona immagine, l'impronta deve essere illuminata in modo che la maggior parte della luce sia riflessa al rilevatore.

Bisogna fare in modo che le creste e le valli assorbano una diversa quantità di luce: il contrasto tra creste e valli, infatti, deve essere sufficientemente elevato per l'estrazione delle features. Sono importanti a questo scopo:

- **Depth of Focus:** il range di profondità in cui il dispositivo di acquisizione è in grado di mettere a fuoco l'oggetto;

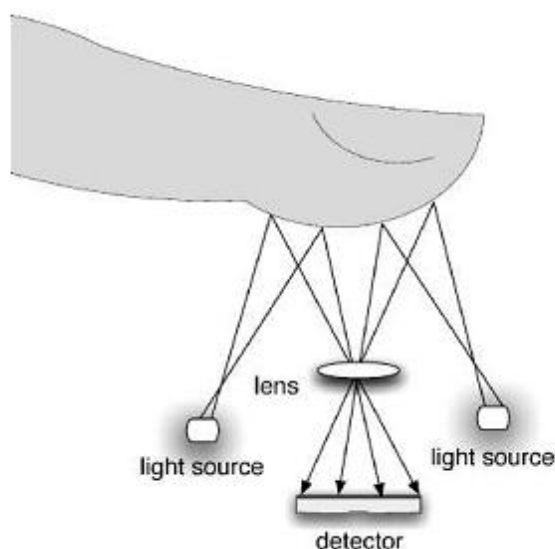


Figura 3.5: Sistemi senza contatto basati su riflettanza

- **Field of view:** il campo visivo del dispositivo;
- **Frequenza della luce:** alcuni colori sono più adatti per illuminare il dito;
- **Velocità dell'otturatore:** più l'otturatore è veloce e meno probabilmente ci saranno sfuocature dovute al movimento;
- **Posizione delle fonti di luce:** il dito deve essere illuminato completamente e non dovrebbero esserci ombre;

Le fonti di luce devono essere posizionate il più possibile vicine al rilevatore e perpendicolari al dito, in modo che riflettano la maggior parte della luce e si evitino creazioni di ombre e spostamento della posizione appartenente delle creste, come mostrato in Figura 3.6.

Anche la lunghezza d'onda della luce è importante: la situazione ottimale si ha nel caso in cui la luce venga riflessa dallo strato superficiale e non penetri negli strati interni. Esperimenti hanno dimostrato che l'emoglobina (vedi Figura 3.7) assorbe maggiormente la luce tra 545 nm e

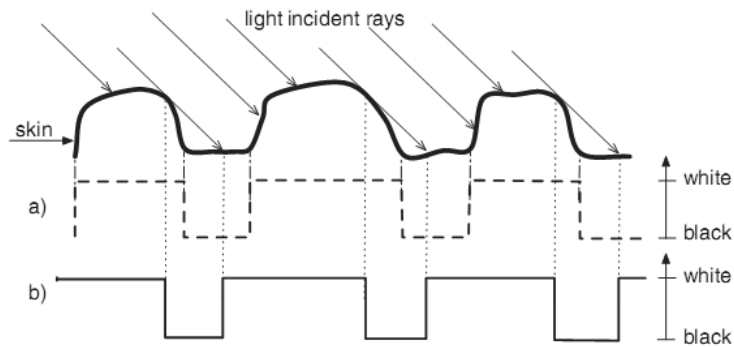


Figura 3.6: *Spostamento delle creste in caso di luce non perpendicolare*

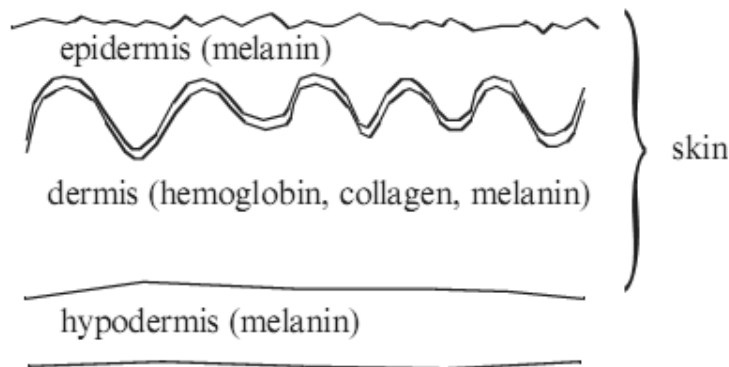


Figura 3.7: *Rappresentazione semplificata degli strati della pelle*

575 nm mentre assorbe in modo minore la luce con lunghezza d'onda intorno ai 500 nm. Inoltre, i sensori CCD sono maggiormente sensibili ai 500 nm. È stato quindi appurato che luce migliore per questo scenario è quella blu. Inoltre, il colore blu è complementare a giallo e rosso, i colori predominanti nella pelle del dito.

L'effetto prospettico, dovuto alla curvatura del dito ed al fatto che esso viene posizionato ad una certa distanza dal CCD, fa in modo che la distanza di ogni regione del dito dal sensore non sia costante. Per questo motivo, il depth of focus deve essere sufficientemente ampio, in modo da mettere a fuoco tutto il dito. La curvatura del dito, inoltre, fa in modo che le creste vicino al bordo del dito tendano ad unirsi nell'immagine acquisita. Questo effetto comporta un'aumento della frequenza apparente ed una

diminuzione della risoluzione utile vicino al bordo.

Il dito, nonostante possa venire appoggiato su una guida, non è perfettamente fermo. Per questo motivo tutti i pixel devono essere acquisiti nello stesso momento ed il tempo di esposizione deve essere molto breve.

Infine, per essere compatibile con gli standard richiesti dai sistemi di identificazione biometrica, la risoluzione deve essere di almeno 500 ppi (*Pixels per inch*) sia lungo l'asse x che lungo l'asse y , e costante in tutta l'immagine ([26]). La risoluzione ottica però, a causa della curvatura, decresce dal centro verso i bordi.

Le tecniche basate su trasmittanza, invece, sfruttano il fatto che esiste uno strato interno della pelle umana che ha caratteristiche che variano in base alla presenza o meno sulla superficie di creste o valli. Più precisamente, la trasmittanza di questo strato, definita come la frazione di luce che lo attraversa, varia in modo analogo all'impronta digitale, senza essere influenzata dai difetti di superficie e soprattutto osservabile anche in caso di impronte danneggiate. La Mitsubishi ha proposto in [27] un setup per sfruttare questa tecnica, illustrato in Figura 3.8. Questa tecnologia usa una luce rossa a 660 nm, che ha la maggiore trasmittanza nella pelle, puntata sull'unghia. La luce che ha penetrato il dito è infine raccolta dal rilevatore dall'altra parte, generando un'immagine come in Figura 3.9. Aumentare la lunghezza d'onda in questo scenario produce assorbimento da parte dell'emoglobina del sangue, il che risulterebbe in un'immagine disturbata dai capillari. Diminuirla comporterebbe l'assorbimento da parte dei tessuti e la conseguente impossibilità di riconoscere i pattern caratteristici.

3.2.2 Preprocessing di impronte digitali senza contatto

Nel processo di riconoscimento biometrico basato su immagini di impronte digitali senza contatto, la fase di preprocessing è probabilmente la più importante. Essa, infatti, consente di sfruttare un sensore di acquisizione senza contatto ed interfacciarlo in modo efficace con un database di impronte digitali esistente.

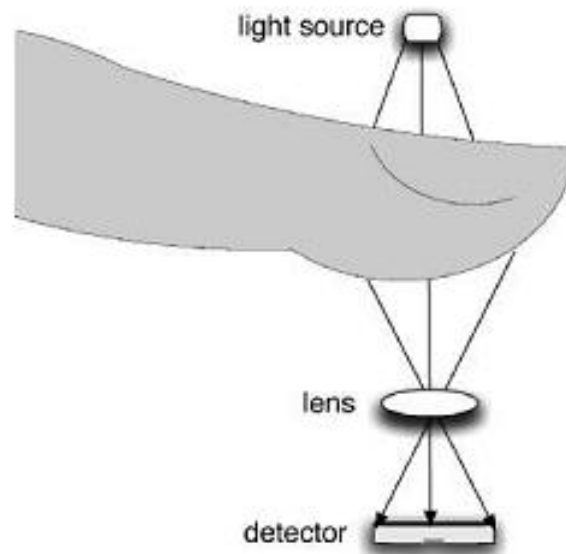


Figura 3.8: *Sistemi basati su trasmittanza*

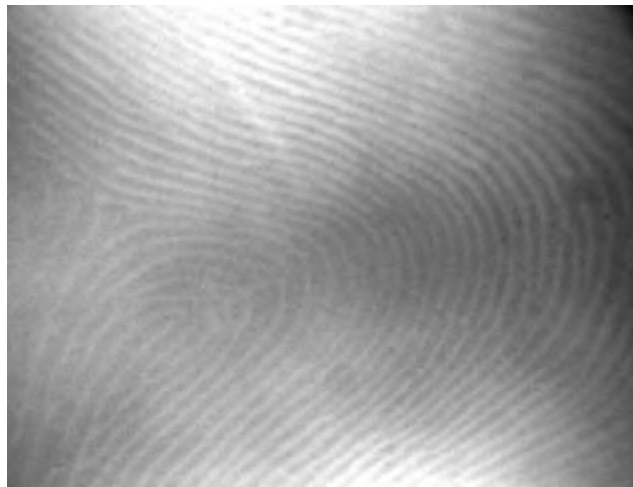


Figura 3.9: *Immagine prodotta da un sistema basato su trasmittanza*

Il processo di elaborazione di un'immagine, come proposto da [28] e [29], è schematizzato in Figura 3.10. Si possono osservare quattro passi principali:

- **Enhancement:** per prima cosa, bisogna compensare gli effetti di sfuocatura introdotti dalle imperfezioni della messa a fuoco e dal movi-

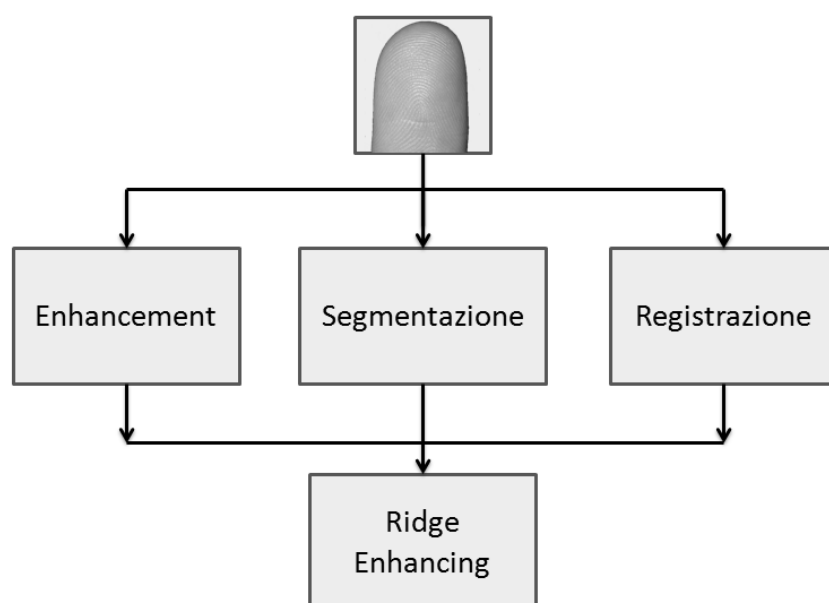


Figura 3.10: *Processo di elaborazione di un impronta digitale acquisita senza contatto*

mento del dito e le imperfezioni dovute a un'illuminazione non uniforme;

- **Segmentazione:** questo step determina la regione di interesse (*Region of Interest*, o ROI) dell'immagine del dito, cioè l'area dell'immagine che sarà utilizzata per le successive elaborazioni.
- **Registrazione:** questo passo è necessario per trasformare qualsiasi immagine in un'immagine con caratteristiche standard. Un'immagine standard presenta il dito in posizione centrale, con orientamento orizzontale e in cui la parte dell'immagine occupata dal dito è una frazione precisa dell'area totale di acquisizione.
- **Ridge enhancing:** l'ultima fase produce immagini direttamente confrontabili con quelle ottenute dai sensori a contatto.

3.2.2.1 Preprocessing basato su filtro passa-banda

Il metodo di preprocessing delle immagini descritto in [28] è schematizzato in Figura 3.11. Si compone dei seguenti passi:

1. **Enhancement:** la sfuocatura presente nell'immagine è compensata con l'algoritmo di Lucy-Richardson e con il filtro di Wiener. La *Point Spread Function* (PSD), necessaria per l'uso del filtro di Wiener, è stimata in base ad un filtro Gaussiano simmetrico.
2. **Segmentazione:** l'immagine viene convertita dallo spazio RGB allo spazio YUV ed ogni canale dello spazio YUV viene modificato con un'operazione di sogliatura. Le soglie sono stabilite in modo da conservare solo le tonalità di colore associate alla pelle del dito. Successivamente, viene effettuata una binarizzazione dell'area utile tramite un'analisi dei componenti connessi ed un'erosione volta ad eliminare i bordi del dito dove, tipicamente, la qualità dell'immagine non è sufficiente. L'operazione di segmentazione è resa più robusta tramite un'analisi della frequenza spaziale dell'immagine, che permette di riconoscere le aree contenenti componenti ad alta frequenza (creste).
3. **Registrazione:** per ottenere un'immagine con le caratteristiche standard di orientamento e posizione del dito, viene sfruttata la maschera binaria ottenuta dalla segmentazione. Da essa è possibile ricavare la posizione del dito e ruotare l'immagine in modo che il dito sia disposto orizzontalmente. Successivamente, l'immagine viene ridimensionata in modo che la larghezza del dito abbia una dimensione fissa.
4. **Ridge enhancing:** viene estratto lo sfondo usando un operatore di Unsharp masking, successivamente viene applicato un filtro passa banda per conservare le sole frequenze spaziali associate alle creste. Le frequenze di taglio del filtro passa banda sono stimate con un approccio basato su un filtro di Gabor.

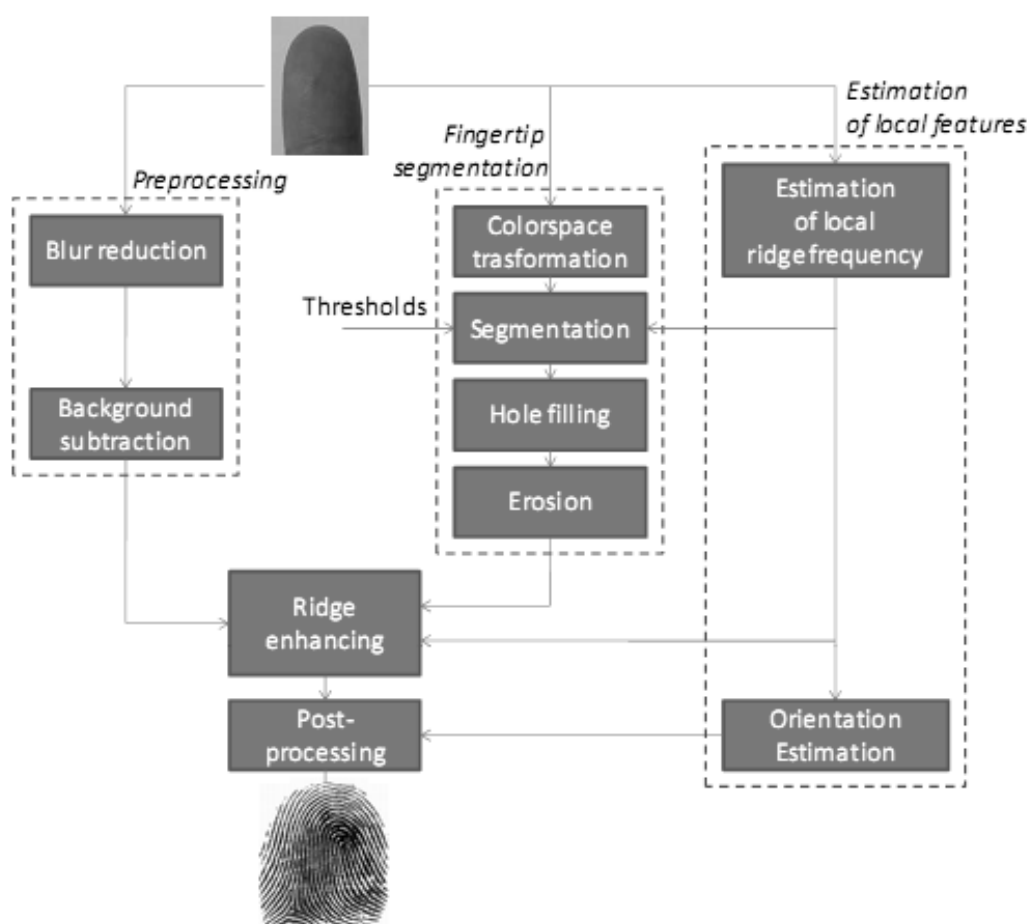


Figura 3.11: Processo di elaborazione basato su filtro passa banda

3.2.2.2 Preprocessing basato su trasformata di Fourier

Il metodo descritto in [29] è schematizzato in Figura 3.12. È suddivisibile nei seguenti passi:

1. **Enhancement:** l'immagine a colori di partenza viene convertita dallo spazio RGB ad un'immagine in scala di grigi. Successivamente viene normalizzata modificando il range dinamico dei valori di intensità dei pixel. In questo modo si compensano le variazioni nell'immagine dovute ad illuminazione non uniforme.

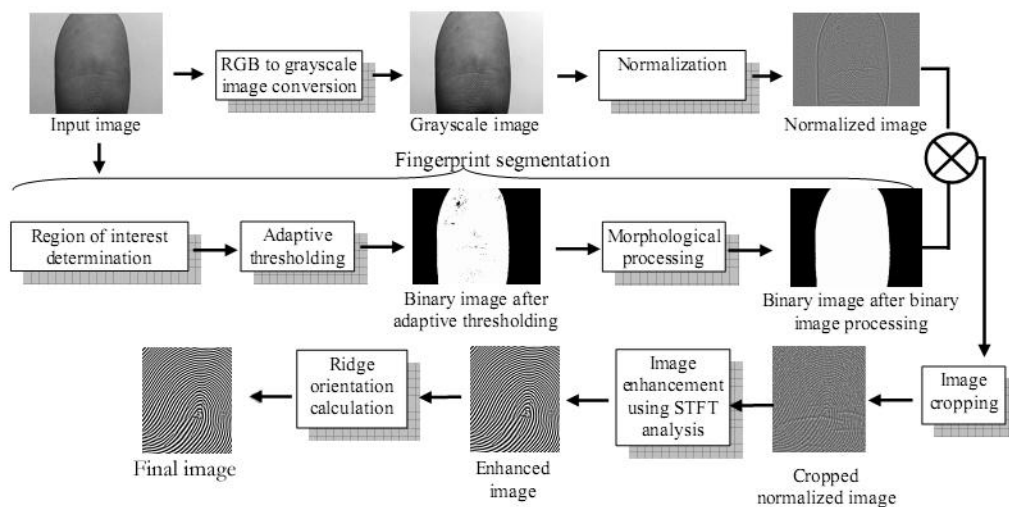


Figura 3.12: Processo di elaborazione basato su trasformata di Fourier

2. **Segmentazione:** l'immagine originale è convertita dallo spazio RGB allo spazio YCbCr. Viene collezionato un ampio numero di campioni di immagini del dito ottenuti da diversi soggetti e vengono calcolate la media e la covarianza delle immagini. Tramite questi valori viene determinata la ROI e quindi l'immagine viene binarizzata tramite un metodo di soglia adattativa. La maschera binaria ottenuta è trattata tramite operatori morfologici.
3. **Ridge enhancing:** l'immagine viene divisa in un prefissato numero di sottoimmagini. Viene calcolata la *Short Term Fourier Transform* (STFT) di ogni immagine, successivamente viene applicato un filtraggio di Gabor con frequenza e angolazione pari a quelle stimate per le creste.

3.2.3 Prestazioni e vulnerabilità dei sistemi senza contatto

Per valutare l'efficacia dei sistemi di riconoscimento senza contatto è necessario utilizzare i software esistenti per l'estrazione ed il matching delle minuzie.

In [28] è stato usato il software NBIS mindtct per l'estrazione delle minuzie da un'immagine acquisita senza contatto e sottoposta ad un'operazione di prefiltraggio. Come mostrato in [28] la maggior parte delle minuzie vengono estratte correttamente, anche se le operazioni di filtraggio a cui è stata sottoposta l'immagine hanno introdotto artefatti, rilevati in modo erroneo come minuzie.

Il confronto tra il sistema di riconoscimento senza contatto e i sistemi classici è stato effettuato creando due database di impronte digitali, ognuno contenente le stesse impronte digitali, rispettivamente acquisite da una fotocamera e da un sensore a contatto allo stato dell'arte.

Il software di matching NBIS bozorth3 è stato successivamente applicato su entrambi i database. I risultati mostrano come la precisione del riconoscimento delle impronte digitali effettuato da sistemi senza contatto sia paragonabile, anche se minore, ai sistemi classici.

Come quelli classici, i sistemi di riconoscimento basati su immagini di dita acquisite senza contatto sono vulnerabili ad alcuni attacchi. In particolare, potrebbero essere ingannati utilizzando dita finte, infatti potrebbe essere sufficiente una fotografia della struttura del dermatoglifo per ingannare il sensore. Per risolvere questo problema in [25] si propone un approccio che fa uso di fotocamere ad alta definizione per rilevare le micro sudorazioni della pelle.

3.3 Sistemi basati su modelli tridimensionali

I sistemi di riconoscimento biometrico basati su immagini di impronte digitali acquisite senza contatto, a causa della curvatura del dito, possono utilizzare una minore area utile del dito rispetto ai sistemi tradizionali. I sensori classici, infatti, richiedono la pressione del dito sul sensore. Ciò comporta una deformazione del dito che lo appiattisce parzialmente, con il risultato che anche le zone periferiche del dito entrano in contatto con la superficie del sensore.

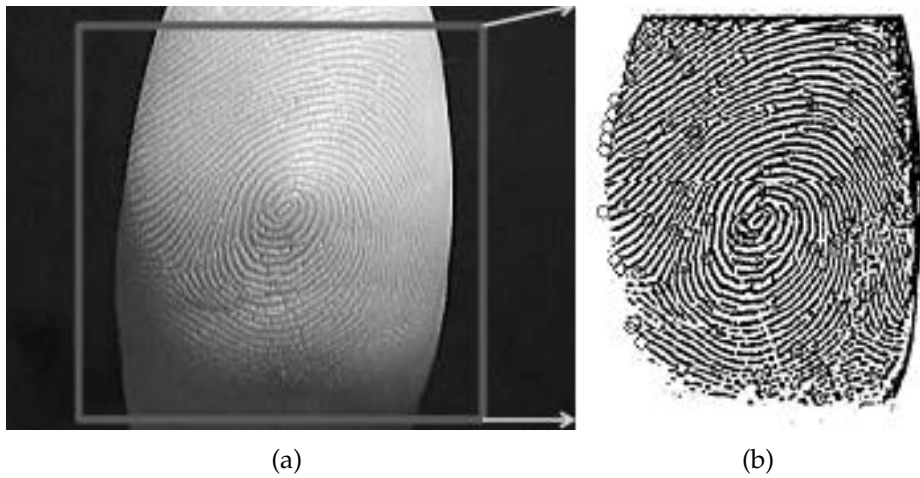


Figura 3.13: Esempio di estrazione delle minuzie da immagine acquisita a distanza (a) Immagine originale (b) Minuzie rilevate e sovrainposte all'immagine elaborata

Una possibile soluzione a questo problema è rappresentato dalla creazione di un modello tridimensionale del dito. Tramite la costruzione di una rappresentazione tridimensionale completa del dito, infatti, è possibile utilizzare l'intera superficie dell'impronta digitale per il riconoscimento biometrico. In questo modo si può disporre di un'area utile persino superiore a quella ottenibile utilizzando un sensore a contatto.

I sistemi di riconoscimento dell'impronta digitale basati sulla costruzione del modello tridimensionale sono schematizzati in Figura 3.14. Si dividono in due tipologie principali:

- sistemi basati su viste multiple;
- sistemi basati su luce strutturata.

Entrambi i sistemi, successivamente alla creazione del modello tridimensionale, richiedono una fase di *unwrapping* per essere compatibili con i database di immagini esistenti.

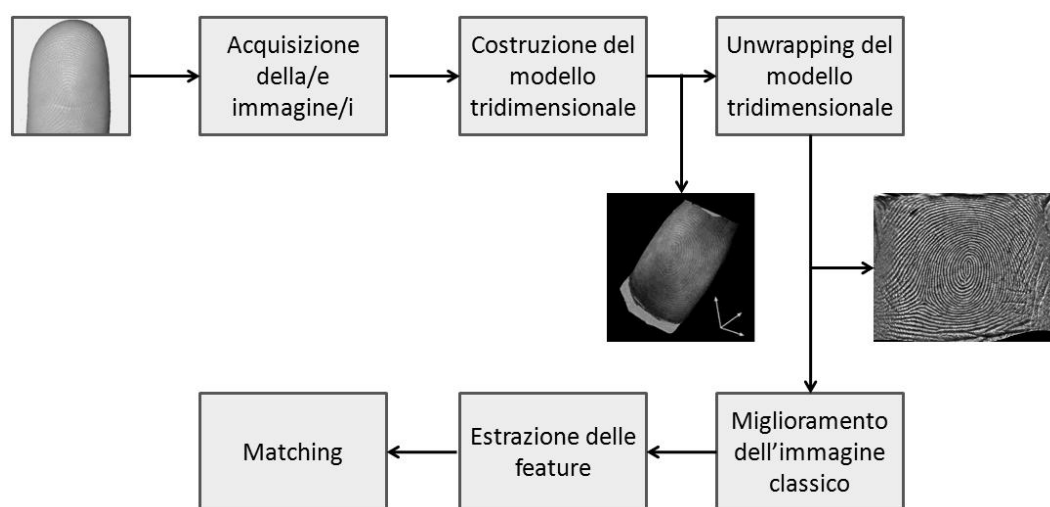


Figura 3.14: *Struttura di un sistema di riconoscimento basato su modello tridimensionale dell'impronta digitale*

3.3.1 Unwrapping del modello tridimensionale

Al fine di rendere compatibile i sistemi basati su modelli tridimensionali dell'impronta digitale con i database di immagini esistenti, in [30] è descritto un metodo per effettuare l'*unwrapping* dell'impronta tridimensionale su un piano bidimensionale. Tramite l'utilizzo di questo metodo le superfici tridimensionali sono mappate su immagini bidimensionali. Diventa quindi possibile utilizzare i normali algoritmi di riconoscimento e matching delle minuzie.

In letteratura, le principali tecniche di unwrapping applicabili ai modelli tridimensionali di impronte digitali sono:

1. **Unwrapping parametrico:** l'oggetto viene proiettato su un solido geometrico (come un cilindro o un cono). La conoscenza del modello geometrico dell'oggetto permette a sua volta la mappatura su un piano bidimensionale. Le tecniche di unwrapping parametriche sono relativamente semplici ed offrono buoni risultati, a patto che il modello geometrico usato sia simile all'oggetto di partenza;
2. **Unwrapping non parametrico:** l'oggetto viene direttamente proiet-

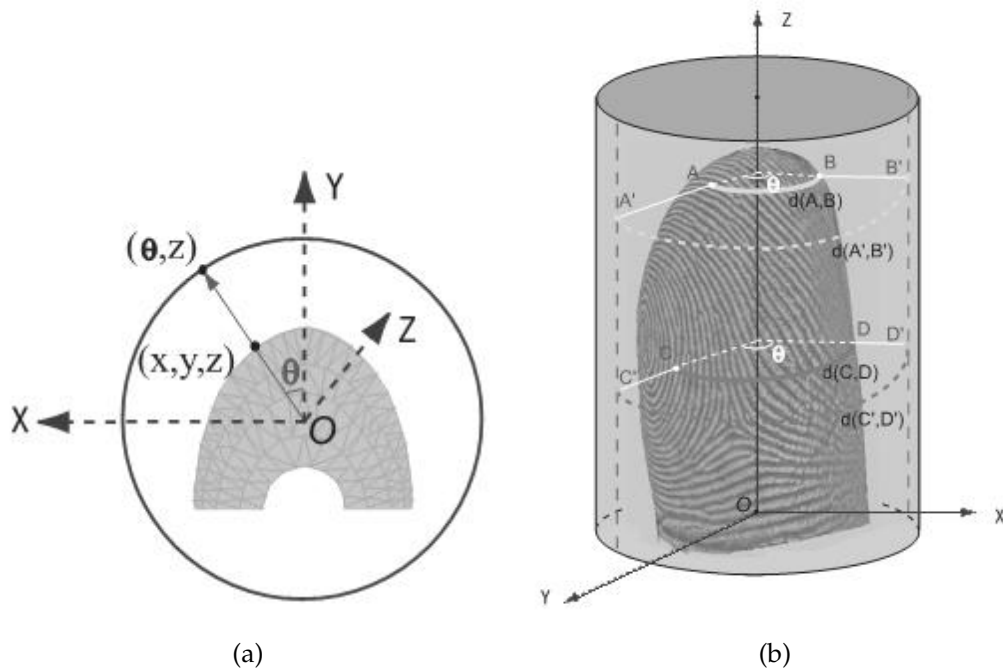


Figura 3.15: *Unwrapping parametrico usando un cono: (a) il punto (x, y, z) è mappato sul punto (θ, z) ; (b) la distanza tra i punti sulla superficie del dito viene modificata in modo non proporzionale*

tato nello spazio bidimensionale cercando di preservare le distanze tra i punti e le relazioni angolari. Questo metodo è particolarmente utilizzato in caso di oggetti dalla forma irregolare.

Nel caso di un dito umano, è possibile effettuare un unwrapping parametrico utilizzando un cilindro come modello di riferimento, come mostrato in Figura 3.15a. Nelle regioni prossime alla punta del dito, però, questa approssimazione è grossolana. In particolare, non vengono preservate le distanze tra le creste (Figura 3.15b).

Il confronto tra le due tecniche effettuato in [30] mostra che usando tecniche non parametriche gli algoritmi di matching delle impronte digitali forniscono risultati maggiormente simili a quelli ottenuti con tecniche basate su contatto. Ciò è dovuto al fatto che molti algoritmi di matching si basano sulla distanza tra le minuzie per effettuare il riconoscimento

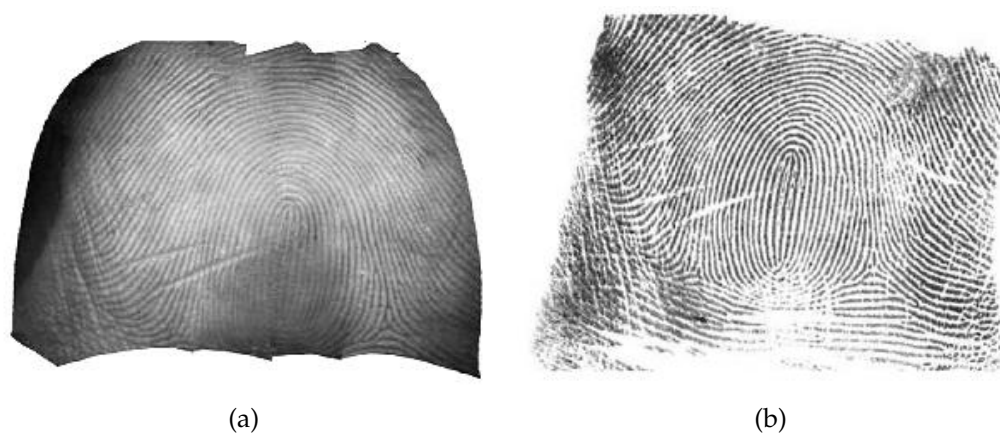


Figura 3.16: *Confronto tra modello tridimensionale dopo unwrapping non parametrico e immagine ottenuta da un sensore a contatto: (a) modello tridimensionale dopo unwrapping; (b) immagine ottenuta da un sensore a contatto*

biometrico. In Figura 3.16 è mostrato il confronto tra un modello tridimensionale sottoposto ad unwrapping non parametrico e la corrispondente immagine acquisita con un sensore a contatto.

3.3.2 Sistemi basati su viste multiple

3.3.2.1 Scansione con specchio ad anello

Un primo approccio viene descritto in [25]. Il sistema proposto fa uso di uno specchio ad anello (Figura 3.17). Mentre l'utente inserisce progressivamente il dito, la telecamera acquisisce sezioni circolari che si sovrappongono parzialmente.

Tramite una trasformazione logaritmico-polare queste sezioni vengono mappate in immagini rettangolari per facilitarne l'elaborazione. Viene poi compensata la distorsione delle sezioni dovuta al posizionamento non esattamente centrale del dito e successivamente le differenziate sezioni vengono correlate in modo da ottenere una rappresentazione completa del dito.

Il problema principale di questo approccio è rappresentato dalla com-

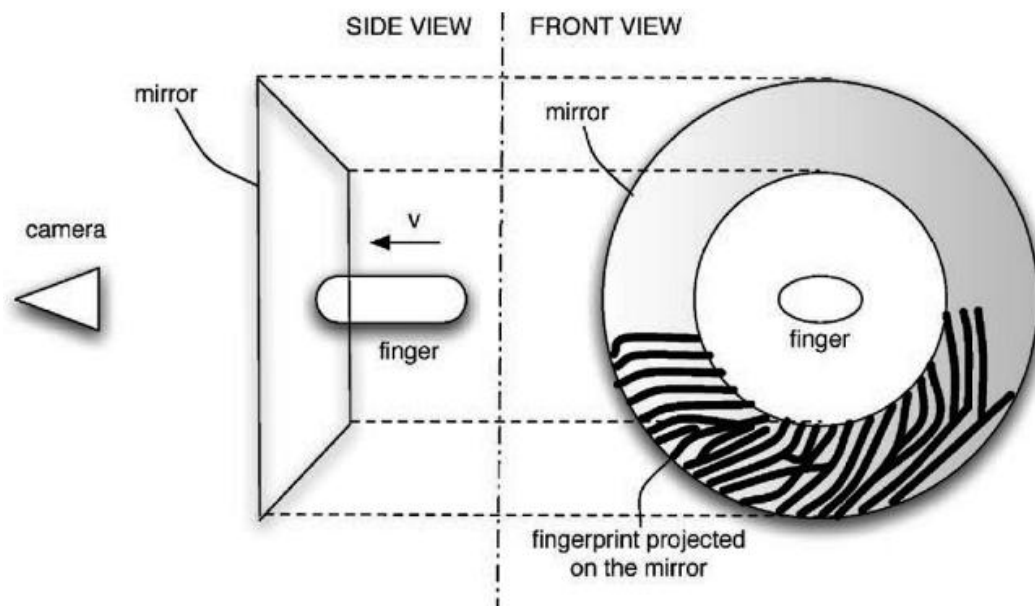


Figura 3.17: Il sistema a specchio circolare

percezione della velocità non costante del dito all'interno dello specchio.

3.3.2.2 Scansione con telecamera a scansione lineare

L'azienda TBS Holding AG ha proposto in [30] un approccio basato su scansione lineare del dito. Questa tecnica fa uso di una telecamera a scansione lineare (*linescan*) e due specchi (Figura 3.18) in modo da sfruttare un'unico dispositivo di acquisizione per ottenere tre differenti viste del dito. Le viste si sovrappongono parzialmente, in modo che rendere possibile la fusione delle immagini in una sola rappresentazione.

La fotocamera percorre tutta la lunghezza del dito ed acquisisce le immagini in modo da ottenere una rappresentazione completa del dito. Durante l'acquisizione delle immagini, vengono proiettati due pattern di linee sul dito. La conoscenza della posizione del pattern, ottenuta tramite una procedura di calibrazione offline, permette di fondere le immagini in un'unica rappresentazione tridimensionale.

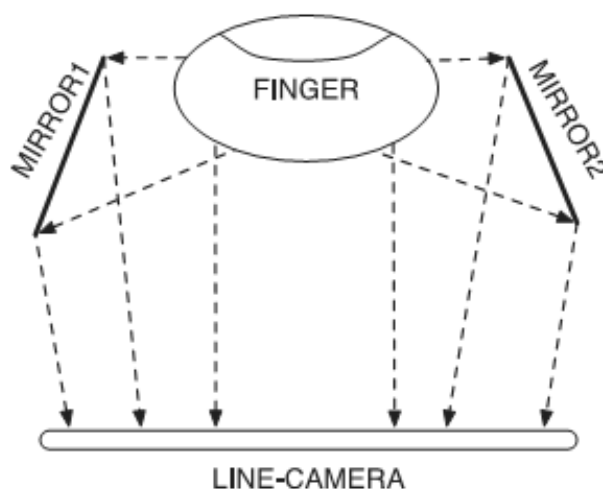


Figura 3.18: *Sistema a scansione circolare*

3.3.2.3 The Surround Imager™

Un ulteriore dispositivo proposto da TBS Holding AG viene descritto in [31]. Il dispositivo, mostrato in Figura 3.19, è composto da cinque telecamere e sedici led puntati verso il centro. I led utilizzano una luce verde in quanto produce immagini con un contrasto maggiore.

Il dito viene poggiato su un supporto per evitare movimenti indesiderati. Un algoritmo, inoltre, aiuta l'utente a posizionarlo correttamente in tempo reale.

La tecnica di ricostruzione del modello tridimensionale del dito utilizzata da questa tecnologia usa tecniche di stereoscopia e fotogrammetria.

- Le telecamere vengono calibrate in precedenza;
- la posizione della telecamera centrale è scelta in modo da catturare i punti singolari (delta e core) dell'impronta digitale;
- le altre telecamere sono posizionate in modo da sovrapporre parzialmente i rispettivi campi visivi;

- il dito è appoggiato su un supporto al fine da mantenere una posizione tale da essere presente nel campo visivo di tutte e cinque le telecamere;
- la risoluzione delle immagini acquisite varia da un minimo di 500 dpi ai bordi dell'immagine fino a 700 dpi nelle zone centrali.

Il procedimento utilizzato per ottenere il modello tridimensionale del dito è il seguente:

1. il dito viene appoggiato sul supporto per iniziare l'acquisizione;
2. l'acquisizione è ripetuta 16 volte in 120 ms. La durata del processo di acquisizione deve essere la più piccola possibile in modo da minimizzare gli effetti dovuti al movimento del dito. L'acquisizione viene ripetuta numerose volte, cambiando ogni volta la potenza dei led che illuminano il dito. In questo modo è possibile disporre di più immagini della stessa porzione di dito, ognuna acquisita con una diversa illuminazione. L'utilizzo di immagini multiple con illuminazione variabile facilita l'estrazione delle feature;
3. vengono eliminate dalle immagini le distorsioni introdotte dall'acquisizione;
4. le immagini vengono rettificate;
5. il volume dell'oggetto viene costruito tramite una tecnica di Shape from silhouette. Viene effettuato l'unwrapping del volume e successivamente viene diviso in sezioni in base alle immagini che lo hanno generato;
6. vengono ricercate le corrispondenze tra le coppie di immagini rettificate tramite il calcolo della cross-correlazione;
7. le corrispondenze tra le coppie di immagini sono raffinate sfruttando le sezioni del volume del dito sottoposto ad unwrapping;

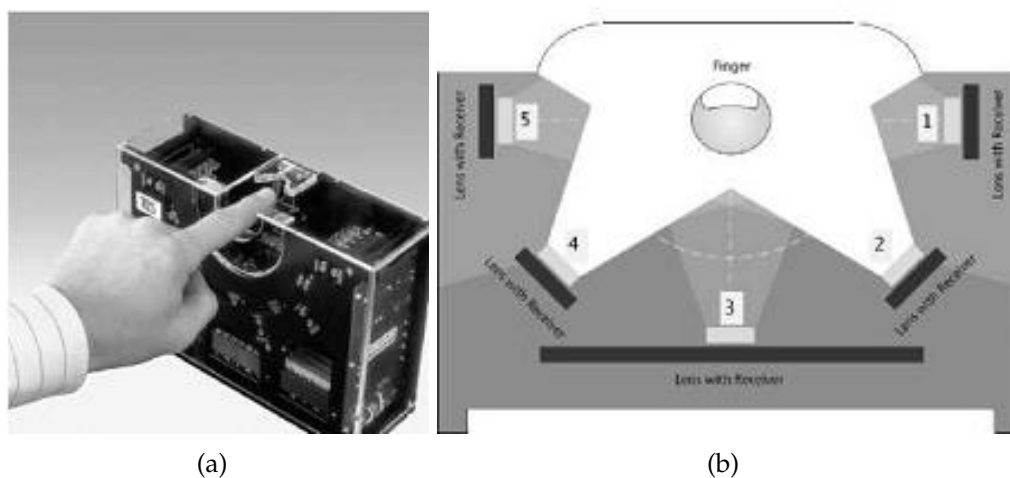


Figura 3.19: Il Surround Imager™ della TBS Holding AG

8. viene effettuata la triangolazione dei punti nello spazio tridimensionale;
9. viene effettuato l'unwrapping non parametrico del modello tridimensionale.

Il modello tridimensionale del dito ricostruito con questa tecnica è mostrato in Figura 3.20.

Il Surround Imager consente, inoltre, di mantenere il modello tridimensionale del dito senza sottoporlo alla fase di unwrapping e senza modificare, quindi, la posizione relativa delle minuzie. In questo modo è possibile effettuare il matching delle impronte digitali in uno spazio tridimensionale.

A questo scopo, è necessario introdurre una nuova rappresentazione delle minuzie. In un modello tridimensionale del dito, si può estendere il sistema di descrizione bidimensionale delle minuzie, che usa tre valori (x, y, θ) , in una tupla di cinque elementi (x, y, z, θ, ρ) , come mostrato in Figura 3.21.

L'algoritmo utilizzato per effettuare il matching delle impronte digitali in uno spazio tridimensionale è una variante dell'algoritmo proposto in [32]. Tramite questo algoritmo, la nuvola di punti generata dalla posizione

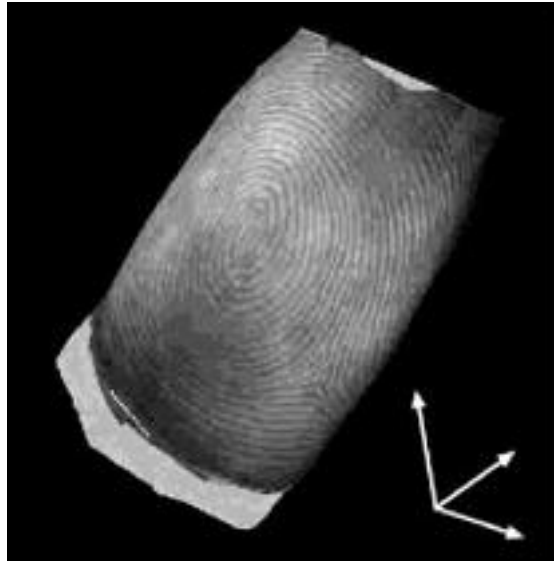


Figura 3.20: Esempio di ricostruzione tridimensionale

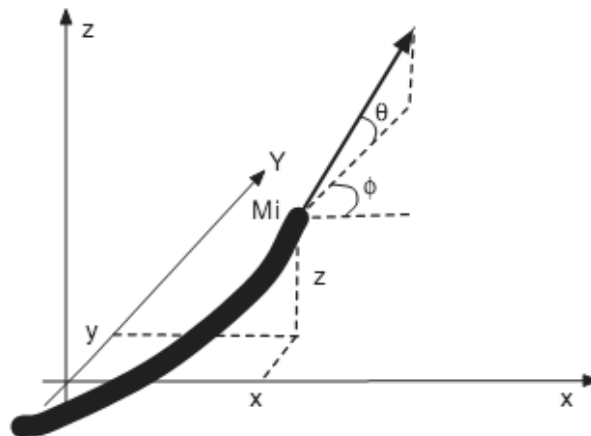


Figura 3.21: Rappresentazione tridimensionale di una minuzia con coordinate (x, y, z) e orientamento (ϕ, θ)

delle minuzie nello spazio tridimensionale viene trasformata in una mesh di triangoli. Successivamente, vengono estratte dai triangoli alcune caratteristiche, come lunghezza dei lati, angoli interni, angoli tra minuzia e lato del triangolo, etc. Queste caratteristiche sono usate per calcolare il grado di correlazione tra le impronte digitali.

3.3.3 Sistema basato su luce strutturata

In [33] è proposta una tecnica per la costruzione di un modello tridimensionale del dito basata su luce strutturata. Il procedimento è riassumibile in quattro passi:

1. un pattern di luce strutturata è proiettato sul dito;
2. la deformazione del pattern è analizzata per costruire il modello tridimensionale;
3. il modello è convertito in un'immagine bidimensionale tramite tecniche di unwrapping parametrico;
4. l'immagine viene migliorata in modo da essere compatibile con gli algoritmi di estrazione delle minuzie;

L'algoritmo usato per la scansione con la luce strutturata è il Phase Measuring Philometry (PMP), derivato dalle tecniche classiche degli interferometri ottici. Il PMP usa meno pattern a parità di precisione rispetto ad altri algoritmi basati su luce strutturata. In particolare, proietta una onda sinusoidale shiftata e cattura il pattern di deformazione per ogni shift impiegato.

Il modello tridimensionale ottenuto dall'elaborazione del pattern di deformazione è convertito in un'immagine bidimensionale con tecniche di unwrapping parametriche, ricercando il cilindro che si adatta meglio alla forma del dito.

L'immagine ottenuta viene convertita in scala di grigi e binarizzata, in modo da estrarre un pattern di creste e valli riconoscibile dagli algoritmi di estrazione delle minuzie.

In Figura 3.22 si vede il confronto tra un'immagine classica e una ottenuta da un modello tridimensionale costruito con questo metodo. La seconda ha una qualità costante ed è esente da distorsioni. Tramite questa tecnologia, le immagini ottenute partendo dal modello tridimensionale presentano un maggiore numero di feature ad alta qualità rispetto a quelle

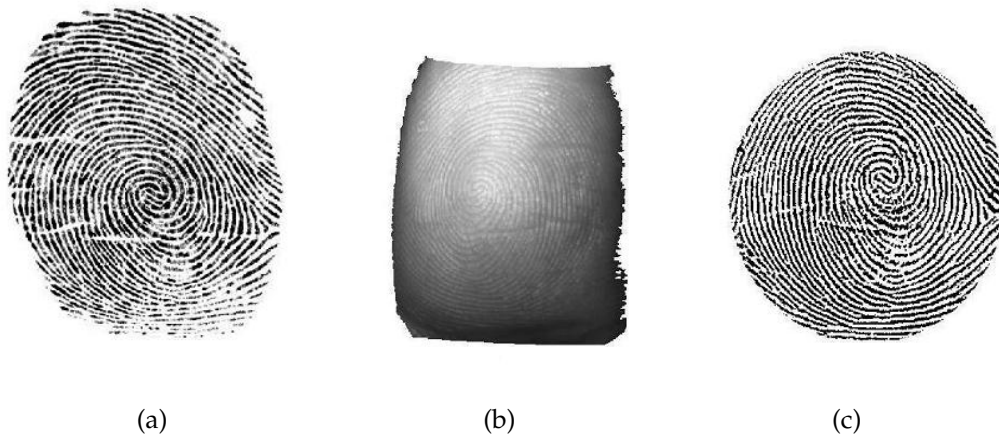


Figura 3.22: *Confronto tra un'immagine classica e un'immagine ottenuta da un modello tridimensionale: (a) immagine acquisita con tecniche classiche; (b) modello tridimensionale; (c) immagine ottenuta da unwrapping e binarizzazione*

ottenute tramite tecniche di acquisizione classiche basate sul contatto del dito con un sensore.

Questa tecnologia impiega, inoltre, una rappresentazione tridimensionale dell'impronta digitale che tiene conto dell'altezza della cresta rispetto alle valli circostanti. Poichè il livello di grigio di un'immagine acquisita senza contatto è direttamente collegato all'altezza della creste, la tupla $(x, y, I(x, y))$ rappresenta una coordinata in uno spazio tridimensionale in cui l'altezza del punto (x, y) è rappresentata dal livello di grigio $I(x, y)$ dell'immagine in quel punto. In Figura 3.23 è mostrata una porzione di un'immagine di un'impronta digitale visualizzata tramite questa rappresentazione.

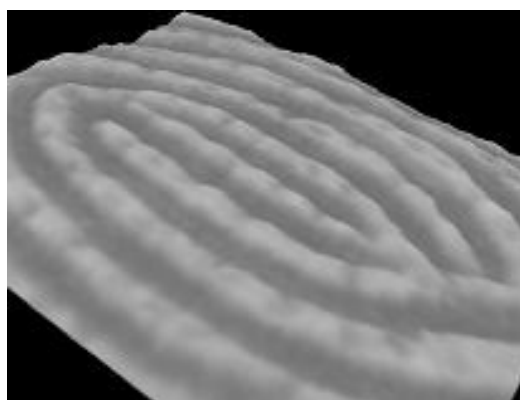


Figura 3.23: *Un'immagine dell'impronta digitale in scala di grigi è rappresentabile come modello tridimensionale*

ELABORAZIONI TRIDIMENSIONALI DI IMPRONTE DIGITALI ACQUISITE SENZA CONTATTO

In questo capitolo vengono presentati alcuni metodi per il calcolo di rappresentazioni tridimensionali delle impronte digitali acquisite senza contatto. Per prima cosa vengono descritti un metodo di miglioramento delle immagini di calibrazione ed il calcolo dell'errore della procedura di calibrazione, successivamente vengono descritti due procedimenti per il calcolo di un modello tridimensionale del dito.

Il primo metodo proposto è basato sulla segmentazione dell'immagine, sulla ricerca delle minuzie e sul calcolo della loro posizione nello spazio tridimensionale. Il secondo metodo, invece, sfrutta la posizione di alcuni punti di riferimento all'interno del dito per calcolare un modello tridimensionale cilindrico dell'ingombro del dito stesso, su cui vengono successivamente applicati i dettagli dell'impronta digitale. Successivamente viene presentato un metodo di elaborazione delle immagini ottenute tramite la proiezione nello spazio bidimensionale del modello tridimensionale.

Infine, è proposto un metodo per il calcolo di un modello tridimension-

ale sintetico del dito.

4.1 Calibrazione del sistema stereoscopico

4.1.1 Miglioramento delle immagini di calibrazione

Durante il procedimento di acquisizione delle immagini di calibrazione, descritto in 2.4.4.2, è possibile che alcune delle immagini ottenute non siano correttamente elaborabili in quanto il software di calibrazione potrebbe non essere in grado di riconoscere tutti i corner presenti nelle scacchiere rappresentate. Ciò può essere causato da un'inquadratura non corretta della scacchiera, oppure da una scarsa qualità dell'illuminazione.

Nella Figura 4.1 è mostrato il confronto tra due acquisizioni di una stessa scacchiera, in posizioni diverse, effettuate con la stessa fotocamera. Nonostante l'intera scacchiera sia stata inquadrata in entrambe le immagini, nell'immagine (a) si vede come tutti i corner siano stati riconosciuti, mentre nell'immagine (b) un corner non è stato riconosciuto.

Una possibile soluzione a questo problema è rappresentata dall'utilizzo di software di acquisizione in tempo reale che verifichino la validità al fine del processo di calibrazione di ogni immagine acquisita. Le eventuali immagini non adatte verrebbero quindi scartate, richiedendo una nuova acquisizione.

Nel caso in cui non fosse possibile realizzare questo tipo di processo, è necessario acquisire le immagini in un primo momento e successivamente utilizzarle ai fini della calibrazione. La qualità delle immagini così ottenute può comunque essere migliorata: infatti, fornendo all'algoritmo di calibrazione un'immagine binaria, creata a partire dall'immagine originale a colori (o in scala di grigi), è più probabile che i corner siano riconosciuti correttamente. In particolare, la situazione ideale si ha quando i corner sono puntiformi ed i quadrati che compongono la scacchiera sono di intensità uniforme. Ogni altra informazione presente nell'immagine risulta superflua ad una corretta calibrazione.

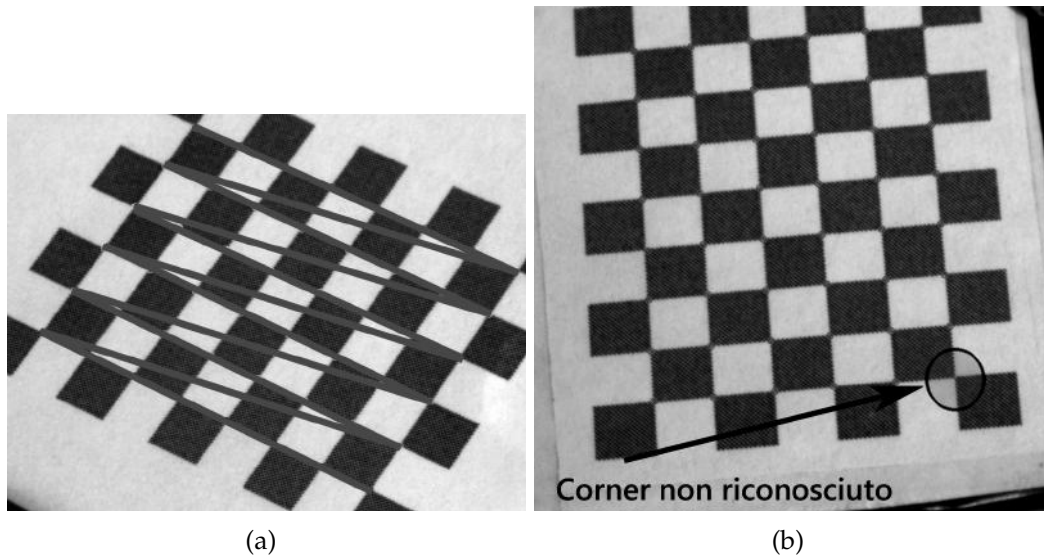


Figura 4.1: Confronto tra due acquisizioni della stessa scacchiera: (a) tutti i corner sono riconosciuti e evidenziati dalle linee che li collegano; (b) un corner (evidenziato in basso a destra) non è stato riconosciuto

L'algoritmo di binarizzazione delle immagini proposto, schematizzato in Figura 4.2, sfrutta il fatto che gli andamenti dei livelli di grigio dei quadrati e dello sfondo sono ben distinti nell'istogramma dell'immagine di una scacchiera, come mostrato in Figura 4.3. È quindi possibile analizzare l'istogramma per ricercare un valore di soglia ottimale tramite il quale binarizzare l'immagine. Tale valore di soglia ha lo scopo di dividere la distribuzione dei livelli di grigio dei quadrati della scacchiera dalla distribuzione dei livelli di grigio dello sfondo.

La soglia ideale corrisponde al punto in cui termina la distribuzione dei livelli di grigio corrispondenti alla scacchiera.

A questo scopo viene calcolata un'approssimazione numerica della derivata dell'istogramma. La funzione usata per il calcolo della derivata calcola la pendenza relativa tra valore dell'istogramma e il valore dell'istogramma cinque punti precedenti, secondo la formula:

$$D_H(i) = \frac{H(i) - H(i-5)}{i - (i-5)} \quad , \quad (4.1)$$

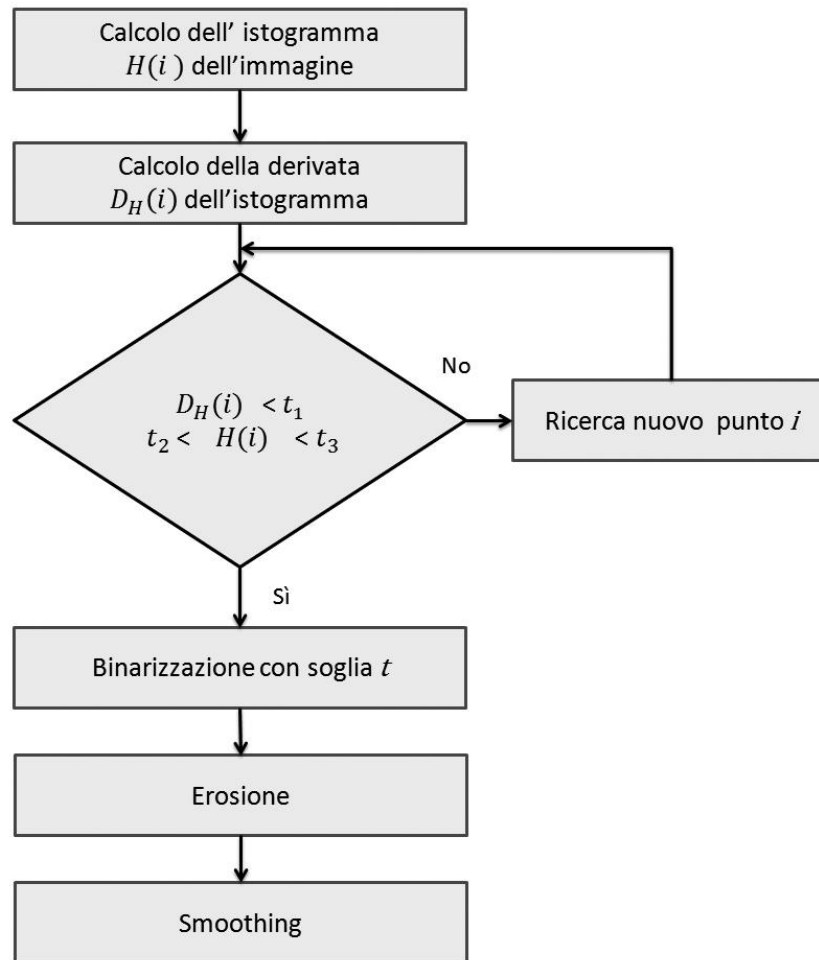


Figura 4.2: Schema a blocchi dell'algoritmo di binarizzazione adottato per le immagini delle scacchiere di calibrazione

dove $H(i)$ rappresenta il valore dell'istogramma e i è l' i -esimo livello di grigio. Il calcolo della derivata numerica campionata ad un passo costante superiore ad 1 permette di ottenere una rappresentazione meno sensibile alle piccole variazioni di pendenza dell'istogramma.

Viene quindi ricercato l'ultimo punto i in cui la derivata numerica sia inferiore ad una soglia fissata e in cui il valore dell'istogramma sia compreso in un determinato intervallo, secondo la formula:

$$t = \arg \max_{0 \leq i \leq 255} ((D_H(i) \leq t_1) \text{ and } (t_2 \leq H(i) \leq t_3)) \quad , \quad (4.2)$$

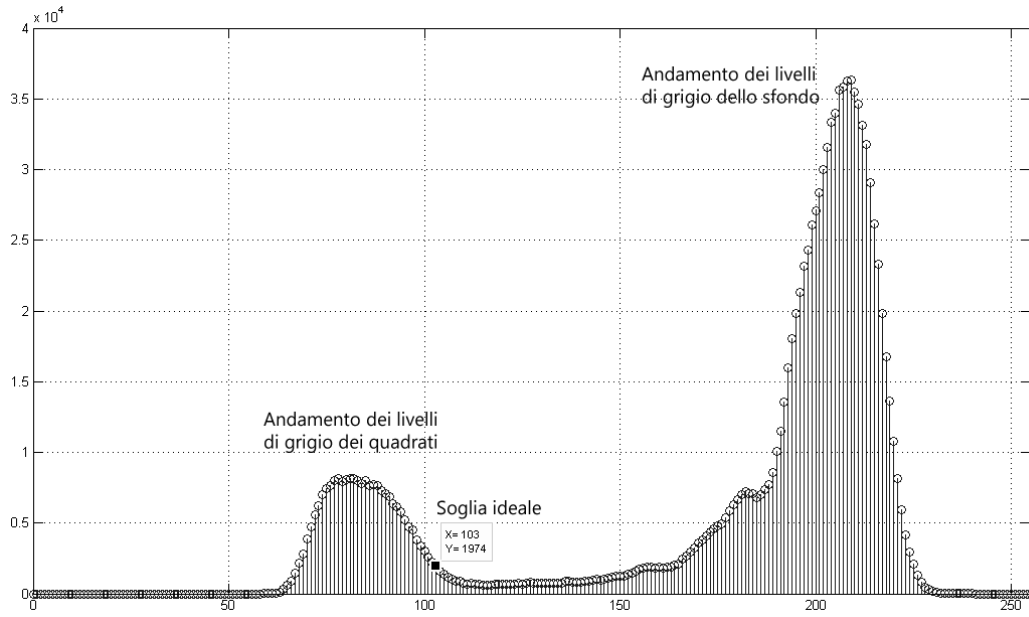


Figura 4.3: Istogramma dell'immagine di una scacchiera

in cui t_1 , t_2 e t_3 sono valori di soglia stimati empiricamente.

L'immagine viene quindi binarizzata utilizzando tale valore, secondo la formula:

$$\begin{cases} B(x, y) = 1 & \text{se } I(x, y) \geq t \\ B(x, y) = 0 & \text{se } I(x, y) < t \end{cases} \quad , \quad (4.3)$$

dove $I(x, y)$ è l'immagine da binarizzare e $B(x, y)$ è l'immagine binarizzata. È possibile, però, che la binarizzazione non avvenga in modo ideale e che alcuni pixel corrispondenti ai quadrati della scacchiera siano considerati come parte dello sfondo. Per questo motivo l'immagine binaria viene sottoposta ad un'operazione morfologica di erosione ([34]).

Infine, per compensare l'effetto di frastagliatura introdotto dall'operatore morfologico e rendere più lineari i bordi, si applica all'immagine un

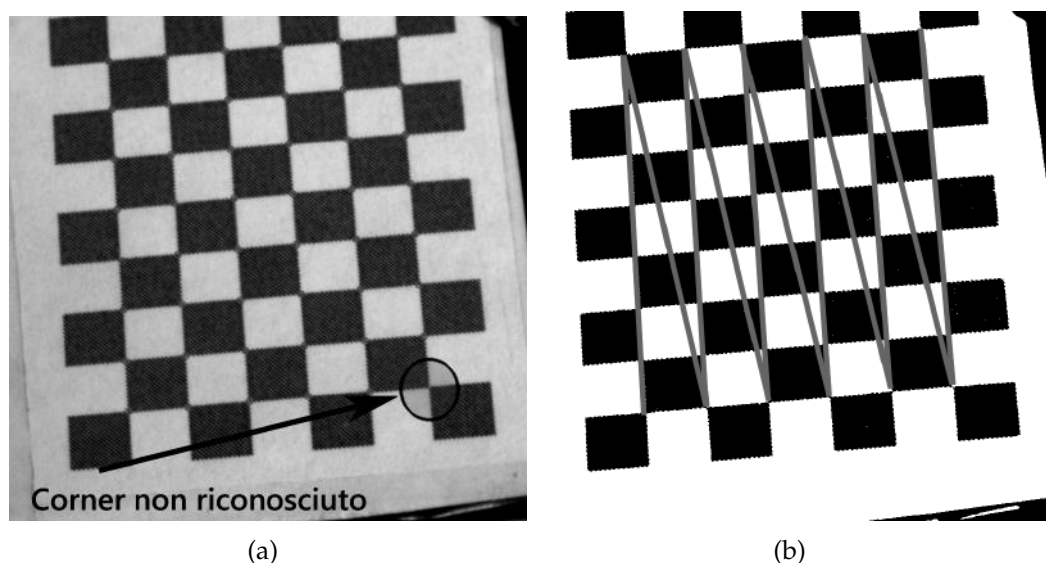


Figura 4.4: Confronto tra il riconoscimento degli angoli sull'immagine originale e binarizzata: (a) non tutti i corner sono riconosciuti; (b) tutti i corner sono riconosciuti correttamente e evidenziati dalle linee che li collegano

filtro di smoothing di tipo *average*, secondo la formula:

$$B'(x,y) = \frac{1}{16} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \star B(x,y) \quad , \quad (4.4)$$

dove $B'(x,y)$ è l'immagine dopo l'operazione di smoothing. In Figura 4.4 è mostrato il confronto tra il riconoscimento dei corner, applicato all'immagine originale e alla corrispondente immagine binarizzata.

4.1.2 Errore di calibrazione

La calibrazione del sistema stereoscopico è effettuata con il procedimento descritto nella sezione 2.6. Utilizzando i parametri ottenuti dalla calibrazione, è possibile sfruttare la posizione dei corner nelle immagini delle

scacchiere per triangolarne la posizione nello spazio tridimensionale. In questo modo, si può valutare l'accuratezza della calibrazione confrontando la posizione dei punti tridimensionali con la posizione dei punti appartenenti ad una scacchiera piana ideale.

Come descritto in [35], la posizione dei corner triangolati nello spazio tridimensionale è utilizzata per interpolare un piano passante per i punti. Idealmente, ogni punto ottenuto dovrebbe essere disposto sul medesimo piano: lo scostamento della posizione dei punti rispetto al piano fornisce una misura dell'errore di calibrazione. Esistono numerosi metodi per calcolare il piano interpolante. Il metodo proposto usa un piano interpolante descritto dall'equazione:

$$z = ax + by + c \quad , \quad (4.5)$$

dove i coefficienti a, b, c sono i componenti del vettore C , calcolato come:

$$C = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ \dots & \dots & \dots \\ x_n & y_n & 1 \end{bmatrix} / \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \dots \\ z_n \end{bmatrix} \quad , \quad (4.6)$$

dove i valori (x_i, y_i, z_i) sono le coordinate dei corner nello spazio tridimensionale. In Figura 4.5 sono mostrati un insieme di corner ricostruiti ed il corrispondente piano interpolante.

Tramite la misura delle distanze euclidee tra i punti ed il piano interpolante, è possibile definire i seguenti parametri di errore:

- differenza tra la diagonale della scacchiera fisica e la diagonale calcolata in seguito alla ricostruzione tridimensionale:

$$|D_f - D_v| \quad , \quad (4.7)$$

dove D_f è la lunghezza della diagonale reale e la lunghezza D_v è calcolata tramite la distanza euclidea tra il primo e l'ultimo corner,

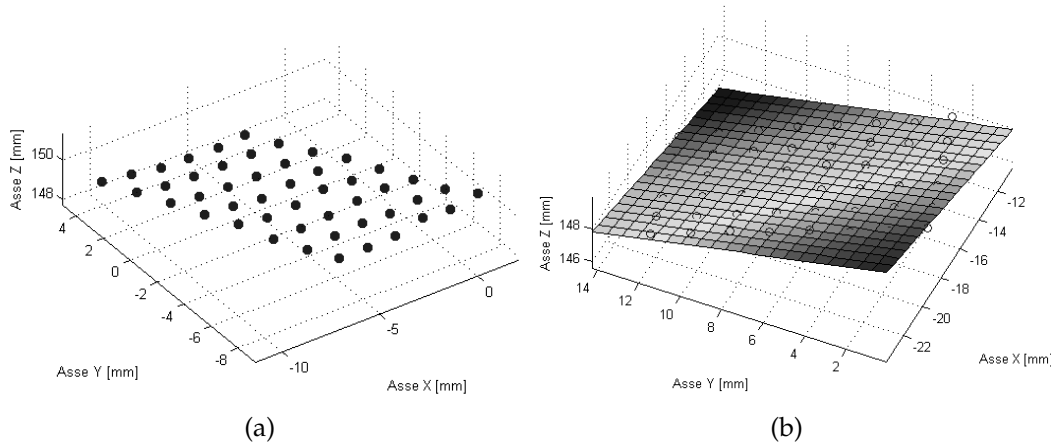


Figura 4.5: *Insieme di corner della scacchiera di calibrazione ricostruiti nello spazio tridimensionale e il corrispondente piano interpolante: (a) i punti nello spazio tridimensionale; (b) il piano interpolante*

posizionati agli estremi di una diagonale:

$$D_v = \sqrt{(x_1 - x_n)^2 + (y_1 - y_n)^2 + (z_1 - z_n)^2} \quad , \quad (4.8)$$

dove (x_1, y_1, z_1) e (x_n, y_n, z_n) sono rispettivamente le coordinate dei due punti;

- differenza tra l'area della scacchiera fisica e l'area calcolata in seguito alla ricostruzione tridimensionale:

$$|A_f - A_v| \quad , \quad (4.9)$$

dove A_f è l'area reale e A_v è l'area calcolata moltiplicando la lunghezza dei due lati:

$$A_v = l_1 l_2 \quad . \quad (4.10)$$

I valori l_1 e l_2 sono a loro volta calcolati tramite la distanza euclidea dei punti corrispondenti:

$$\begin{aligned} l_1 &= \sqrt{(x_3 - x_4)^2 + (y_3 - y_4)^2 + (z_3 - z_4)^2} \\ l_2 &= \sqrt{(x_5 - x_6)^2 + (y_5 - y_6)^2 + (z_5 - z_6)^2} \quad , \end{aligned} \quad (4.11)$$

in cui (x_3, y_3, z_3) e (x_4, y_4, z_4) sono rispettivamente le coordinate dei punti disposti in corrispondenza delle estremità di un lato corto della scacchiera e (x_5, y_5, z_5) e (x_6, y_6, z_6) rispettivamente corrispondono alle estremità di un lato lungo della scacchiera;

- distanza media dei punti ricostruiti dal piano interpolante:

$$E_m = \frac{1}{n} \sum_{i=1}^n E_i \quad , \quad (4.12)$$

dove E_i sono le distanze del punto triangolato i dal piano interpolante:

$$E_i = z_i - z_p \quad , \quad (4.13)$$

dove z_i è la profondità del punto i e z_p è la profondità del piano interpolante alle coordinate (x_i, y_i) ;

- distanza massima dei punti ricostruiti dal piano interpolante:

$$E_{Max} = \max_{1 \leq i \leq n} \{E_i\} \quad ; \quad (4.14)$$

- deviazione standard della distanza dei punti ricostruiti dal piano interpolante:

$$\sigma_E = \sqrt{\frac{\sum_{i=1}^n (E_i - E_m)^2}{n}} \quad . \quad (4.15)$$

Per poter confrontare i valori fisici con quelli calcolati in seguito alla ricostruzione tridimensionale, è necessario rapportare la scala del modello tridimensionale a dei valori reali. Il software OpenCV utilizza una descrizione dello spazio tridimensionale che fa corrispondere un passo della scala alla lunghezza di un quadrato della scacchiera. Conoscendo la dimensione fisica dei quadrati, è possibile assegnare alle coordinate dei punti triangolati un valore confrontabile con le dimensioni fisiche, tramite le formule:

$$\begin{bmatrix} x_f \\ y_f \\ z_f \end{bmatrix} = l_q \begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} \quad , \quad (4.16)$$

dove l_q è la lunghezza fisica del lato dei quadrati della scacchiera e (x_f, y_f, z_f) e (x_v, y_v, z_v) sono rispettivamente le coordinate reali e le coordinate dei punti triangolati.

4.2 Ricostruzione tridimensionale basata sulle minuzie

Il primo metodo proposto per il calcolo di un modello tridimensionale si basa sulla ricerca delle correlazioni tra le immagini, come descritto in 2.4.4.5, per determinare il valore di profondità degli oggetti. In particolare, vengono ricercate le corrispondenze tra le minuzie dell'impronta digitale per ottenerne una rappresentazione tridimensionale discreta. Il modello così ottenuto viene poi completato tramite l'aggiunta del contorno del dito, come schematizzato in Figura 4.6.

Per determinare la posizione delle minuzie nello spazio tridimensionale è necessario, per prima cosa, calcolare la regione di interesse delle immagini corrispondente alla zona dell'impronta digitale con la qualità migliore.

Successivamente, viene effettuata un'operazione di preprocessing in modo da ottenere immagini di impronte digitali elaborabili con i correnti software per l'estrazione delle minuzie.

Le minuzie estratte dalle immagini sono confrontate tramite l'utilizzo di un software di matching, in modo da determinare le corrispondenze stereoscopiche. Le coordinate delle minuzie corrispondenti sono infine utilizzate per triangolarne la posizione nello spazio tridimensionale.

Una volta estratto il contorno del dito tramite un'operazione di segmentazione, è necessario determinare la corrispondenza stereoscopica di almeno un punto appartenente al contorno per calcolarne la posizione nello spazio tridimensionale. Se l'acquisizione dell'immagine del dito è avvenuta frontalmente, infatti, è possibile assumere che l'intera regione del contorno sia alla medesima profondità. Una volta determinato il valore di profondità del bordo tramite la corrispondenza stereoscopica, è quindi possibile assegnare tale valore a tutti i punti del bordo, ottenendo una rappresentazione tridimensionale dello spazio occupato dal dito.

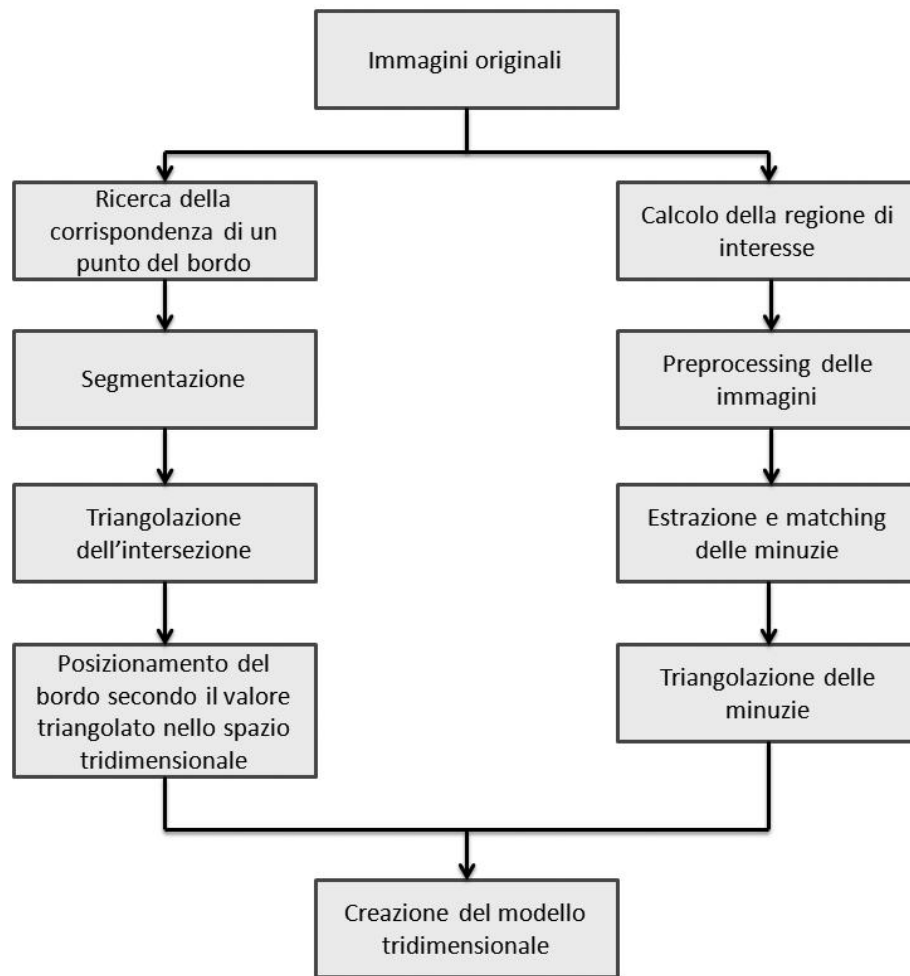


Figura 4.6: *Schema a blocchi della ricostruzione tramite triangolazione delle minuzie*

4.2.1 Segmentazione dell'immagine

L'algoritmo di estrazione del bordo del dito si basa sull'operatore di Roberts per l'estrazione dei bordi presenti nell'immagine. Per prima cosa, è necessario applicare alle immagini un filtro mediano per ridurre il rumore. Il filtro mediano, come descritto in [34], è dato dalla formula:

$$f(x, y) = \text{mediana}\{g(s, t)\}_{(s, t) \in S_{xy}} \quad , \quad (4.17)$$

dove $f(x, y)$ è il nuovo valore del pixel alla posizione (x, y) e $g(s, t)$ è un intorno dell'immagine di dimensione fissata;

Successivamente, viene applicato all'immagine l'operatore di Roberts, che approssima la derivata secondo la formula:

$$\nabla f = \sqrt{(z_9 - z_5)^2 + (z_8 - z_6)^2} \quad , \quad (4.18)$$

dove z_i rappresenta l' i -esimo pixel di un intorno 3×3 dell'immagine. Il risultato è sottoposto quindi a sogliatura in modo da mantenere il più possibile le regioni corrispondenti al bordo del dito ed eliminare dettagli superflui.

L'immagine risultante dall'applicazione del filtro di Roberts è quindi sottoposta ad un'operazione morfologica al fine di eliminare i dettagli superflui ed estrarre solo le componenti relative al bordo. In particolare, viene effettuata un'operazione di chiusura. Tramite un'analisi dei componenti connessi è possibile quindi mantenere solo il componente connesso più esteso, corrispondente alla zona del bordo del dito.

Infine un'operazione di thinning, seguita da un'operazione morfologica di chiusura, mantiene solo la linea principale del bordo. La maschera di segmentazione finale è ottenuta tramite il riempimento della regione compresa dal bordo.

Il procedimento proposto è schematizzato in Figura 4.7.

In Figura 4.8 sono mostrate l'immagine originale, il bordo del dito, la maschera binaria di segmentazione e l'immagine segmentata risultante.

4.2.2 Calcolo della regione di interesse

Il calcolo della regione di interesse dell'immagine permette di ottenere una maschera binaria che mantiene solo le zone dell'immagine contenenti i dettagli dell'impronta digitale con la qualità necessaria per l'estrazione delle minuzie. L'algoritmo si basa sul riconoscimento dell'ultima giunzione del dito e sull'erosione della maschera binaria ottenuta dalla segmentazione, calcolata come descritto in 4.2.1.

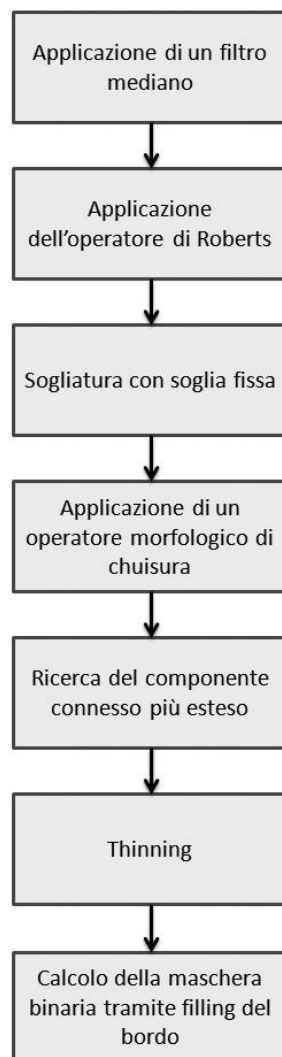


Figura 4.7: *Schema a blocchi del procedimento di segmentazione dell'immagine*

Il riconoscimento dell'ultima giunzione del dito permette di delimitare in modo automatico la regione dell'immagine contenente la punta del dito e di considerare, quindi, solo la zona contenente l'impronta digitale.

Per determinare la posizione della giunzione, che è disposta verticalmente nelle immagini di test, si effettua il calcolo del gradiente orizzontale dell'immagine. La zona corrispondente alla giunzione del dito presenta un gradiente più elevato rispetto alle altre zone dell'immagine, ma inferiore

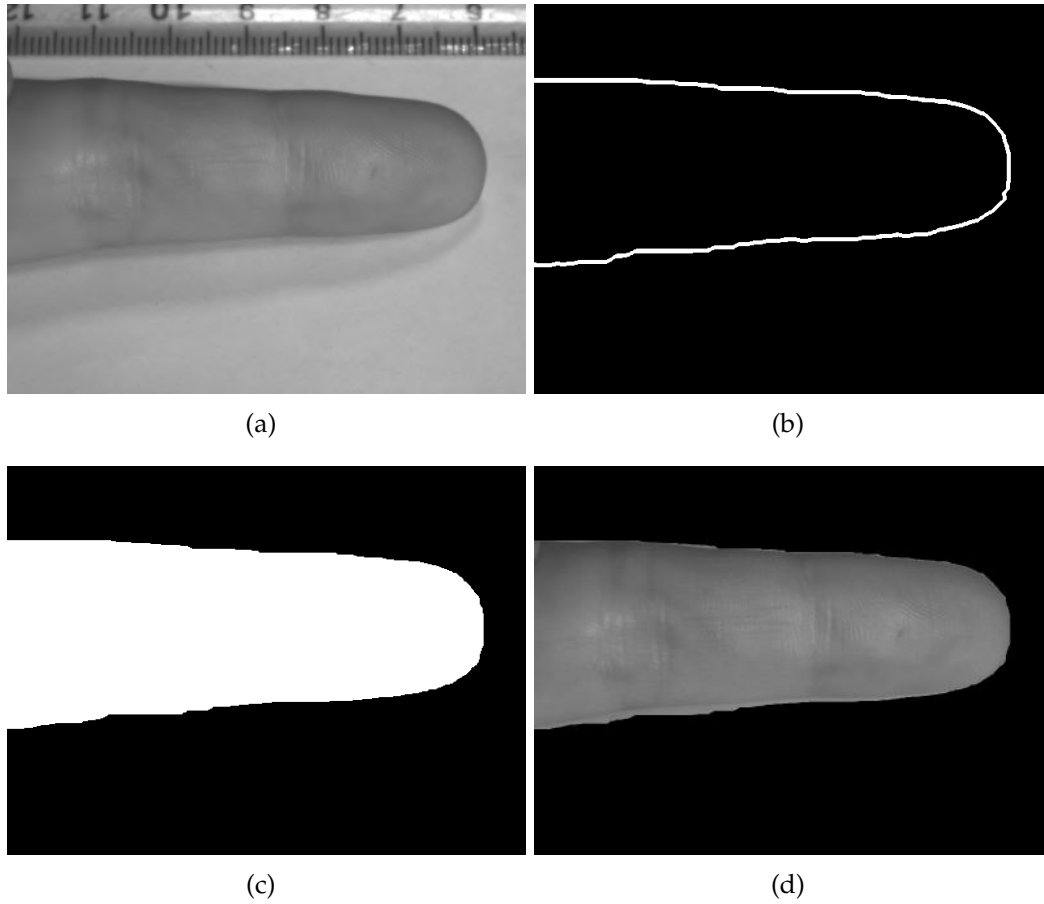


Figura 4.8: Estrazione del bordo e della maschera binaria di segmentazione: (a) immagine originale; (b) bordo del dito; (c) maschera binaria di segmentazione; (d) immagine segmentata

rispetto ai bordi del dito. Per questo motivo il risultato dell'applicazione del gradiente viene sottoposto ad un'operazione volta ad eliminare le zone con gradiente massimo e quelle con gradiente minimo ed a estendere il range dei valori intermedi alla scala di grigi, secondo la formula:

$$I'(x, y) = 255 \cdot \frac{I(x, y) - t_{min}}{t_{max} - t_{min}}, \quad (4.19)$$

dove $I(x, y)$ è l'immagine originale, $I'(x, y)$ è l'immagine risultante e t_{min} e t_{max} sono rispettivamente due valori di soglia stimati empiricamente.

L'immagine risultante viene quindi binarizzata in modo da essere elaborata con operatori morfologici. In particolare viene sottoposta ad un'operazione di chiusura usando come elemento strutturante una linea verticale, in modo da congiungere i pixel appartenenti alla giunzione. Successivamente, tramite il calcolo dei componenti connessi nell'immagine binaria, viene mantenuto solo il componente connesso più esteso.

La posizione della giunzione è quindi approssimata alla media delle coordinate x e y dei punti della zona estratta. Tale posizione è inoltre utilizzata per calcolare la larghezza di riferimento del dito, definita come la larghezza del dito in corrispondenza dell'ultima giunzione.

La posizione dell'ultima giunzione è utilizzata per eliminare dalla maschera binaria di segmentazione le zone che non fanno parte della punta del dito. Per ottenere la regione di interesse (ROI) dell'immagine, corrispondente alla zona centrale della punta del dito, è necessario procedere ad un'operazione morfologica di erosione sulla maschera così ottenuta. È necessario, infatti, eliminare le zone dell'immagine corrispondenti alle regioni periferiche del dito. I test hanno dimostrato che utilizzando una dimensione dell'elemento strutturante pari a circa un quarto della larghezza del dito si mantiene solo la zona dell'immagine con la qualità sufficiente.

Il procedimento proposto è schematizzato in Figura 4.9. In Figura 4.10 è mostrato il risultato dell'estrazione della giunzione del dito, mentre in Figura 4.11 sono mostrate la maschera di segmentazione limitata alla punta del dito e la maschera di segmentazione erosa.

4.2.3 Preprocessing delle immagini

La fase di preprocessing delle immagini ha il compito di trasformare le immagini originali, acquisite da una fotocamera, in immagini quanto più simili a quelle ottenute con un sensore dedicato. A questo scopo sono stati testati differenti approcci:

- **Metodo A: preprocessing basato su filtro passa-banda:** consiste nel-

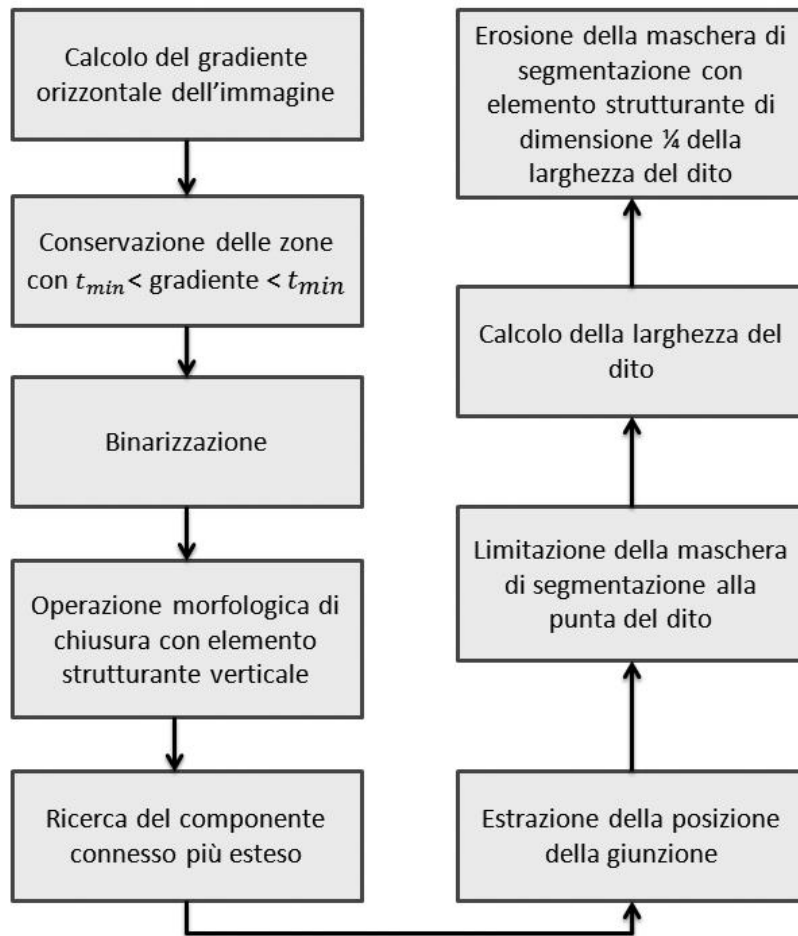


Figura 4.9: Schema a blocchi del calcolo della regione di interesse

l'utilizzo di un filtraggio in frequenza che mantiene solo le frequenze dell'immagine corrispondenti alle creste dell'impronta digitale. A questo scopo l'immagine viene convertita in scala di grigi e viene calcolata la trasformata di Fourier. Successivamente all'immagine trasformata viene applicato un filtro calcolato secondo la formula:

$$F = e^{-\left((X-\frac{\xi^2}{2}) + \left(\frac{Y-r^2}{\sigma^2}\right)\right)} - e^{-\left((X-\frac{\xi^2}{2}) + \left(\frac{Y-r^2}{\sigma^3}\right)\right)} \quad (4.20)$$

L'immagine finale è infine calcolata tramite l'antitrasformata di Fourier;

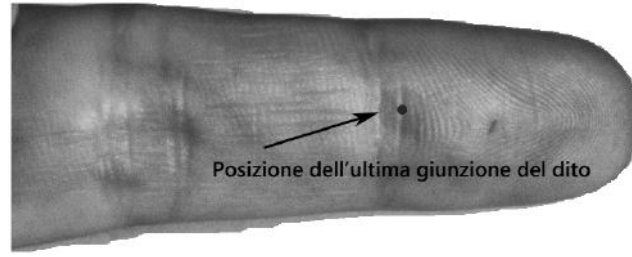


Figura 4.10: *Posizione dell'ultima giunzione del dito*

- **Metodo B: preprocessing basato su operatore Fast Radial:** l'operatore Fast Radial, descritto in [36], risalta le circonferenze con un determinato raggio ed è usato per rilevare l'andamento delle creste;
- **Metodo C: preprocessing basato sul calcolo del gradiente perpendicolare alle creste:** questo metodo determina l'orientamento delle creste in modo che il gradiente, applicato in direzione perpendicolare alle creste stesse, risalti l'alternarsi dei livelli di grigio tipico delle creste e delle valli. Il calcolo del gradiente è effettuato utilizzando gli operatori di Sobel, secondo le formule:

$$\begin{aligned}
 G_x(x, y) &= \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \star I(x, y) \quad ; \\
 G_y(x, y) &= \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \star I(x, y) \quad , \quad (4.21)
 \end{aligned}$$

dove $G_x(x, y)$ e $G_y(x, y)$ rappresentano rispettivamente il gradiente lungo x e lungo y . Successivamente vengono calcolati il modulo $M(x, y)$ e la fase $\alpha(x, y)$ del gradiente, secondo le formule:

$$\begin{aligned}
 M(x, y) &= \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad ; \\
 \alpha(x, y) &= \arctan \left[\frac{G_x(x, y)}{G_y(x, y)} \right] \quad . \quad (4.22)
 \end{aligned}$$

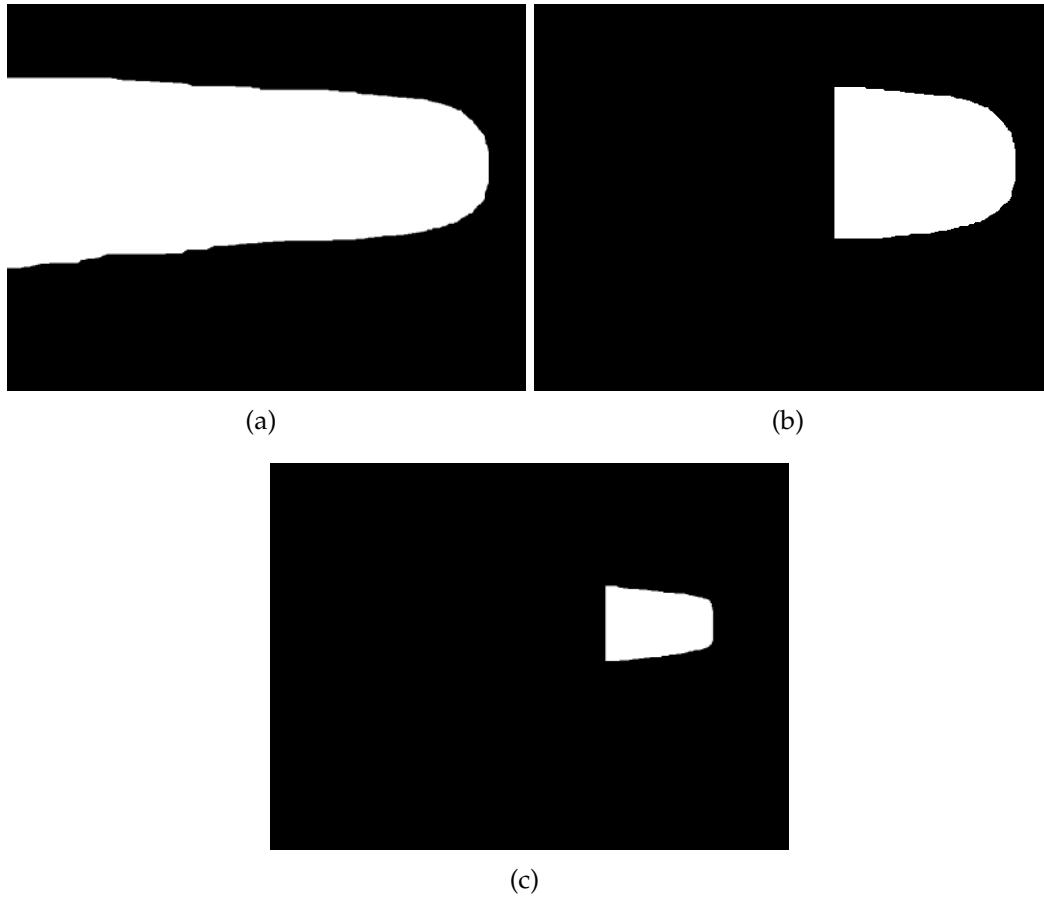


Figura 4.11: *Limitazione della ROI dell'immagine: (a) maschera di segmentazione originale; (b) maschera di segmentazione limitata alla punta del dito; (c) maschera di segmentazione limitata ed erosa*

L'immagine finale è quindi calcolata tramite il prodotto del modulo $M(x, y)$ e il seno della differenza tra la fase del gradiente e l'angolazione locale delle creste:

$$R(x, y) = M(x, y) \cdot \sin(\alpha(x, y) - D(x, y)) \quad , \quad (4.23)$$

dove $D(x, y)$ è la mappa degli orientamenti delle creste;

- **Metodo D: preprocessing basato su sottrazione dello sfondo e binarizzazione:** questo metodo si basa sulla sottrazione dall'immagine delle zone corrispondenti allo sfondo ed alla pelle del dito, secondo

la formula:

$$g(x, y) = f(x, y) - \bar{f}(x, y) \quad ,$$

dove $g(x, y)$ è l'immagine risultante e $f(x, y)$ e $\bar{f}(x, y)$ sono rispettivamente l'immagine originale e l'immagine sottoposta a smoothing. L'immagine risultante viene quindi binarizzata utilizzando aree locali di differenti dimensioni. Le diverse immagini binarizzate sono infine sommate le une con le altre.

In Figura 4.12 sono mostrati i risultati dei filtraggi descritti, applicati ad un'immagine di un impronta digitale acquisita senza contatto.

4.2.4 Estrazione ed elaborazione delle minuzie

In un'impronta digitale, esistono tre livelli di analisi dei dettagli: globale, locale e ultra-fine. A livello globale si osservano l'orientamento e la frequenza delle creste, la posizione dei punti singolari e la forma dell'impronta.

A livello locale si analizzano le caratteristiche locali delle creste, dette minuzie. Esistono circa 150 diversi tipi di minuzie, anche se la maggior parte dei sistemi di identificazione ne considera solo due: terminazioni e biforcazioni, come mostrato in Figura 4.13.

A livello ultra-fine è possibile distinguere, a patto di disporre di una scansione ad alta risoluzione (almeno 1000 dpi), le caratteristiche intra-creste (come i pori per la sudorazione) e inter-creste (come le creste incipienti).

Il processo di estrazione delle minuzie proposto si basa sui dettagli di livello locali e sfrutta il pacchetto software NBIS (*Nist Biometric Image Software*) ([37]) del NIST (*National Institute of Standards and Technology*). In particolare, il software considera solo terminazioni e biforcazioni.

L'estrazione delle minuzie è effettuata con il software NBIS mindtct che restituisce, oltre alla posizione delle minuzie, l'orientamento, la qualità e il

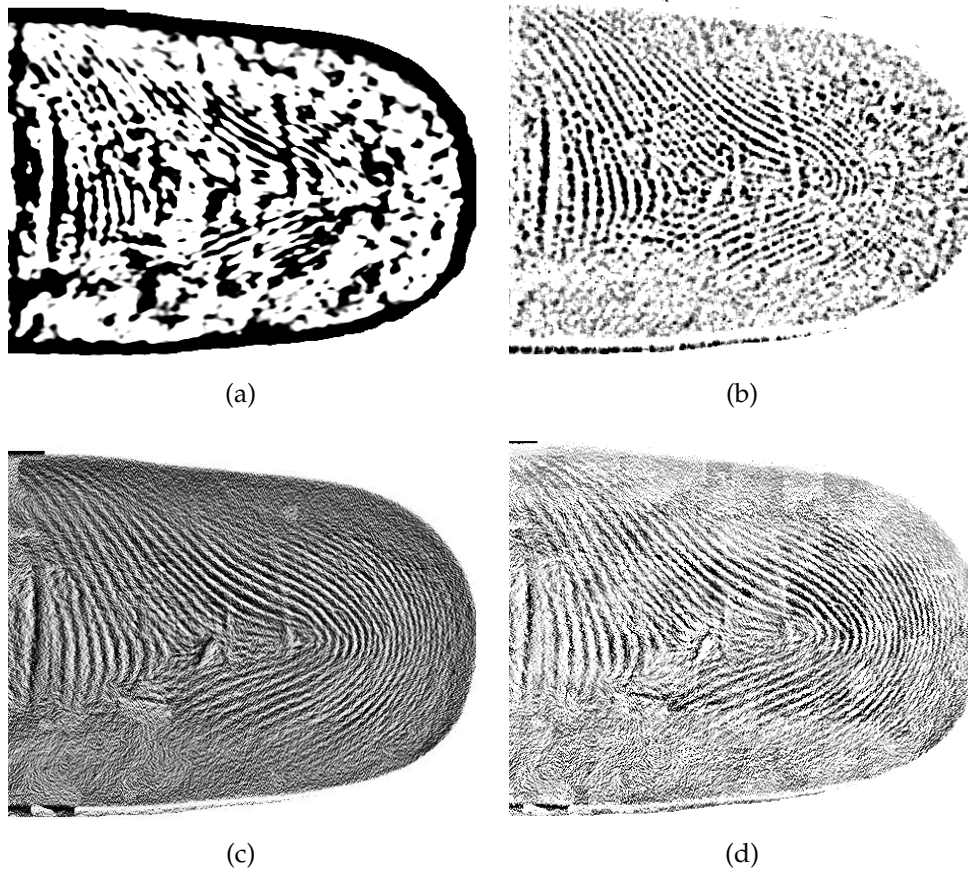


Figura 4.12: *Differenti tipi di preprocessing delle immagini: (a) Metodo A; (b) Metodo B; (c) Metodo C; (d) Metodo D*

tipo delle minuzie stesse. Il software è stato creato per elaborare in modo ottimale immagini acquisite a 500 ppi e con 256 livelli di grigio. L'algoritmo alla base del software mindtct, come descritto in [38] si compone di sei passi principali:

1. **Generazione delle mappe dell'immagine:** in questa fase si calcolano la mappa di qualità delle zone dell'immagine e l'orientamento locale delle creste;
2. **Binarizzazione dell'immagine:** l'informazione dell'orientamento delle creste è utilizzata per determinare se un determinato pixel appartiene

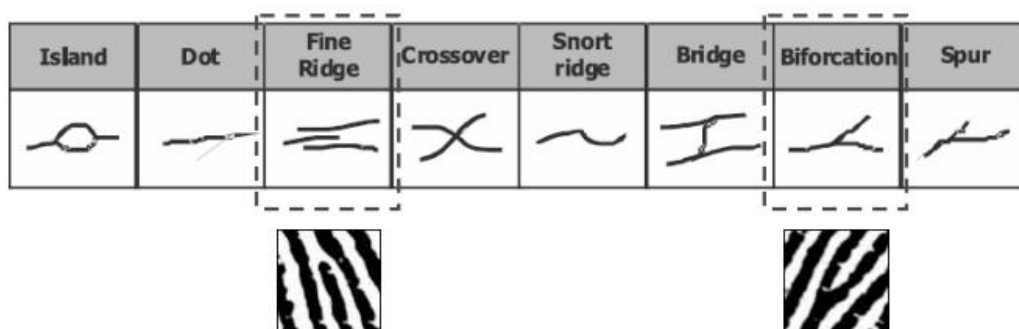


Figura 4.13: *I principali tipi di minuzie*

ad una cresta o ad una valle. L'immagine ottenuta assegna quindi valore 1 alle creste e valore 0 alle valli ed allo sfondo;

3. **Rilevamento delle minuzie:** vengono ricercati nell'immagine alcuni pattern conosciuti, corrispondenti a terminazioni e biforcazioni;
4. **Rimozione delle minuzie false:** si eseguono controlli per rilevare artefatti erroneamente identificati come minuzie;
5. **Calcolo della qualità delle minuzie:** la qualità delle minuzie estratta viene determinata in base alla mappa di qualità e al contrasto delle zone circostanti;
6. **Generazione del file di output:** posizione, orientamento, qualità e tipo delle minuzie vengono salvate su un file.

In Figura 4.14 è mostrato il risultato dell'estrazione delle minuzie effettuando il pre-processing dell'immagine con il metodo basato su sottrazione dello sfondo e binarizzazione. Osservando l'immagine in modo approfondito si rileva come solo una parte delle minuzie estratte dal software mindtct corrisponda a minuzie realmente presenti nell'immagine. In particolare, il software di estrazione delle minuzie è particolarmente sensibile alle zone prossime al bordo dell'immagine, nelle quali la terminazione artificiale delle creste può essere interpretata come minuzia di tipo fine cresta.

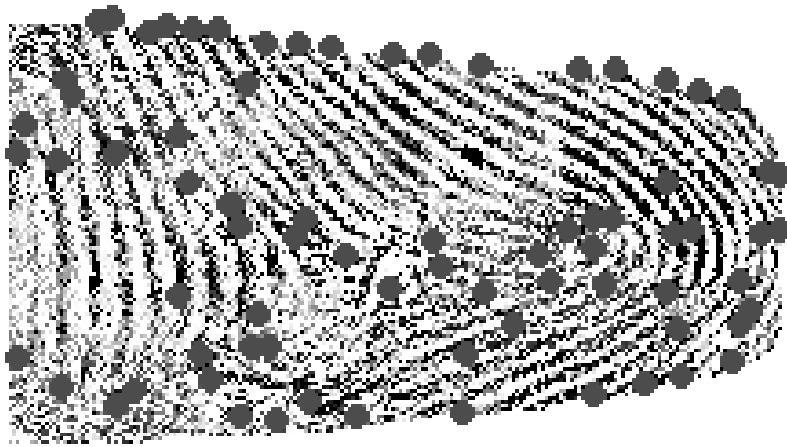


Figura 4.14: Risultato dell'estrazione delle minuzie

Come evidenziato nel particolare in Figura 4.15a, le coordinate delle minuzie estratte dal software del NIST non sono sempre corrispondenti a quelle reali. Osservando l'immagine sottoposta all'operazione di skeleton, inoltre, si osserva come alcune minuzie corrispondano a zone dell'immagine prive di creste.

Per raggiungere una maggiore precisione, è possibile modificare la posizione di ciascuna minuzia in modo da farla coincidere con il pixel nero più vicino. Per ogni minuzia sono quindi calcolate le distanze tra i pixel neri appartenenti ad un intorno di dimensione fissata. Il pixel più vicino tra essi è scelto come nuova posizione della minuzia. In Figura 4.15b è mostrato il risultato di tale operazione.

4.2.4.1 Matching delle minuzie basato su bozorth3

Un metodo per effettuare il calcolo delle corrispondenze tra due immagini stereoscopiche di impronte digitali si basa sul matching delle minuzie estratte dalle due immagini. In questo modo, infatti, si può utilizzare un algoritmo di correlazione robusto, invariante alla rotazione ed alla traslazione delle immagini.

Il software di matching utilizzato è stato il programma bozorth3 del

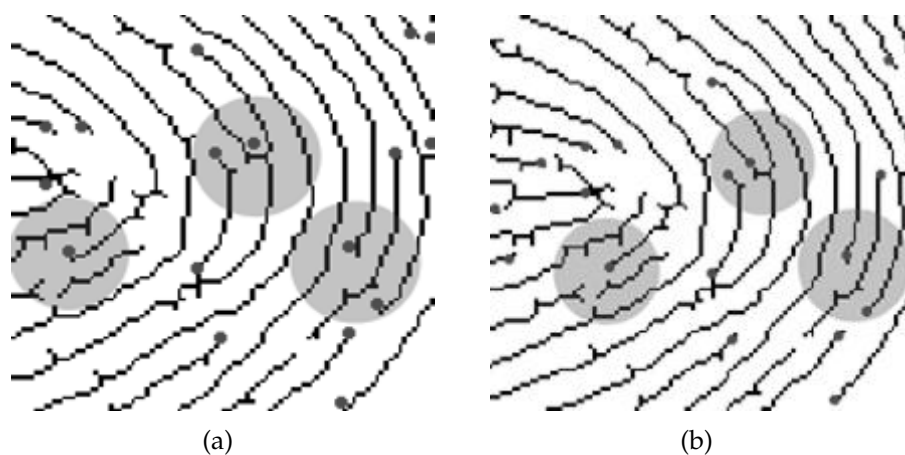


Figura 4.15: *Le coordinate delle minuzie ottenute tramite il software mindtct non corrispondono sempre alla loro posizione effettiva: (a) posizione originale; (b) posizione raffinata*

pacchetto NBIS, che fornisce però come uscita un solo valore che rappresenta il livello di somiglianza dei due pattern di minuzie. A questo scopo, però, effettua una ricerca delle minuzie corrispondenti nelle due immagini. È stato quindi necessario modificarne il codice sorgente in modo da ottenere in uscita la posizione delle minuzie considerate come corrispondenti tra di loro.

Il software si compone di quattro passi principali, come descritto in [38]:

1. **Calcolo delle tabelle di comparazione intra-fingerprint:** queste tabelle, una per ogni insieme di minuzie su cui si effettua il confronto, contengono la distanza e l'orientamento relativo tra le minuzie di un'impronta digitale. Ogni riga della tabella di comparazione corrisponde a un singolo arco che collega due minuzie;
2. **Calcolo di una tabella di compatibilità inter-fingerprint:** si confrontano le tabelle di comparazione relative a due impronte digitali e si considerano le coppie di righe che hanno distanza e orientamen-

to tra loro simili. Ogni riga della tabella di compatibilità mette in relazione un arco nella prima immagine con un arco nella seconda;

3. **Scansione della tabella di compatibilità:** si analizza la tabella di compatibilità in modo da collegare i singoli archi e determinare i possibili grafi di compatibilità. L'analisi viene iniziata da più punti di partenza: per ogni punto il software ricerca il percorso più lungo attraverso gli archi della tabella di compatibilità. Quando non è possibile andare oltre, l'analisi si interrompe e viene ripresa da un differente punto di partenza;
4. **Calcolo del match score:** si confrontano i grafi di compatibilità alla ricerca di quello con il maggior numero di archi. Il numero di archi del grafo più esteso costituisce il match score.

Modificando il codice sorgente dell'algoritmo, è stato quindi possibile ottenere dal software le righe della tabella di compatibilità analizzate durante la fase 3 dell'algoritmo, oltre al match score risultante dal confronto fra le immagini.

Nella Tabella 4.1 è mostrata una parte delle righe della tabella di compatibilità analizzate dal software. Nella porzione riportata, la dimensione temporanea del grafo aumenta fino a 67 archi corrispondenti tra le due immagini, dopodichè il software inizia a calcolare un nuovo grafo di compatibilità. Nel caso considerato, il software ha riportato un match score pari a 67, ciò significa che la miglior soluzione ottenuta consiste nella corrispondenza tra 64 archi e che il grafo considerato riporta le migliori corrispondenze stimate tra le minuzie delle due impronte digitali.

Tramite la conoscenza dell'indice della tabella di compatibilità è possibile quindi risalire alla posizione delle minuzie corrispondenti nelle due immagini. Le minuzie sono infine memorizzate come coppie di archi, così come nella tabella di compatibilità: a due minuzie nella prima immagine corrispondono due minuzie nella seconda immagine.

Dimensione temporanea del grafo	Indice della tabella di compatibilità
...	...
62	1033
63	1072
64	1073
65	1077
66	1036
67	1115
1	57
2	58
3	59
4	106
5	107
6	108
7	241
...	...

Tabella 4.1: *Esempio di righe della tabella di compatibilità analizzate dal software bozorth3*

Il risultato così ottenuto dimostra che il software bozorth3 non realizza un matching biunivoco, né esente da errori, tra le minuzie. I principali problemi riscontrati sono:

- la corrispondenza tra le minuzie non è sempre corretta; visivamente, infatti, non tutte le minuzie corrispondono;
- una coppia di minuzie in un'immagine può corrispondere a diverse coppie di minuzie nella seconda immagine, e viceversa;
- una minuzia in un'immagine può corrispondere a diverse minuzie nella seconda immagine, e viceversa.

4.2.4.2 Matching delle minuzie basato su bozorth3 e raffinamento stereoscopico

Per ottenere un matching biunivoco esatto tra le minuzie è necessario raffinare il risultato. Il processo di raffinamento proposto prende in esame una coppia di archi alla volta, e si suddivide in quattro passi:

1. **Eliminazione degli archi con minuzie vicino al bordo dell'area di riconoscimento:** questo passo è necessario in quanto il bordo dell'immagine difficilmente contiene minuzie reali, a causa della terminazione artificiale delle creste dovuta alla segmentazione ed alla scorretta frequenza delle creste dovuta ad un effetto prospettico legato alla forma del dito. Il processo di raffinamento proposto calcola la distanza di ogni minuzia dai punti del bordo del dito: gli archi le cui minuzie distano dal bordo meno di una soglia sono scartate;
2. **Eliminazione delle coppie di archi le cui minuzie sono situate in posizioni differenti nelle due dita:** per eliminare le minuzie che non si trovano nella stessa regione del dito, si confrontano le coordinate di una minuzia con le coordinate della minuzia corrispondente. La differenza tra le coordinate deve essere inferiore ad una soglia fissata, in caso contrario gli archi contenenti le minuzie vengono scartati.:

$$\begin{cases} L_i \in S_L & \text{se } d(L_{i,A}, R_{i,A}) \leq s \quad \text{and } d(L_{i,B}, R_{i,B}) \leq s_1 \\ L_i \notin S_L & \text{se } d(L_{i,A}, R_{i,A}) \geq s \quad \text{or } d(L_{i,B}, R_{i,B}) \geq s_1 \end{cases} ; \quad (4.24)$$

$$\begin{cases} R_i \in S_R & \text{se } d(R_{i,A}, L_{i,A}) \leq s_1 \quad \text{and } d(R_{i,B}, L_{i,B}) \leq s_1 \\ R_i \notin S_R & \text{se } d(R_{i,A}, L_{i,A}) \geq s_1 \quad \text{or } d(R_{i,B}, L_{i,B}) \geq s_1 \end{cases} , \quad (4.25)$$

dove S_L e S_R sono gli insiemi di archi dell'immagine sinistra e dell'immagine destra, d indica la distanza tra due punti, s_1 è un valore di soglia, $L_{i,A}$ e $L_{i,B}$ sono le due minuzie dell'arco i -esimo dell'immagine sinistra e $R_{i,A}$ e $R_{i,B}$ sono le due minuzie dell'arco i -esimo dell'immagine destra;

3. **Eliminazione delle corrispondenze non univoche tra archi:** questo passo fa in modo che ad un arco in un'immagine non possano corrispondere più archi nella seconda immagine, e viceversa. L'eliminazione degli archi duplicati è realizzata ricercando tutti gli archi differenti le cui minuzie hanno le stesse coordinate. Viene quindi mantenuto uno solo di questi archi;
4. **Eliminazione delle minuzie ripetute:** questo passo fa in modo che a minuzie uguali in un'immagine corrispondano minuzie uguali nella seconda immagine, e viceversa. Si ricercano nell'immagine minuzie appartenenti ad archi diversi con posizione uguale tra di loro. Si controlla quindi che le minuzie corrispondenti nella seconda immagine siano a loro volta uguali tra di loro.

$$\begin{aligned} L_{i,A} = L_{j,A} &\Rightarrow R_{i,A} = R_{j,A} \\ L_{i,B} = L_{j,B} &\Rightarrow R_{i,B} = R_{j,B} \end{aligned} \quad (i \neq j) \quad . \quad (4.26)$$

In Figura 4.16 sono mostrate le corrispondenze tra minuzie calcolate dal software bozorth3 e le corrispondenze tra minuzie dopo il processo di raffinamento.

4.2.4.3 Matching delle minuzie basato sulla vicinanza

Viene proposto inoltre un algoritmo di matching che ricerca nelle due immagini le minuzie con posizione simile. L'algoritmo confronta ogni minuzia nella prima immagine con tutte le minuzie nella seconda immagine: la corrispondenza tra le minuzie viene stabilita solo nel caso in cui i valori relativi di posizione, orientamento e qualità siano inferiori alle rispettive soglie fissate:

$$\begin{aligned} \{ML_i(x_l, y_l, \theta_l, Q_l), MR_i(x_r, y_r, \theta_r, Q_r)\} \in S_C \quad &\text{se } |x_l - x_r| \leq s_2 \\ &\text{and } |y_l - y_r| \leq s_3 \\ &\text{and } |\theta_l - \theta_r| \leq s_4 \\ &\text{and } |Q_l - Q_r| \leq s_5 \quad , \quad (4.27) \end{aligned}$$

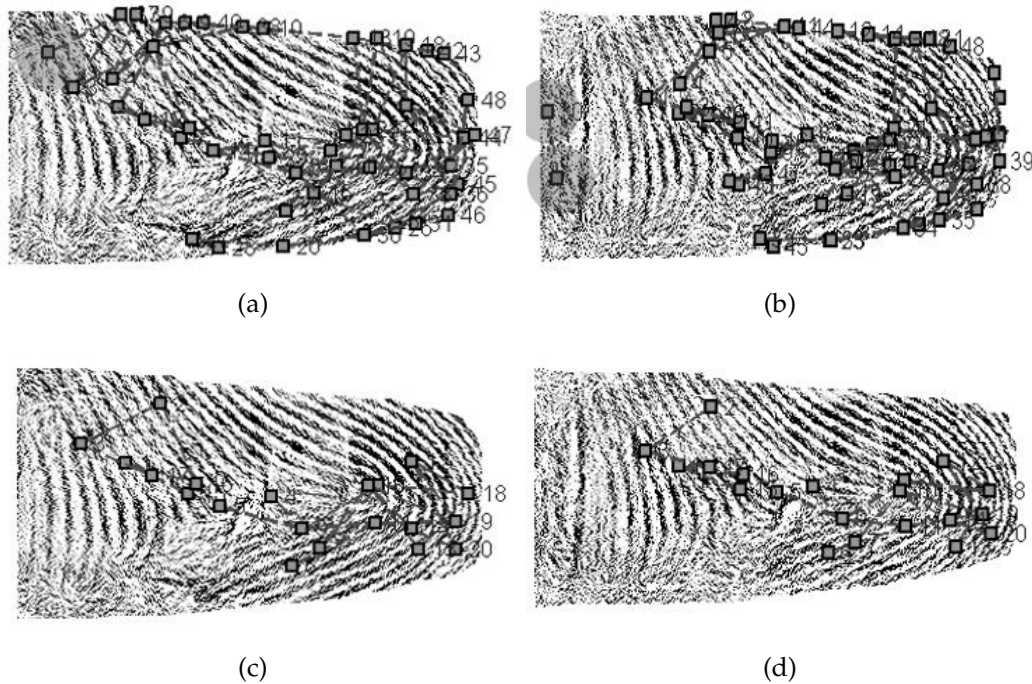


Figura 4.16: *Matching delle minuzie e raffinamento dei risultati: (a) minuzie nell'immagine sinistra; (b) minuzie nell'immagine destra; (c) minuzie raffinate nell'immagine sinistra; (d) minuzie raffinate nell'immagine destra;*

dove $ML_i(x_l, y_l, \theta_l, Q_l)$ e $MR_i(x_r, y_r, \theta_r, Q_r)$ sono rispettivamente la posizione, l'orientamento e la qualità della minuzia i -esima nell'immagine sinistra e nell'immagine destra, s_2, s_3, s_4, s_5 sono valori di soglia e S_C è l'insieme delle coppie di minuzie corrispondenti.

Quest'algoritmo è utilizzato come misura di confronto nella ricerca delle corrispondenze stereoscopiche. Ha però due difetti principali: in primo luogo, non è invariante alla rotazione delle immagini. È possibile, infatti, compensare la traslazione della posizione delle minuzie in un'immagine rispetto all'altra; non è possibile compensare globalmente, però, gli effetti della rotazione: ciascuna minuzia risulta spostata in modo differente a seconda della sua posizione nell'immagine.

Secondariamente, è poco accurato nelle zone dell'immagine che presen-

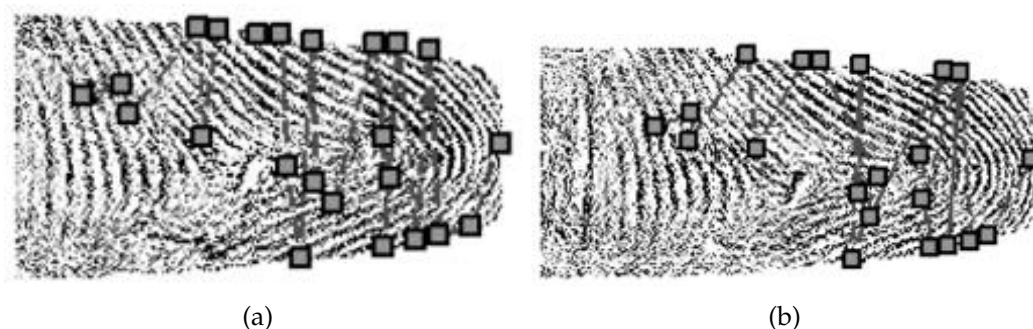


Figura 4.17: *Matching delle minuzie basato sulla vicinanza: (a) minuzie nell'immagine sinistra; (b) minuzie nell'immagine destra*

tano minuzie ravvicinate. Non impiegando tecniche di analisi tra coppie di minuzie, il metodo non è in grado di trovare la corrispondenza esatta nel caso in cui sia presente più di una minuzia all'interno della zona circoscritta dal valore di soglia.

In Figura 4.17 è mostrato un esempio di matching utilizzando l'algoritmo proposto.

4.2.5 Triangolazione del bordo del dito e delle minuzie

Il modello tridimensionale risultante dall'estrazione del bordo del dito e delle minuzie viene ottenuto calcolando per prima cosa la corrispondenza stereoscopica di un punto appartenente al contorno. Il metodo descritto usa la punta del dito poichè è facilmente individuabile nelle due immagini.

Conoscendo le coordinate x_l e x_r del punto scelto rispettivamente nell'immagine sinistra e nell'immagine destra, la sua profondità z nello spazio tridimensionale è calcolata secondo la formula di triangolazione 2.19:

$$z = \frac{fT}{x_l - x_r} \quad , \quad (4.28)$$

dove f è la lunghezza focale e T è la traslazione tra le due fotocamere. Il valore di profondità ottenuto è quindi assegnato a tutti i punti del bordo. L'operazione di triangolazione è quindi effettuata sulle minuzie corrispondenti, come mostrato in Figura 4.18a.

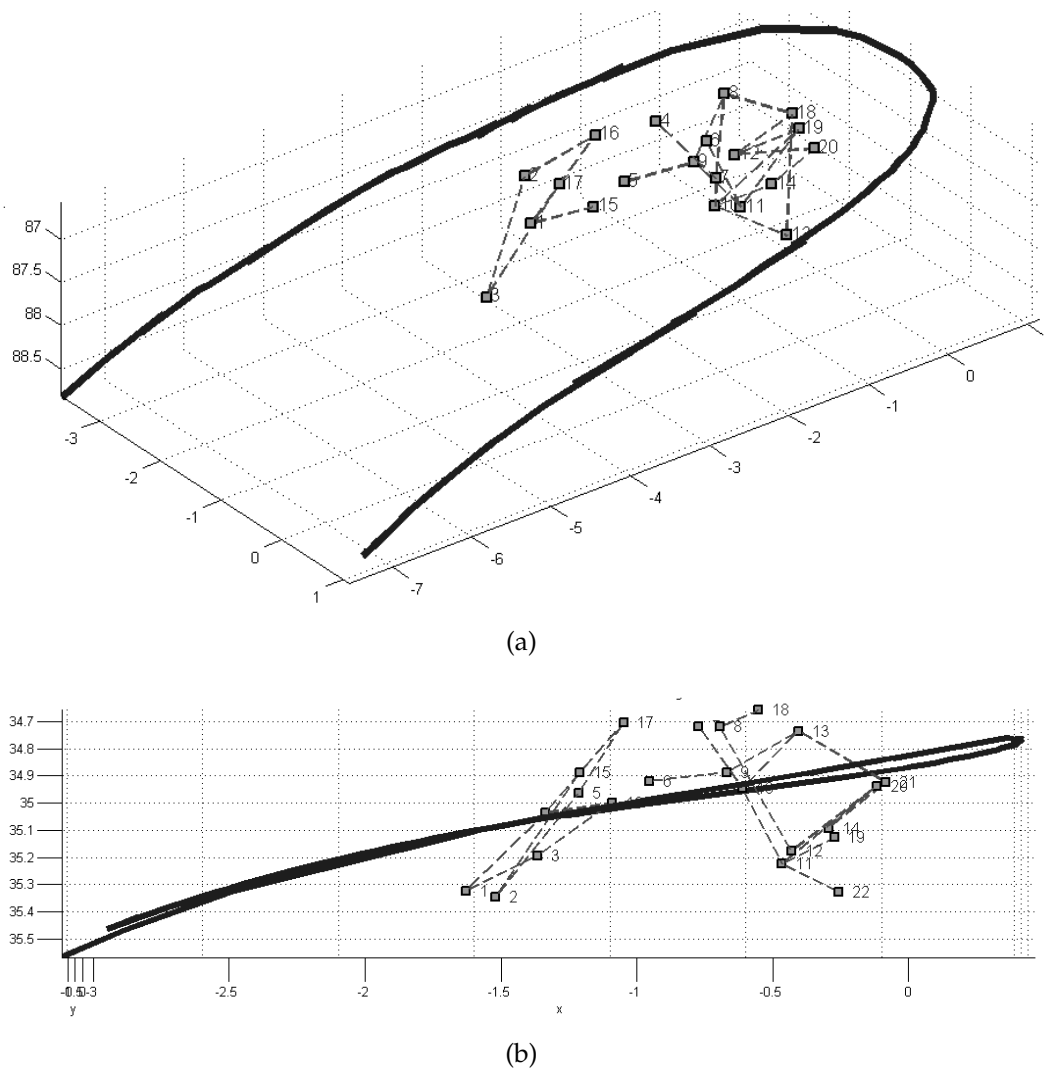


Figura 4.18: *Triangolazione del bordo del dito e delle minuzie corrispondenti: (a) vista dall'alto; (b) vista laterale*

Come si può osservare nell'immagine 4.18b, la profondità delle minuzie nello spazio tridimensionale non rispecchia la realtà con approssimazione accettabile. Le minuzie, infatti, risultano triangolate secondo delle posizioni che non riflettono la curvatura del dito.

Il calcolo della profondità delle minuzie nello spazio tridimensionale, infatti, necessita di estrema accuratezza nella stima della posizione delle minuzie. Il software di estrazione delle minuzie, però, non ricerca la

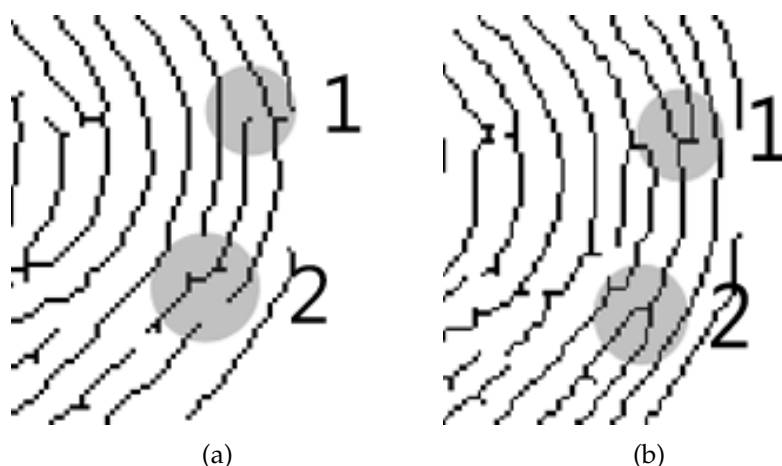


Figura 4.19: *La posizione ed il tipo delle minuzie risultano modificate tra le due immagini: (a) due minuzie nell'immagine sinistra; (b) le due minuzie corrispondenti nell'immagine destra*

posizione delle minuzie con una precisione sufficientemente elevata ai fini della triangolazione. Una variazione della posizione della minuzia di pochi pixel, infatti, può essere influente ai fini del calcolo del grado di somiglianza tra le due impronte digitali, ma i test hanno mostrato che risulta determinante ai fini del calcolo della posizione delle stesse nello spazio tridimensionale.

Una soluzione possibile è rappresentata dall'utilizzo dell'immagine sottoposta all'operazione morfologica di skeleton per determinare la posizione delle minuzie con maggiore precisione. Il problema di questo approccio, però, consiste nel fatto che gli algoritmi di preprocessing elaborano in modo diverso le due immagini stereoscopiche. Le minuzie presenti nell'impronta digitale, infatti, risultano spesso modificate all'interno delle due viste.

In Figura 4.19 è mostrata la stessa regione dell'impronta digitale nelle due immagini stereoscopiche. Nell'immagine 4.19a le due minuzie mostrate sono di tipo terminazione, mentre nell'immagine 4.19b le minuzie corrispondenti sono di tipo biforcazione.

4.3 Ricostruzione tridimensionale basata su modello cilindrico

Il secondo metodo proposto per il calcolo di un modello tridimensionale del dito e dell'impronta digitale è basato sul calcolo di un cilindro che approssimi la forma reale del dito. Il volume così ottenuto è quindi completato tramite la sovrapposizione dei dettagli dell'impronta digitale.

L'algoritmo proposto è schematizzato in Figura 4.20. Per prima cosa, viene stimata l'altezza del dito rispetto al bordo; questo passo richiede il calcolo della corrispondenza stereoscopica di un punto di riferimento all'interno dell'impronta digitale. Il metodo proposto assume che tale punto coincida con il punto di massima curvatura delle creste. Nella maggior parte dei casi, inoltre, tale punto è collocato nella zona centrale dell'impronta digitale.

Successivamente, viene calcolata la larghezza del dito, come descritto in 4.2.2, e la profondità nello spazio tridimensionale del bordo del dito, come descritto in 4.2.5. In questo modo, risulta possibile ottenere due punti tridimensionali, distanziati di un valore pari alla larghezza del dito e collocati alla profondità del bordo del dito.

Il punto di massima curvatura e i due punti calcolati contengono le informazioni relative alla larghezza e all'altezza del dito nello spazio tridimensionale. Viene quindi calcolato il polinomio del secondo ordine passante per i tre punti. La circonferenza così ottenuta viene successivamente applicata a tutta la lunghezza del dito, al fine di ottenere un cilindro.

Il modello cilindrico così creato è completato applicando su di esso i dettagli dell'impronta digitale, secondo una procedura denominata "wrapping". Il metodo di ricostruzione proposto descrive il wrapping sia delle minuzie che delle immagini stesse.

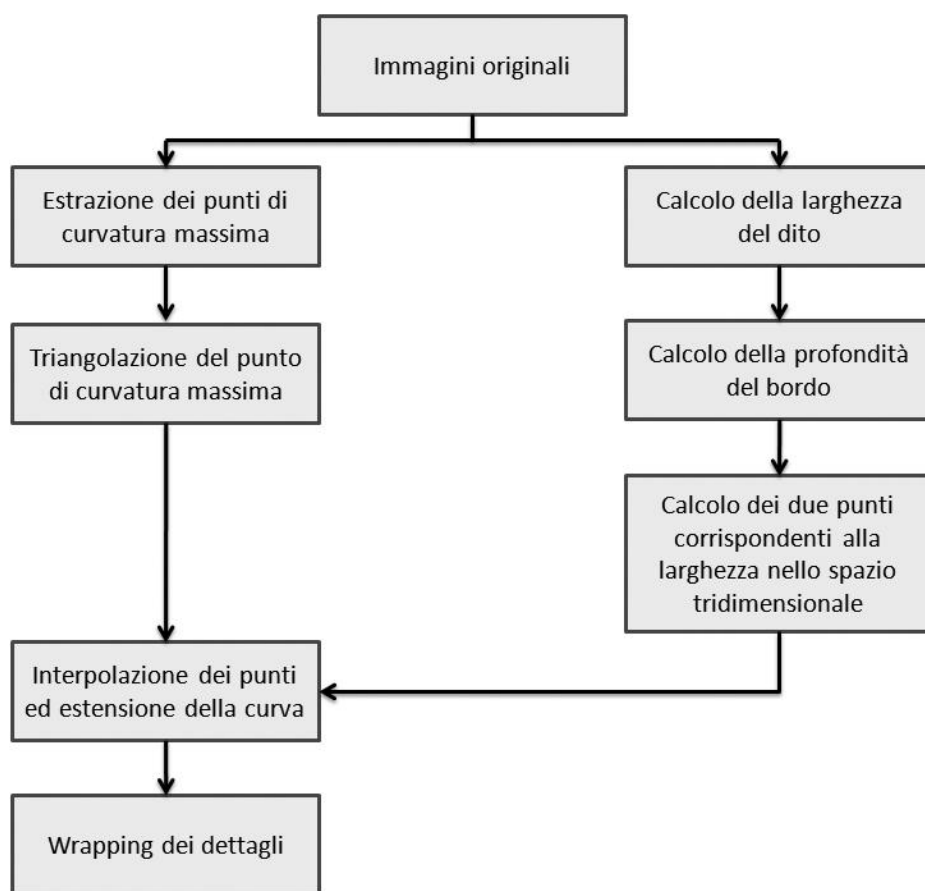


Figura 4.20: Schema a blocchi dell'algoritmo per il calcolo di un modello tridimensionale cilindrico

4.3.1 Calcolo del punto di massima curvatura

La funzione per la stima degli orientamenti delle creste, riportata in [39], permette di calcolare, con una certa approssimazione, la direzione delle creste dell'impronta digitale. Il risultato ottenuto è una mappa degli orientamenti in cui ogni pixel ha un valore di intensità proporzionale all'angolo formato dalla direzione della cresta rispetto all'asse orizzontale.

Il risultato del calcolo dell'orientamento dei ridge è mostrato in Figura 4.21. Nell'immagine (a) si può osservare come l'orientamento dei ridge cambi a seconda della zona dell'immagine.

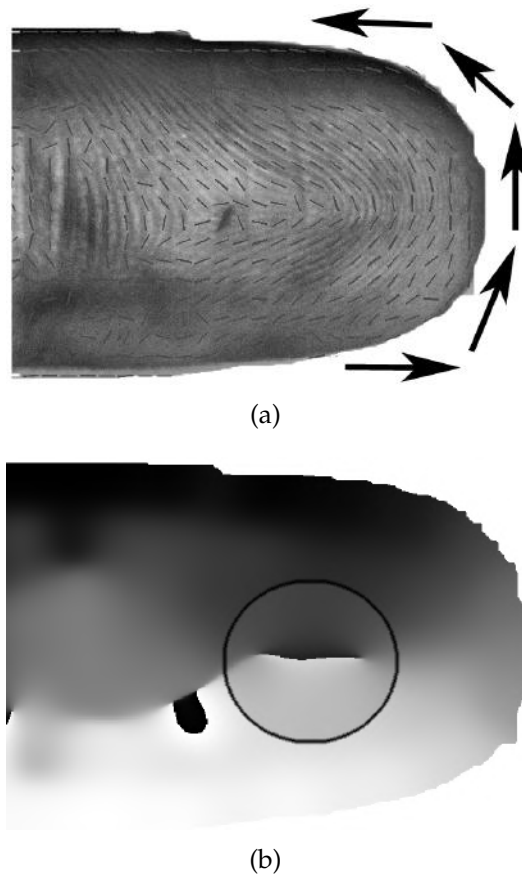


Figura 4.21: *Mappa degli orientamenti delle creste: (a) variazione dell'orientamento delle creste; (b) zona di massima curvatura*

Nell'immagine (b), ottenuta mappando il livello di intensità dei pixel sulla scala di grigi, si nota come esista una zona centrale in cui l'orientamento dei ridge varia rapidamente da un valore prossimo allo 0 (bianco) ad un valore vicino a π (nero). L'estremità di questa zona più vicina alla punta del dito corrisponde alla posizione del punto di massima curvatura dell'impronta digitale.

Poichè tale punto è definito come la zona della mappa degli orientamenti con variazione più repentina, al fine di estrarne la posizione, il metodo proposto calcola il gradiente della mappa degli orientamenti ed estrae le zone con il valore massimo. Per evitare che il calcolo del gra-

diente sia influenzato da valori elevati in prossimità del bordo del dito, la maschera di segmentazione viene precedentemente erosa in modo da eliminare dall'immagine le regioni periferiche del dito.

L'algoritmo proposto per il calcolo del punto di massima curvatura è riassumibile nei seguenti passi:

1. calcolo del gradiente verticale della mappa degli orientamenti;
2. erosione della maschera di segmentazione con elemento strutturante di dimensione pari a circa $1/10$ della larghezza del dito: in questo modo si eliminano dall'immagine solo le regioni prossime al bordo del dito;
3. estrazione della zona dell'immagine con gradiente superiore al 50% del valore massimo;
4. elaborazione con operatore morfologico di chiusura, utilizzato con elemento strutturante orizzontale in modo da unire le zone della regione con curvatura massima;
5. calcolo dei componenti connessi per mantenere solo il componente più esteso;
6. estrazione, dalla regione calcolata, della posizione del punto più vicino alla punta del dito.

In Figura 4.22 è mostrato il risultato del calcolo della posizione del punto di massima curvatura.

4.3.2 Calcolo del modello cilindrico

Il metodo proposto per il calcolo del modello tridimensionale cilindrico è sviluppato a partire da tre punti nello spazio tridimensionale: il primo punto è ottenuto tramite la triangolazione del punto di massima curvatura, calcolato come descritto in 4.3.1. Gli altri due punti sono stimati basandosi sulla conoscenza della larghezza del dito, calcolata come descritto in

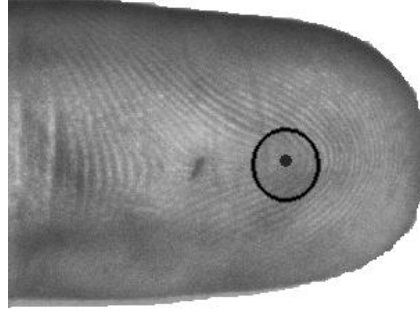


Figura 4.22: Posizione del punto di massima curvatura

4.2.2, e del valore di profondità del bordo, ottenuto come descritto in 4.2.5. Questi punti sono infatti posizionati nello spazio tridimensionale secondo tale profondità, e distanziati in base alla larghezza del dito.

Successivamente, viene calcolato il polinomio del secondo ordine passante per i tre punti, mostrati nella Figura 4.23a, tramite funzioni *spline* cubiche. Le funzioni *spline* cubiche sono utilizzate separatamente per interpolare i valori delle coordinate (x, y, z) , secondo la formula:

$$S_i(x) = \frac{z_i(x - t_{i-1})^3}{6h_i} + \frac{z_{i-1}(t_i - x)^3}{6h_i} + \left[\frac{f(t_i)}{h_i} - \frac{z_i h_i}{6} \right] (x - t_{i-1}) + \left[\frac{f(t_{i-1})}{h_i} - \frac{z_{i-1} h_i}{6} \right] (t_i - x) \quad , \quad (4.29)$$

dove $S_i(x)$ è la funzione interpolante risultante, $f(t_i)$ sono i punti da interpolare, $z_i = f''(t_i)$ e $h_i = t_i - t_{i-1}$.

Tramite l'uso di una funzione di interpolazione di circonferenza, descritta in [40], è possibile trovare il centro della circonferenza passante per i punti interpolati, come mostrato in Figura 4.23b. Il calcolo del centro della circonferenza permette di esprimere la curva C secondo l'equazione parametrica:

$$C : \begin{cases} x = x_0 + r \cos(t) \\ y = y_0 + r \sin(t) \end{cases} \quad t \in [0, 2\pi] \quad , \quad (4.30)$$

dove (x_0, y_0) e R sono rispettivamente le coordinate del centro ed il raggio della circonferenza. La circonferenza è quindi applicata a tutta la lunghezza del dito. Poichè è computazionalmente costoso visualizzare in tre dimensioni una nuvola di punti di grandi dimensioni, per una maggiore efficienza di calcolo si utilizza una funzione di interpolazione di superficie proposta in [41]. Il metodo di interpolazione si basa sulla creazione di una mesh di triangoli a partire dalla nuvola di punti; successivamente viene applicata una funzione di interpolazione lineare all'interno di ogni triangolo. In Figura 4.23d è mostrato il risultato dell'interpolazione.

4.3.3 Wrapping dei dettagli sul modello cilindrico

Dopo aver calcolato il modello cilindrico del volume del dito, è necessario sovrapporre ad esso i dettagli dell'impronta digitale. Questo procedimento, definito *wrapping*, può essere eseguito in diversi modi. Un primo metodo consiste nell'aggiungere dettagli precedentemente calcolati, come le minuzie oppure il pattern di creste, estratti secondo il metodo descritto in 4.2.4.

Il problema principale dell'utilizzo delle minuzie consiste nel fatto che i metodi di elaborazione delle immagini necessari alla ricostruzione basata sulle minuzie, descritti in 4.2, risultano poco accurati nelle regioni dell'immagine corrispondenti a quelle di massima curvatura del dito.

Una soluzione proposta consiste nell'applicare al modello cilindrico direttamente le immagini originali acquisite senza contatto. Effettuando un mapping tra l'immagine del dito ed il modello cilindrico, le regioni dell'immagine corrispondenti alla massima curvatura del dito vengono fatte corrispondere alla curvatura del modello cilindrico. In questo modo sono ripristinate le distanze originali tra i punti, come mostrato in Figura 4.24.

La tecnica di wrapping proposta consiste nell'effettuare la proiezione dei dati sul modello del volume, in modo inverso a quanto succede quando l'oggetto reale viene proiettato sul piano dell'immagine. Il metodo proposto elabora quindi la posizione dei dati confrontandola con il modello

cilindrico. A seconda della posizione in cui il dato si trova, viene assegnato ad esso un valore di profondità in modo da corrispondere alla superficie del modello.

Poichè non è necessaria l'operazione di triangolazione per determinare la profondità dei punti, i dettagli possono essere estratti da una singola immagine.

L'operazione di wrapping è quindi effettuata analizzando ogni punto bidimensionale appartenente alla nuvola di punti dei dettagli. Per ognuno di questi punti viene ricercato nel modello cilindrico il punto con coordinate planari corrispondenti. Il valore di profondità di quest'ultimo punto viene quindi assegnato al punto bidimensionale, in modo da trasformarlo in un punto tridimensionale e posizionarlo in corrispondenza della superficie del modello cilindrico.

Poichè il modello cilindrico è ottenuto estendendo una curva secondo la lunghezza del dito, come descritto in 4.3.2, è sufficiente condurre un'analisi monodimensionale per calcolare il valore di profondità del punto di cui effettuare il wrapping. Nella situazione illustrata in Figura 4.25, è necessario determinare il valore della coordinata z da assegnare ai punti: tale valore non è influenzato dalla posizione del punto lungo l'asse x in quanto il modello non cambia lungo tale asse. È quindi sufficiente analizzare la coordinata y del punto per ottenerne il corretto valore lungo l'asse z .

Il procedimento di wrapping proposto è quindi suddivisibile in due passi:

1. confronto della coordinata y della nuvola di punti bidimensionale dei dettagli con la coordinata y dei punti del modello cilindrico;
2. assegnamento ai punti bidimensionali del valore z corrispondente al punto con coordinata y uguale.

In Figura 4.26 sono mostrati un esempio di wrapping delle minuzie e di wrapping delle creste sul modello cilindrico.

Il procedimento di wrapping delle immagini originali sul modello cilindrico, invece, calcola la nuvola di punti bidimensionale a partire dai pixel

dell'immagine. Ogni punto è quindi descritto da una terna $(x, y, I(x, y))$, dove $I(x, y)$ rappresenta il valore di intensità del pixel alla posizione (x, y) . Il wrapping dei punti viene quindi effettuato secondo il metodo già descritto, producendo una rappresentazione pseudo-quadrimensionale, in cui ogni punto è descritto dalla t -upla $(x, y, z, I(x, y, z))$.

Inoltre, poichè la larghezza del modello cilindrico è calcolata sulla base della larghezza del dito in prossimità dell'ultima giunzione, calcolata come descritto in 4.2.2, le zone dell'immagine eccedenti tale larghezza sono escluse dalla ricostruzione, come mostrato in Figura 4.27. Una rappresentazione visuale del risultato è mostrata in Figura 4.28.

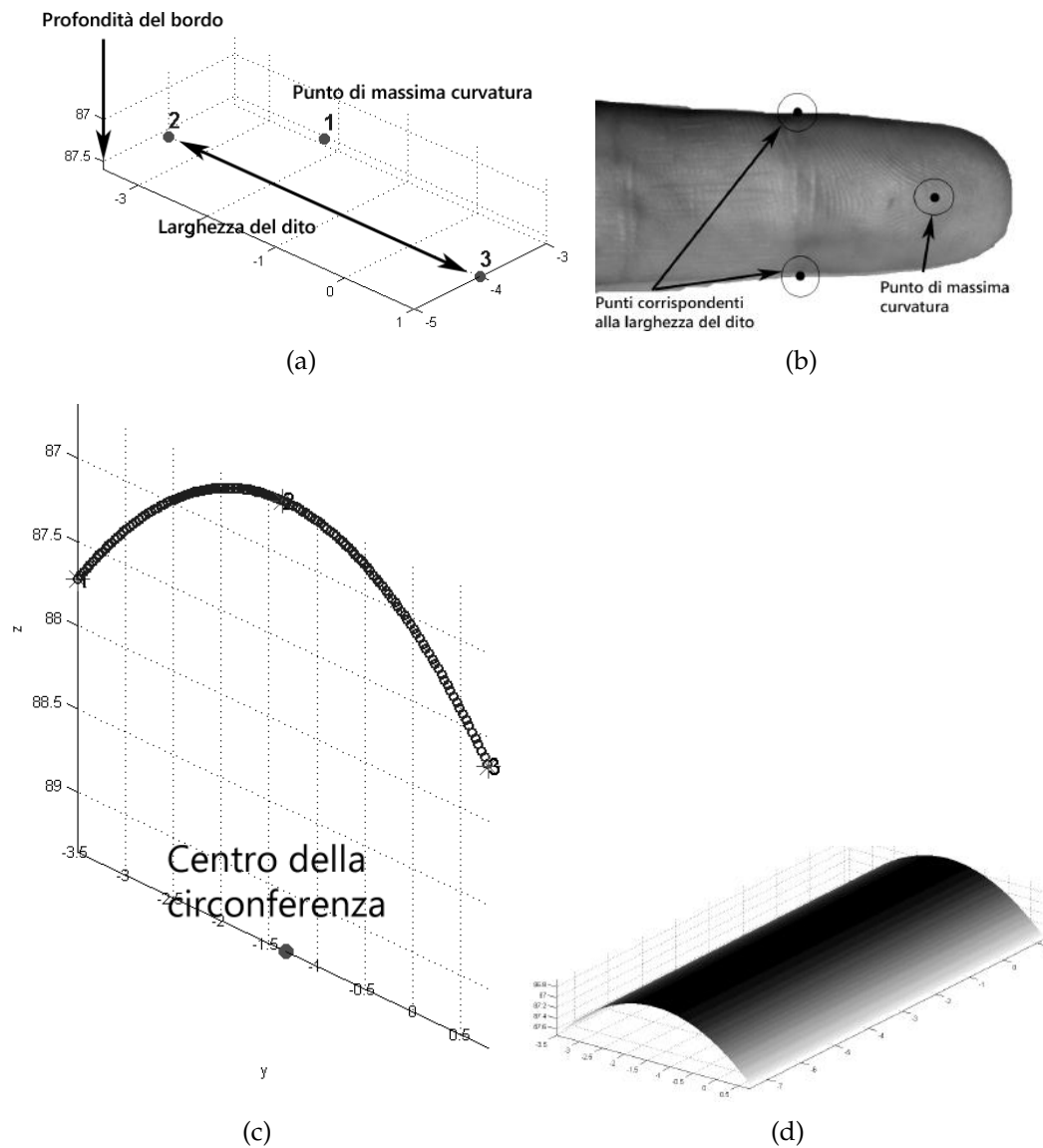


Figura 4.23: Calcolo del modello cilindrico del dito: (a) i tre punti principali; (b) la posizione dei tre punti nel dito; (c) interpolazione con spline e centro della circonferenza interpolante; (d) interpolazione della superficie del modello esteso

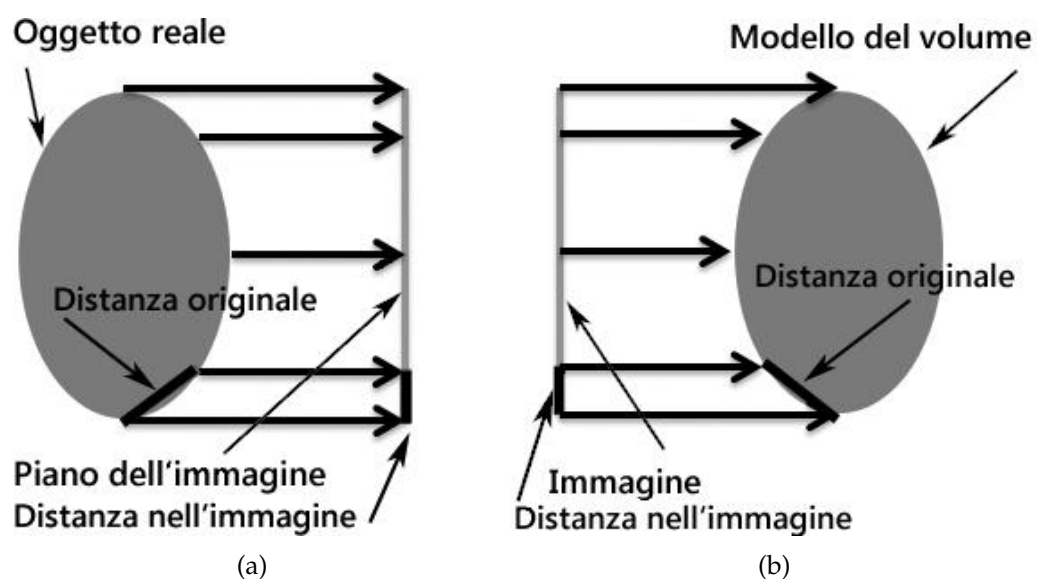


Figura 4.24: Spiegazione del procedimento di wrapping: (a) proiezione dell'oggetto sull'immagine; (b) proiezione dell'immagine sul modello

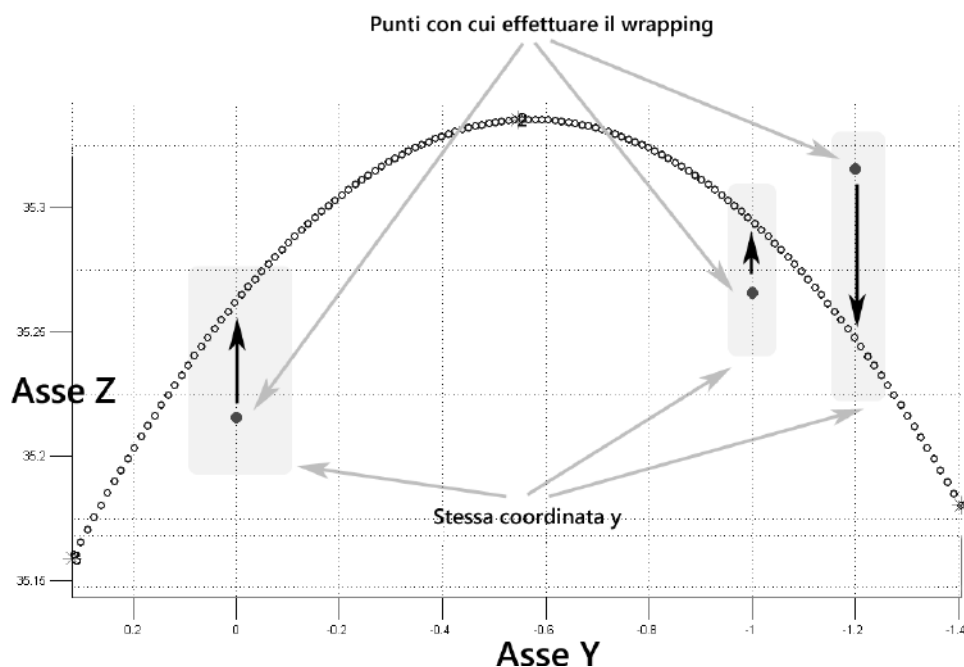
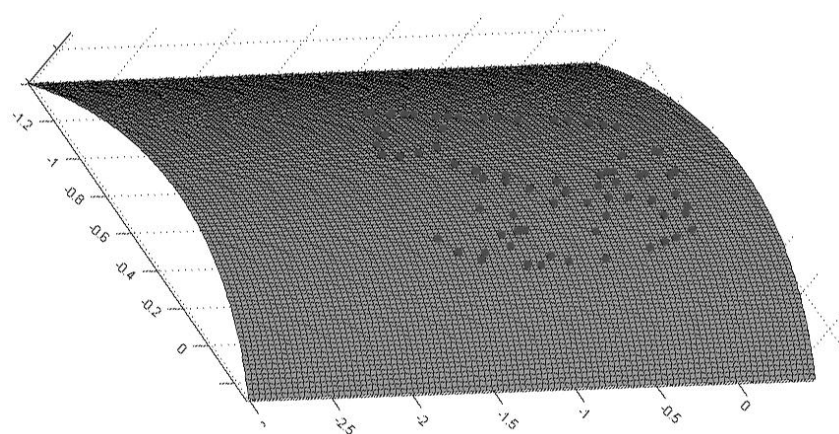
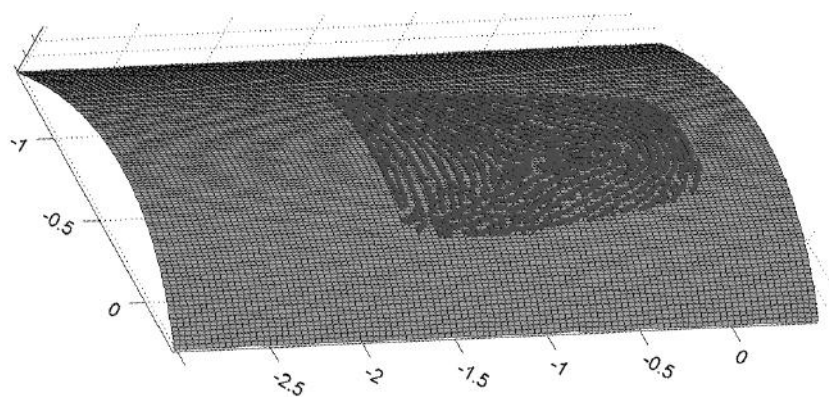


Figura 4.25: Tecnica di wrapping tramite analisi monodimensionale



(a)



(b)

Figura 4.26: Wrapping dei dettagli sul modello cilindrico: (a) wrapping delle minuzie; (b) wrapping del pattern di creste

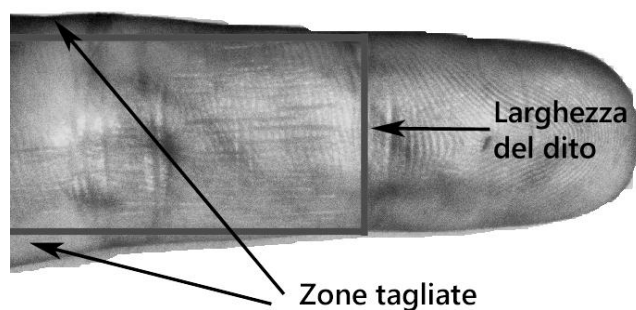


Figura 4.27: Zone dell'immagine escluse dall'operazione di wrapping

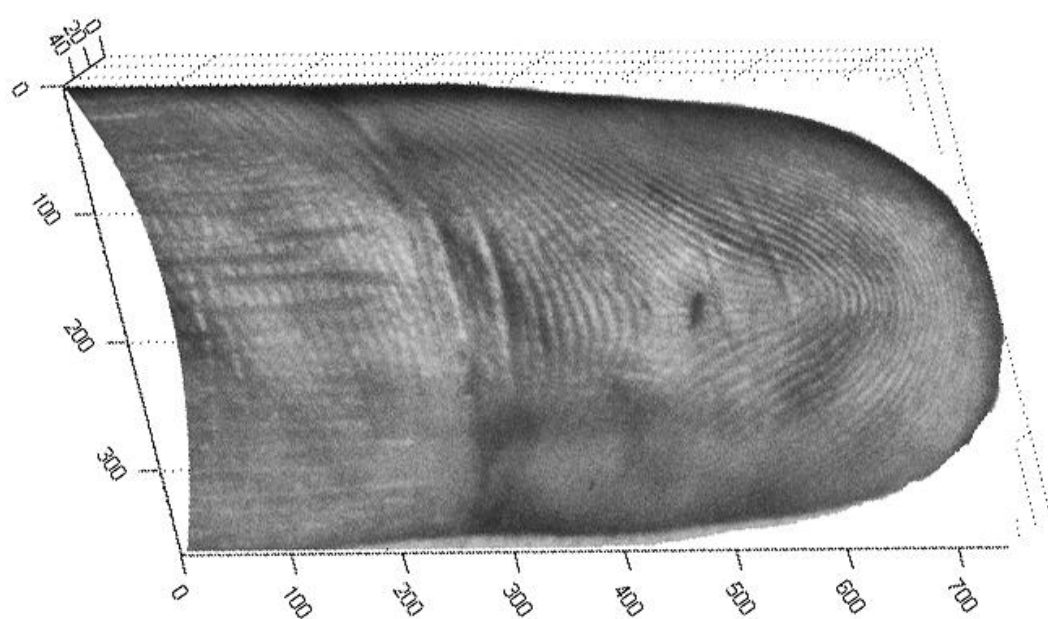


Figura 4.28: *Esempio di risultato dell'operazione di wrapping dell'immagine sul modello cilindrico*

4.4 Unwrapping del modello tridimensionale cilindrico

Il procedimento di unwrapping è la tecnica mediante la quale una superficie tridimensionale viene mappata su un piano bidimensionale. Questa procedura ha il vantaggio di non produrre immagini influenzate dalla curvatura del modello tridimensionale di partenza. Nell'immagine risultante, infatti, le zone periferiche possiedono il medesimo grado di dettaglio delle zone centrali.

Il metodo di unwrapping del modello cilindrico proposto consiste in due fasi:

1. conversione in coordinate polari;
2. trasformazione del modello polare in nuove coordinate cartesiane.

Il valore della fase θ di una circonferenza in coordinate polari è lineare. È quindi possibile sostituire le due coordinate cartesiane (x, y) con una sola coordinata θ . Analogamente, poichè la circonferenza interpolante del modello cilindrico è disposta lungo gli assi (Y, Z) , le tre coordinate cartesiane (x, y, z) di una superficie cilindrica possono essere espresse tramite due sole coordinate (x, θ) .

Il procedimento di unwrapping delle immagini proposto consiste nel rimappare la t -upla di valori $(x, y, z, I(x, y, z))$, ottenuta dopo la fase di wrapping, in una nuova immagine bidimensionale $(x, \theta, I(x, \theta))$ in cui le coordinate y sono sostituite di valori della fase.

Durante la fase di wrapping e la successiva fase di unwrapping, quindi, viene effettuato un mapping tra le immagini, in particolare viene tenuta traccia della corrispondenza tra la posizione dei pixel nell'immagine originale e quella nell'immagine risultante dalla fase di unwrapping, passando per la posizione nello spazio tridimensionale:

$$(x, y) \rightarrow (x, y, z) \rightarrow (x, \theta) \quad . \quad (4.31)$$

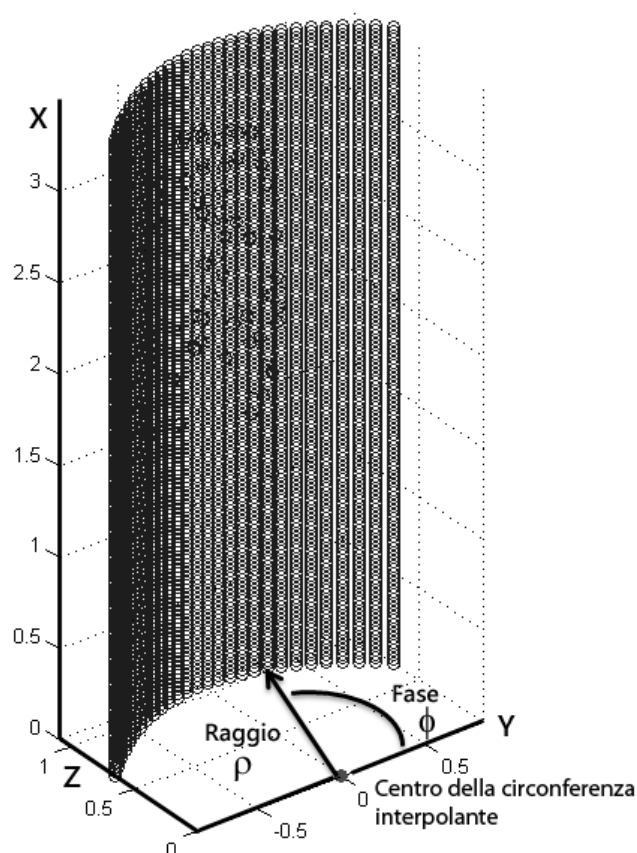


Figura 4.29: *Calcolo delle coordinate polari del modello cilindrico*

In Figura 4.30 è mostrato il risultato dell'unwrapping di un modello cilindrico dopo la fase di wrapping delle minuzie: come si può osservare, tutti i punti risultano disposti su un unico piano. Inoltre, le distanze tra i punti rimangono costanti anche nelle zone periferiche dell'impronta digitale.

Le coordinate θ risultanti dalla fase di unwrapping hanno un range di valori più ampio delle coordinate y originali e quindi l'immagine risultante ha una larghezza maggiore. Il passo di campionamento delle coordinate spaziali non risulta inoltre costante nell'immagine finale. Per questo motivo, sono visualmente riscontrabili aree di valore indefinito, come mostrato in Figura 4.31a.

La soluzione proposta a questo problema sfrutta l'interpolazione dei valori mancanti. Dal momento che è necessario garantire solamente la

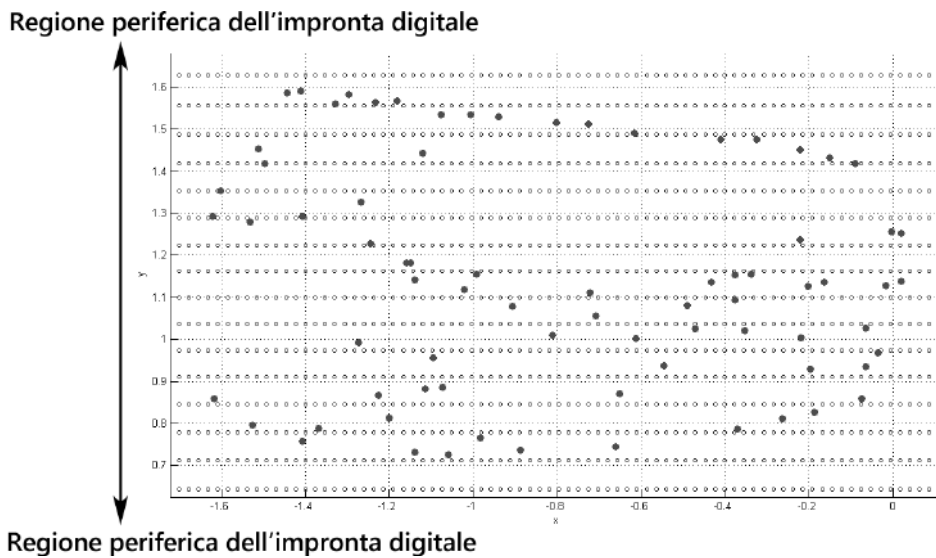
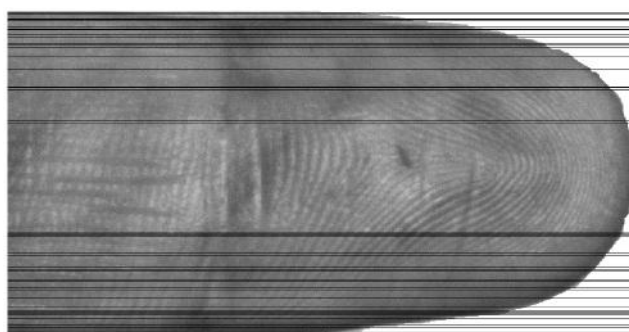


Figura 4.30: *Modello cilindrico delle minuzie dopo la fase di unwrapping*

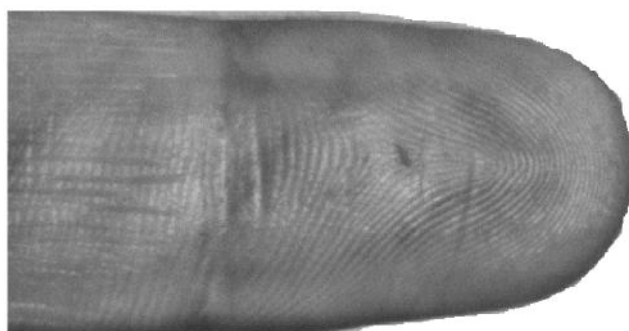
continuità delle creste nell'immagine, il metodo di interpolazione utilizzato si basa sulla funzione *nearest neighbor*, che assegna ad ogni riga di valore indefinito dell'immagine i valori di intensità dei pixel della riga vicina. Il risultato dell'interpolazione è mostrato in Figura 4.31b.

Le immagini ottenute tramite unwrapping del modello tridimensionale hanno la proprietà di rappresentare il modello tridimensionale del dito tramite un'immagine bidimensionale, senza introdurre deformazioni e senza possedere i difetti derivanti dalla mappatura bidimensionale delle zone prossime alle estremità laterali del dito, caratterizzate da un'elevata curvatura. Per questo motivo, tali immagini risultano più adeguate al confronto con le immagini ottenute da sensori tradizionali rispetto ad immagini acquisite senza contatto tradizionali. Appoggiando il dito sul sensore, infatti, la sua superficie tende ad appiattirsi in modo simile a quanto simulato tramite il metodo di unwrapping.

Risulta quindi possibile applicare alle immagini ottenute metodi di estrazione delle caratteristiche biometriche, come il metodo descritto in 4.2.4, al fine di effettuare il riconoscimento dell'identità di un individuo.



(a)



(b)

Figura 4.31: *Unwrapping di immagini applicate al modello cilindrico del dito: (a) l'immagine prima dell'interpolazione; (b) l'immagine dopo l'interpolazione*

4.5 Calcolo di un modello tridimensionale sintetico del dito

In questa sezione è proposto un metodo per il calcolo di un modello tridimensionale sintetico del dito a partire da immagini di impronte digitali acquisite tramite contatto. Il calcolo di un modello tridimensionale virtuale permette di ottenere immagini simili a quelle prodotte da acquisizioni stereoscopiche di impronte digitali senza disporre di attrezzature di laboratorio quali fotocamere, ecc. Inoltre, dal momento che è semplice ottenere dataset di impronte digitali acquisite tramite contatto, permette di calcolare velocemente un'insieme di immagini di prova. La realizzazione di un dataset di grandi dimensioni di immagini stereoscopiche acquisite senza contatto, infatti, richiede una quantità di tempo e di persone coinvolte molto superiori.

Il procedimento proposto di calcolo di un dataset sintetico si compone di quattro fasi:

1. calcolo del modello tridimensionale sintetico;
2. definizione del setup ambientale relativo alle immagini stereoscopiche;
3. calcolo del modello tridimensionale del pattern di calibrazione;
4. calcolo delle immagini di calibrazione.

4.5.1 Calcolo del modello sintetico tridimensionale

Il metodo proposto per il calcolo di un modello tridimensionale sintetico è schematizzato in Figura 4.32 e consiste nell'elaborazione delle immagini tradizionali tramite una tecnica simile al metodo di ricostruzione tridimensionale basato su modello cilindrico, descritto in 4.3.

Per prima cosa, infatti, è necessario calcolare un modello cilindrico della superficie a partire da un'immagine: tale modello ha caratteristiche standard di larghezza e di altezza. In particolare, poichè durante un'acqui-

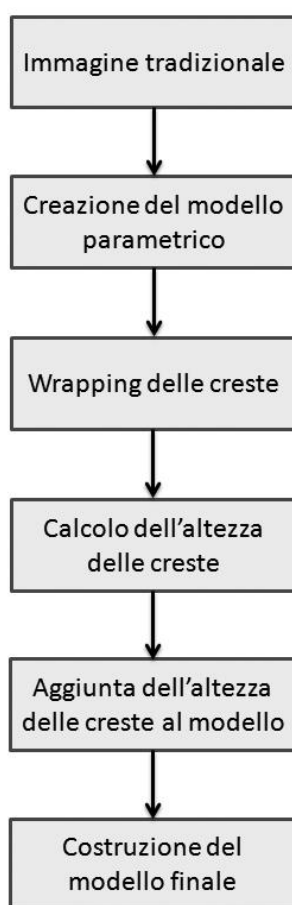


Figura 4.32: *Schema a blocchi del metodo di calcolo di un modello tridimensionale sintetico*

sizione tramite sensori tradizionali non tutto il dermatoglifo è a contatto con il sensore, la larghezza del modello cilindrico è calcolata estendendo in modo proporzionale la larghezza dell'impronta digitale.

Successivamente, il modello è completato con l'aggiunta del pattern delle creste estratte dall'immagine binarizzata. Dopo aver effettuato il wrapping delle creste sul modello cilindrico, è necessario modificare tale modello al fine di aggiungervi l'altezza delle creste rispetto alla superficie.

Per modificare il modello in modo che le creste risultino sempre perpendicolari alla superficie cilindrica del modello, si interviene sul valore

del raggio delle coordinate polari. Le coordinate cartesiane della nuvola di punti delle creste sono quindi convertite in coordinate polari. Poichè il modello non varia lungo l'asse x , è sufficiente considerare le coordinate (y, z) :

$$\begin{aligned}\rho &= \sqrt{y^2 + z^2} \quad ; \\ \theta &= \arctan\left(\frac{z}{y}\right) \quad .\end{aligned}\tag{4.32}$$

Il valore di ρ è quindi modificato sommando ad esso l'altezza delle creste:

$$\rho' = \rho + \rho_0 \quad ,\tag{4.33}$$

dove ρ' è il nuovo valore del raggio e ρ_0 rappresenta l'altezza delle creste. In Figura 4.33 è mostrato un esempio di calcolo della coordinata delle creste. Le coordinate polari sono quindi trasformate nuovamente in coordinate cartesiane, secondo le formule:

$$\begin{aligned}y' &= \rho' \cos \theta \quad ; \\ z' &= \rho' \sin \theta \quad ,\end{aligned}\tag{4.34}$$

dove y' e z' sono le nuove coordinate cartesiane dei punti delle creste.

Il valore di altezza delle creste è calcolato considerando la risoluzione dell'immagine di partenza. In un'immagine con risoluzione standard di 500 dpi, un pixel corrisponde ad una dimensione fisica in pollici di $1/500$, pari a circa 0.05 mm. Utilizzando un valore di altezza delle creste pari a 6 voxel, si impone una dimensione fisica di $6 \times 0.05 = 0.30$ mm, molto simile all'altezza media delle creste nella popolazione, pari a circa 0.27 mm ([42]).

La nuvola di punti finale è quindi elaborata tramite un'interpolazione di superficie, in modo da ottenere una rappresentazione visuale migliore. In Figura 4.34 sono mostrati un esempio di immagine tradizionale di impronta digitale acquisita tramite contatto ed il corrispondente modello tridimensionale sintetico.

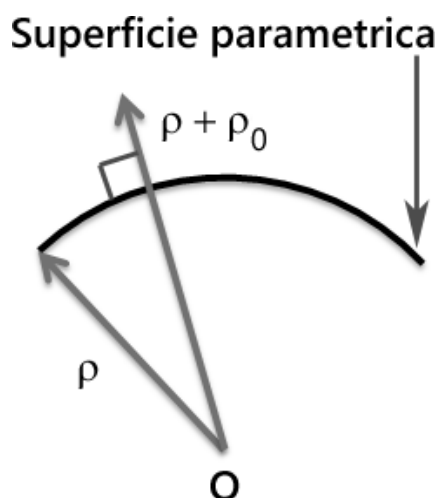


Figura 4.33: Calcolo dell'altezza delle creste

4.5.2 Definizione del setup stereoscopico

Così come è necessario definire l'ambiente reale di acquisizione delle immagini stereoscopiche, come descritto in 2.4.4.2, risulta necessario definire un ambiente virtuale in cui inserire il modello tridimensionale sintetico. L'ambiente utilizzato dal metodo proposto è definito da tre caratteristiche:

1. dimensioni dello spazio virtuale;
2. posizione e direzione della fotocamera e della fonte di luce;
3. distanza tra le viste stereoscopiche.

Per prima cosa, si definiscono le dimensioni dello spazio virtuale, lungo le coordinate (x, y, z) . Tali dimensioni sono scelte a priori e non dipendono dalle dimensioni dell'impronta digitale e del corrispondente modello sintetico. In questo modo l'acquisizione di immagini stereoscopiche viene simulata nel medesimo spazio a prescindere dalle differenti dimensioni delle immagini di partenza.

La fotocamera virtuale è posizionata sulla sommità di questo spazio, centrale rispetto alle coordinate planari (x, y) e diretta verticalmente verso

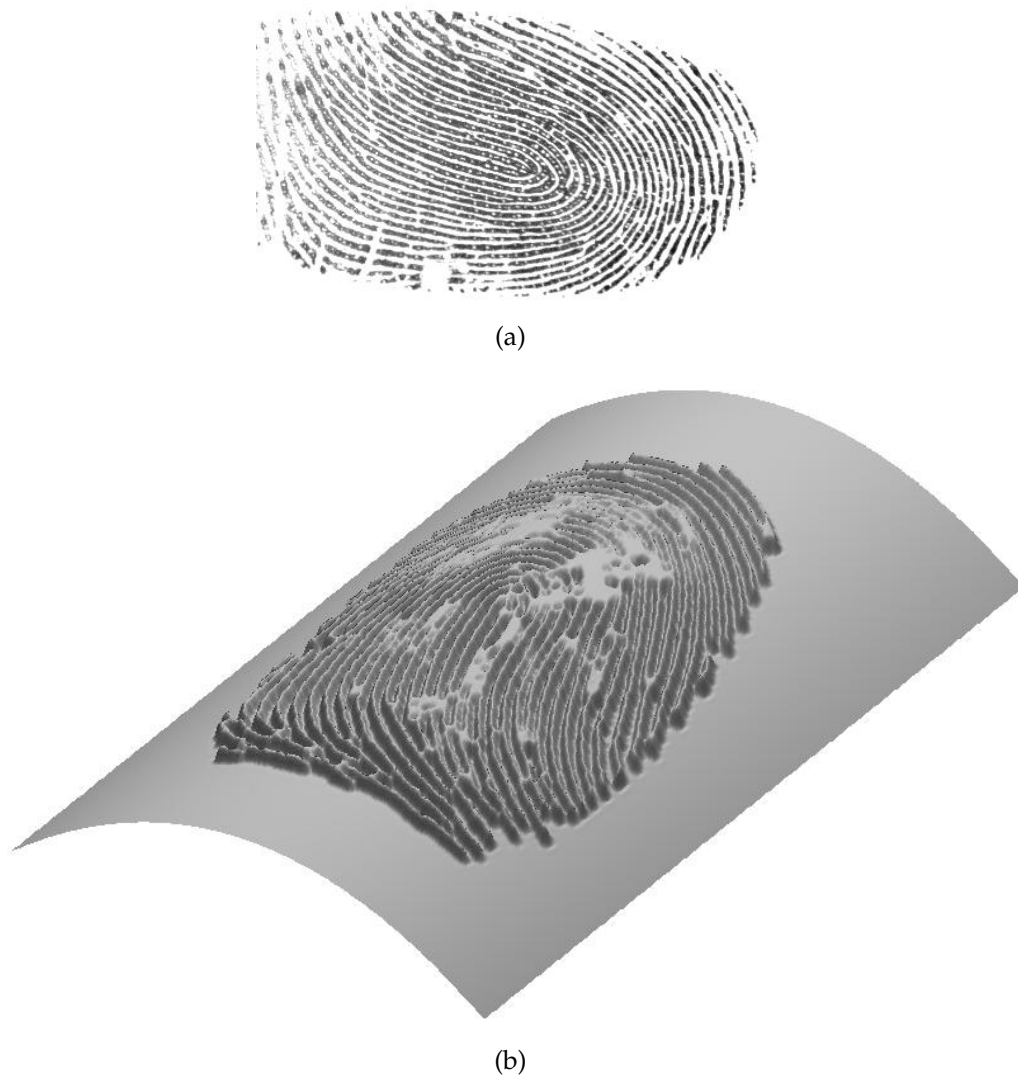


Figura 4.34: *Modello tridimensionale sintetico del dito: (a) immagine acquisita con un sensore tradizionale; (b) modello tridimensionale sintetico basato sull'immagine (a)*

il basso. In modo analogo è posizionata la fonte di luce. Invece di utilizzare due fotocamere su una determinata baseline come avviene nella realtà, il metodo proposto definisce una sola fotocamera virtuale, spostando invece il modello degli oggetti secondo il valore di baseline, come mostrato in Figura 4.35. Questa soluzione, non applicabile nella realtà, è di fatto

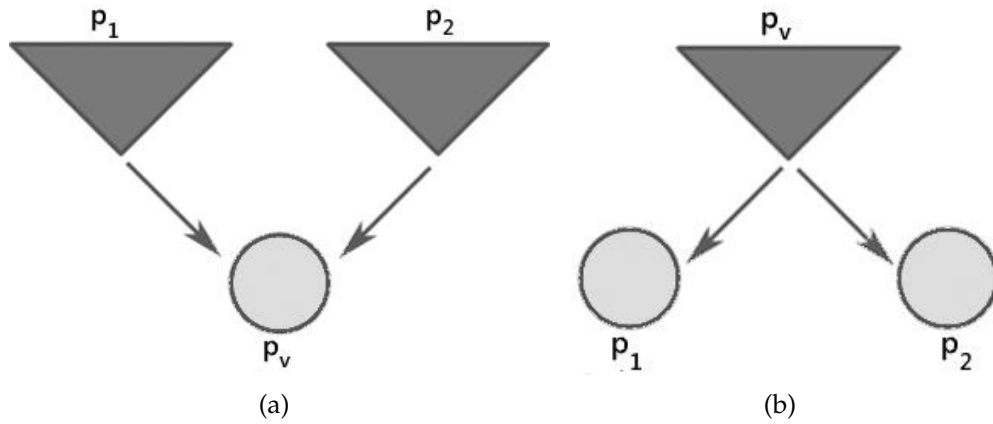


Figura 4.35: Confronto tra setup stereoscopico reale e setup stereoscopico virtuale:
 (a) situazione reale; (b) situazione virtuale

analoga al setup stereoscopico reale.

Il modello tridimensionale del dito è quindi acquisito in due posizioni, simmetriche rispetto alla posizione della fotocamera e spostate tra di loro di un intervallo Δx . Il setup stereoscopico utilizzato per la simulazione di acquisizioni di immagini stereoscopiche è mostrato in Figura 4.36.

I parametri del setup stereoscopico da considerare sono i seguenti:

- **Coordinate dei limiti dello spazio virtuale:** si definisce la terna di valori (x_v, y_v, z_v) , che rappresentano rispettivamente le dimensioni dello spazio virtuale lungo gli assi x, y e z ;
- **Posizione della fotocamera e della fonte di luce:** la fotocamera virtuale e la fonte di luce sono posizionate sulla sommità dello spazio tridimensionale e centralmente rispetto alle coordinate (x, y) . Le loro coordinate sono quindi calcolate come:

$$(x_p, y_p, z_p) = \left(\frac{x_v}{2}, \frac{y_v}{2}, z_v \right) ; \quad (4.35)$$

- **Direzione della fotocamera e della fonte di luce:** la fotocamera è diretta verso il basso, perpendicolarmente all'asse z . La fonte di luce è diretta in modo analogo in modo da illuminare in modo uniforme lo spazio virtuale;

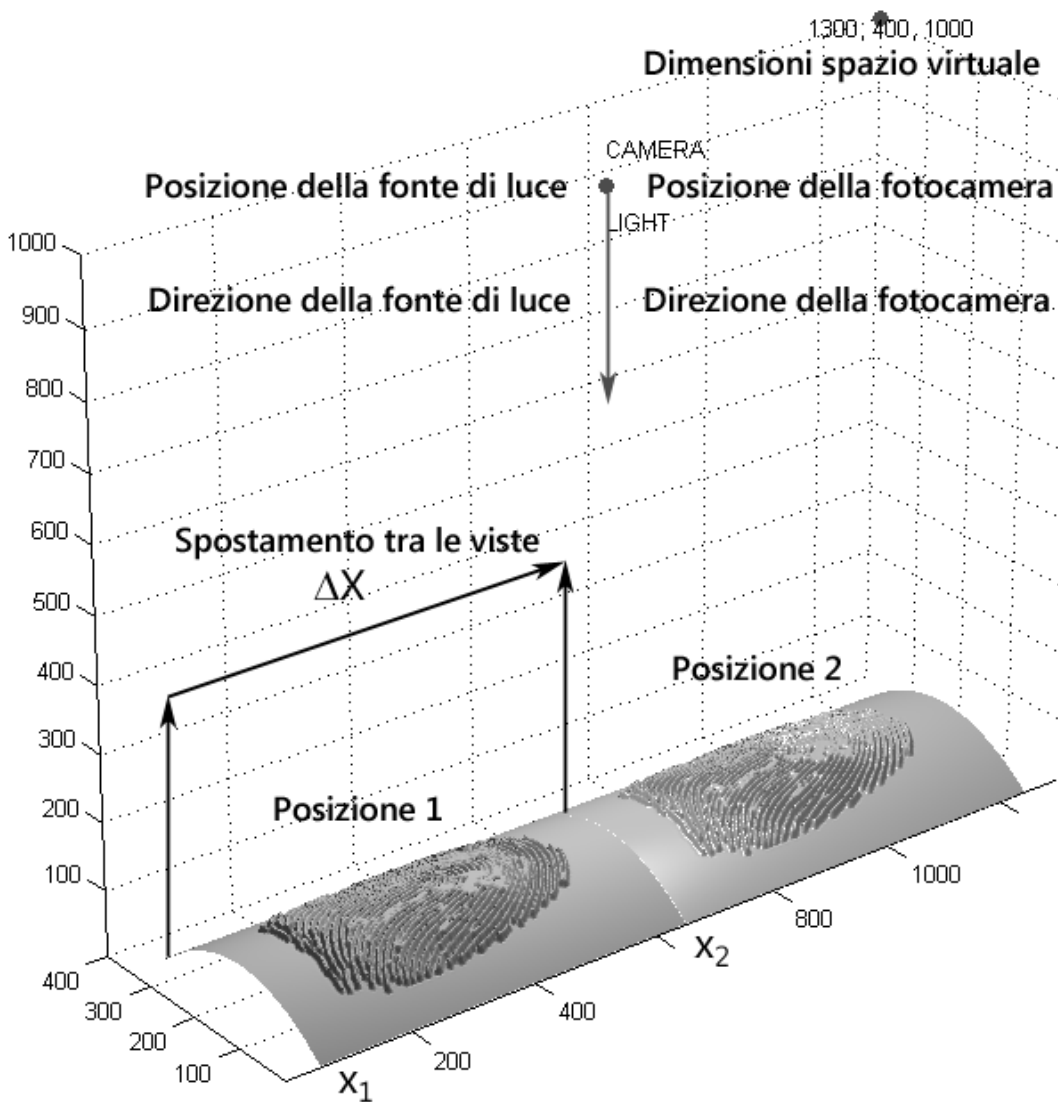


Figura 4.36: Setup stereoscopico virtuale

- **Spostamento tra le posizioni del modello degli oggetti:** le due posizioni del modello sono distanziate di un valore Δx , secondo le formule:

$$\begin{aligned} x_1 &= \frac{x_v - (x_l + \Delta x)}{2} \\ x_2 &= x_1 + \Delta x \quad , \end{aligned} \quad (4.36)$$

dove x_l è la lunghezza del modello sintetico, x_v è la dimensione lungo

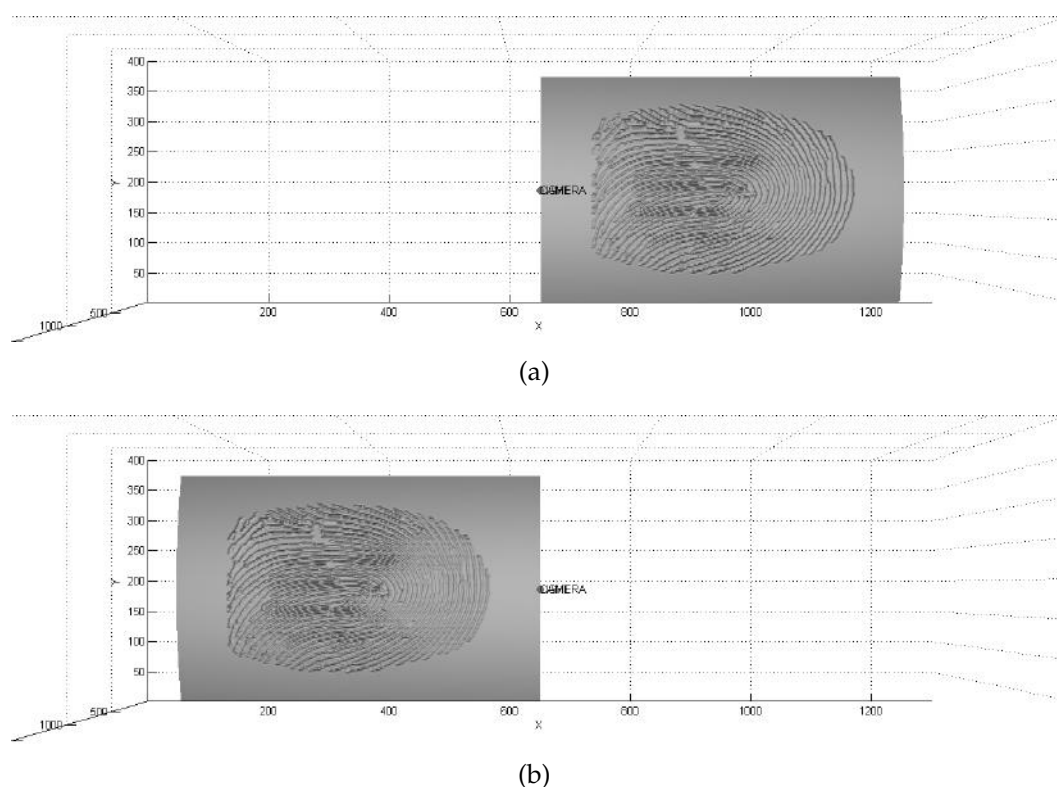


Figura 4.37: *Simulazione di acquisizione stereoscopica del modello sintetico del dito: (a) immagine sinistra; (b) immagine destra*

l'asse x dello spazio virtuale e x_1 , x_2 sono rispettivamente le coordinate lungo l'asse x delle due posizioni del modello. Le posizioni sono scelte in modo che la fotocamera si trovi in posizione centrale rispetto ad esse, al fine di simulare il modello di acquisizione stereoscopica reale.

In Figura 4.37 è mostrato un'esempio di risultato ottenuto tramite la simulazione di acquisizione stereoscopica del modello tridimensionale sintetico.

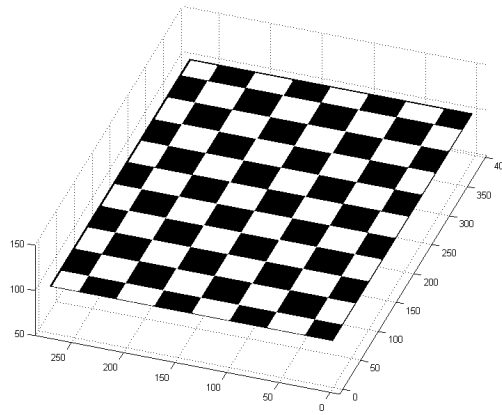


Figura 4.38: Modello tridimensionale sintetico del pattern di calibrazione

4.5.3 Calcolo del modello sintetico del pattern di calibrazione

Il modello tridimensionale sintetico del pattern di calibrazione viene ottenuto interpolando la superficie di un piano nello spazio tridimensionale, con una tecnica simile a quella descritta in 4.1.2. La superficie risultante viene infine modificata tramite un'operazione di *texture mapping*, sovrapponendo ad essa l'immagine di una scacchiera.

La tecnica di texture mapping consente di mappare un'immagine su una superficie tridimensionale modificando il colore della superficie stessa. In questo modo è possibile simulare un oggetto tridimensionale applicandone una sua foto su una superficie, senza effettuare un procedimento di ricostruzione delle sue proprietà geometriche.

Il modello risultante è infine ridimensionato in modo che la larghezza della scacchiera sia pari alla larghezza del modello sintetico dell'impronta digitale.

In Figura 4.38 è mostrato il risultato del calcolo di un modello tridimensionale sintetico del pattern di calibrazione.

Il dataset di calibrazione del sistema stereoscopico virtuale è ottenuto acquisendo il pattern di calibrazione in differenti posizioni di rotazione. La scacchiera virtuale è ruotata intorno al suo punto centrale secondo due assi

di rotazione, rispettivamente paralleli all'asse x e all'asse y , con differenti gradi di rotazione. I punti ruotati sono dati dalle formule:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_x(\psi)R_y(\phi) \begin{bmatrix} x \\ y \\ z \end{bmatrix} , \quad (4.37)$$

dove (x', y', z') sono le coordinate dei punti ruotati e $R_x(\psi)$ e $R_y(\phi)$ sono rispettivamente le matrici di rotazione lungo gli assi x e y , descritte nella formula 2.13:

$$R_x(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & \sin \psi \\ 0 & -\sin \psi & \cos \psi \end{bmatrix} ,$$

$$R_y(\phi) = \begin{bmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{bmatrix} ,$$

Alcune prove visuali hanno permesso di definire un dataset valido di immagini di calibrazione utilizzando 30° , 40° e 150° come valori di ψ e ϕ .

Il procedimento di acquisizione della scacchiera è quindi effettuato inserendo il modello virtuale della scacchiera all'interno del setup stereoscopico e ripetendo il procedimento per entrambe le posizioni separate dal valore Δx .

Un esempio di rotazione delle scacchiere in entrambe le posizioni è mostrato in Figura 4.39, mentre in Figura 4.40 è mostrata un'acquisizione stereoscopica della scacchiera.

4.5.4 Elaborazione delle immagini sintetiche

Per poter confrontare le immagini ottenute tramite acquisizione stereoscopica del modello sintetico con quelle tradizionali ottenute tramite con-

tatto, è necessario sottoporle al procedimento di ricostruzione tridimensionale tramite modello cilindrico, come descritto in 4.3.2, ed alla successiva fase di unwrapping, come descritto in 4.4.

La separazione tra *foreground* e *background* è già molto netta dal momento che lo sfondo è perfettamente bianco. L'operazione di segmentazione dell'immagine è realizzata quindi tramite una sogliatura dei livelli di grigio, secondo la formula:

$$\begin{cases} M(x, y) = 1 & \text{se } I(x, y) > 0 \\ M(x, y) = 0 & \text{se } I(x, y) = 0 \end{cases} , \quad (4.38)$$

dove $I(x, y)$ è l'immagine e $M(x, y)$ è la maschera binaria.

Successivamente, la ricerca della zona di interesse è eseguita tramite la funzione di segmentazione delle creste descritta in [39]. La segmentazione è effettuata calcolando la deviazione standard dei valori di intensità dei pixel in una regione S dell'immagine $I(x, y)$, secondo la formula:

$$\sigma_S = \sqrt{\frac{\sum_{s,t \in S} (I(s, t) - \bar{I}(s, t))^2}{S}} , \quad (4.39)$$

dove (s, t) sono le coordinate dei pixel appartenenti alla regione S . Se la deviazione standard σ_S è maggiore di una soglia fissata, la regione è considerata parte dell'impronta digitale:

$$\begin{cases} R(s, t) = 1 & \text{se } \sigma_S \geq r \\ R(s, t) = 0 & \text{se } \sigma_S < r \end{cases} , \quad (4.40)$$

dove r è una soglia stimata empiricamente e $R(s, t)$ è la maschera binaria che delimita la regione di interesse.

Il punto corrispondente all'ultima giunzione del dito, inoltre, non è presente e viene approssimato ad uno degli estremi dell'area di interesse.

In Figura 4.41 sono mostrati un esempio di immagine acquisita da un modello tridimensionale sintetico e le due maschere binarie di segmentazione e della regione di interesse.

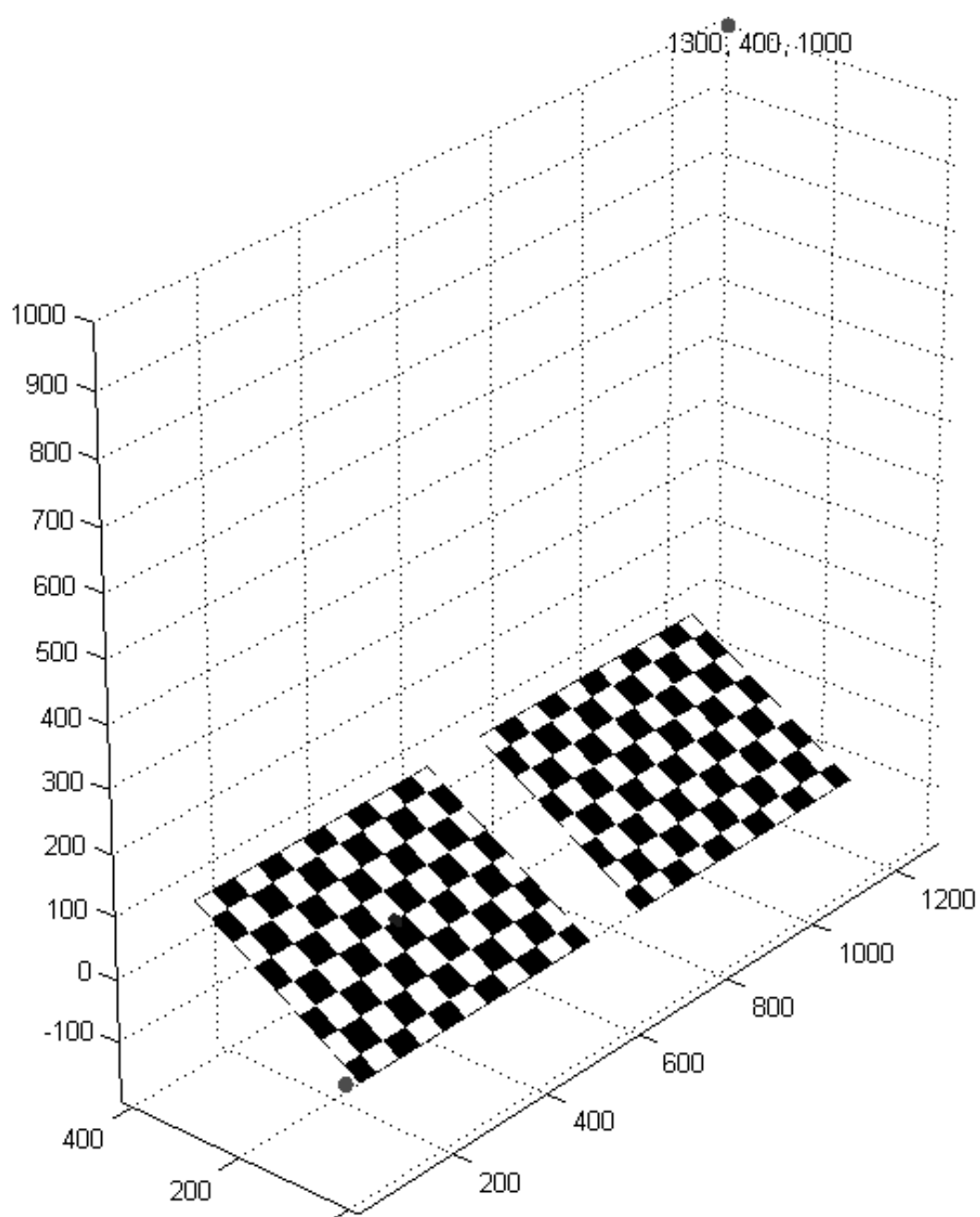
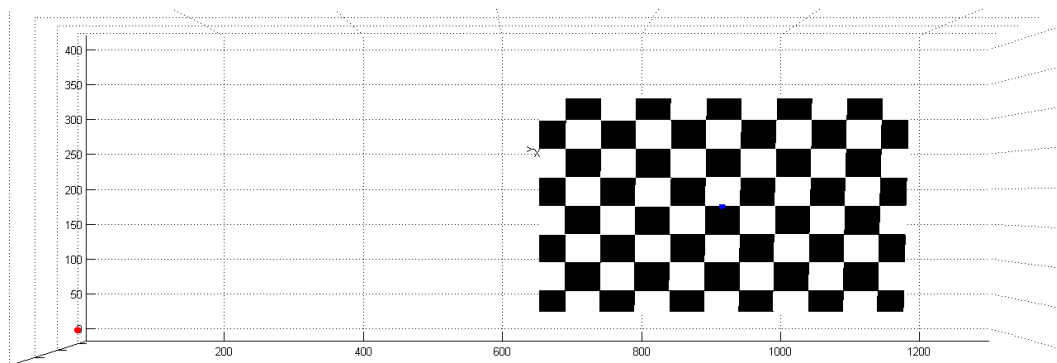
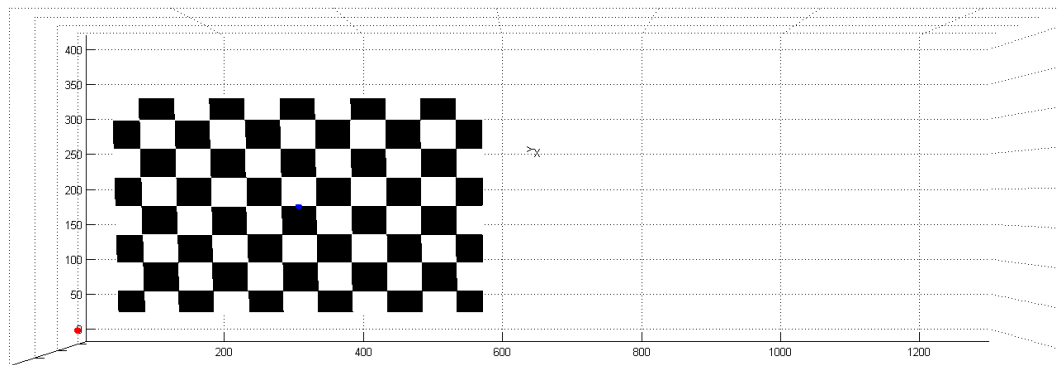


Figura 4.39: *Rotazione del pattern di calibrazione*



(a)



(b)

Figura 4.40: *Acquisizione stereoscopica del pattern di calibrazione: (a) immagine sinistra; (b) immagine destra*

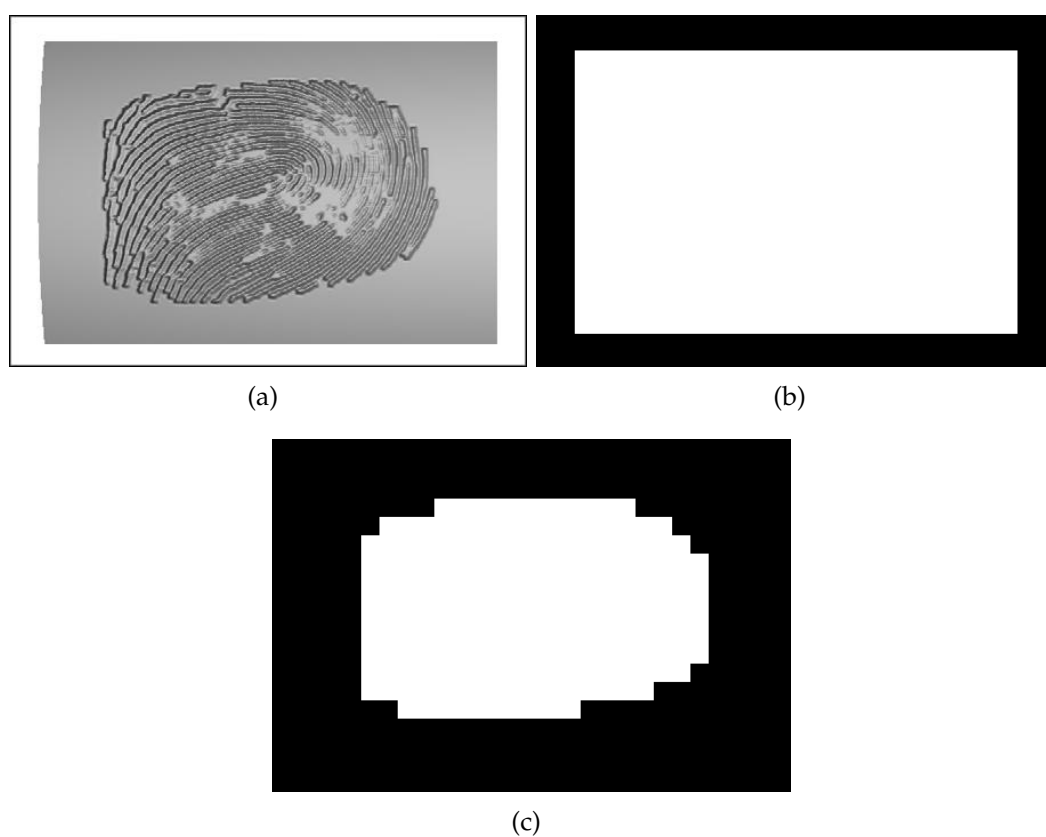


Figura 4.41: Segmentazione dell'immagine ottenuta dal modello sintetico: (a) immagine originale; (b) maschera di segmentazione; (c) maschera della regione di interesse

RISULTATI SPERIMENTALI

In questo capitolo sono descritti i risultati degli esperimenti effettuati sui metodi proposti. Per prima cosa, sono stati condotti i test riguardanti la precisione del procedimento di calibrazione. Successivamente, è stata valutata l'accuratezza degli algoritmi di preprocessing e quella dei metodi di matching su immagini stereoscopiche.

È stata infine valutata l'accuratezza ottenuta dal sistema biometrico utilizzando le immagini tradizionali ed immagini risultanti dalla fase di unwrapping dei modelli tridimensionali sintetici.

I test sono stati effettuati su un computer portatile dalle seguenti caratteristiche;

- processore *Intel Core 2 Duo* 2.00 GHz;
- memoria *RAM* 2 GB *DDR2* 667 MHz;
- sistema operativo *Windows* 7;
- Matlab R2010a [43];
- OpenCV 1.1a [1].

5.1 Descrizione delle immagini

I test sono stati effettuati utilizzando i seguenti dataset di immagini:

- **Dataset N.1:** immagini stereoscopiche acquisite senza contatto del dito con un sensore;
- **Dataset N.2:** immagini di impronte digitali acquisite tramite sensore dedicato;
- **Dataset N.3:** immagini stereoscopiche ottenute dai modelli tridimensionali sintetici;
- **Dataset N.4:** immagini ottenute tramite unwrapping dei modelli tridimensionali parametrici.

Il dataset N.1 contiene 6 immagini composte da 3 coppie stereoscopiche di dita, di cui 2 coppie sono state acquisite da un individuo e una coppia è stata acquisita da un secondo individuo. Il dataset comprende inoltre 14 coppie stereoscopiche di immagini di calibrazione. La scacchiera di calibrazione è composta da 5×7 corner con lato di 4.3 mm. L'acquisizione è stata effettuata con una fotocamera SONY XCD-V90 in toni di grigio, con risoluzione 1280×1024 . Le immagini sono state salvate in formato bitmap. Il setup stereoscopico usato ha le seguenti caratteristiche:

- Baseline: 20 mm;
- Lunghezza focale: 25 mm;
- Distanza fotocamera-dito: 300 mm;
- Illuminazione: lampada a luce bianca.

Un esempio di immagini di questo dataset è mostrato in Figura 5.1.

Il dataset N.2, costituito da impronte digitali acquisite con un sensore tradizionale, è composto da 138 immagini, che comprendono 45 individui

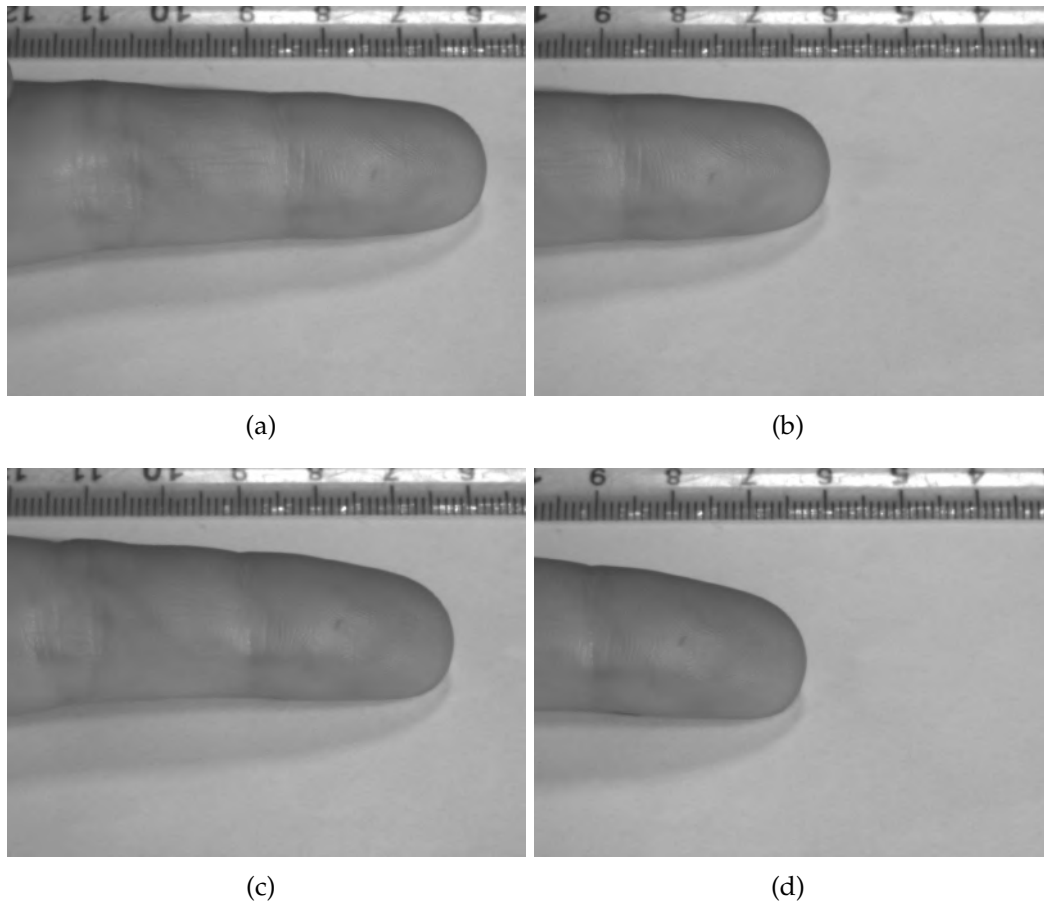


Figura 5.1: Due coppie di immagini del dataset N.1: (a),(c) immagini acquisite tramite la vista di sinistra; (b),(d) immagini acquisite tramite la vista di destra

diversi. Per ogni individuo sono presenti 3 acquisizioni dell'impronta digitale. Le immagini sono state acquisite utilizzando un sensore ottico *Cross Match Verifier 300 Classic* alla risoluzione di 500 dpi. Le immagini sono state salvate in formato tiff non compresso. In Figura 5.2 sono mostrati alcuni esempi di immagini appartenenti al dataset.

Il dataset N.3 è ottenuto con il procedimento descritto in 4.5 tramite le immagini del dataset N.2 di impronte digitali acquisite tramite contatto. Il dataset comprende le impronte digitali di 45 individui diversi, ognuna acquisita 3 volte. Ogni acquisizione comprende un'immagine sin-



Figura 5.2: Esempi di immagini del dataset N.2: le immagini (a),(b),(c),(d) appartengono a persone diverse

istra e un'immagine destra. Ogni modello tridimensionale sintetico ricostruito è stato inserito in un setup stereoscopico virtuale dalle seguenti caratteristiche:

- Dimensione x : 1300;
- Dimensione y : 400;
- Dimensione z : 1500.

A partire da queste dimensioni sono stati utilizzati due diversi valori di Δx :

- Δx_1 : lunghezza del modello;
- Δx_2 : $\frac{1}{2}$ lunghezza del modello.

Le immagini sono state salvate in formato tiff non compresso. In Figura 5.3 è mostrato un esempio di coppia stereoscopica ottenuta dal modello tridimensionale sintetico, acquisita con spostamento Δx_1 tra le posizioni.

Il Dataset N.4 è stato calcolato utilizzando le immagini stereoscopiche del dataset N.3 per effettuare la ricostruzione tridimensionale tramite modello cilindrico. Successivamente, le immagini sono state sottoposte ad un-

wrapping. Le immagini sono state salvate in formato tiff non compresso. In Figura 5.4 è mostrato un esempio di immagini appartenenti al dataset.

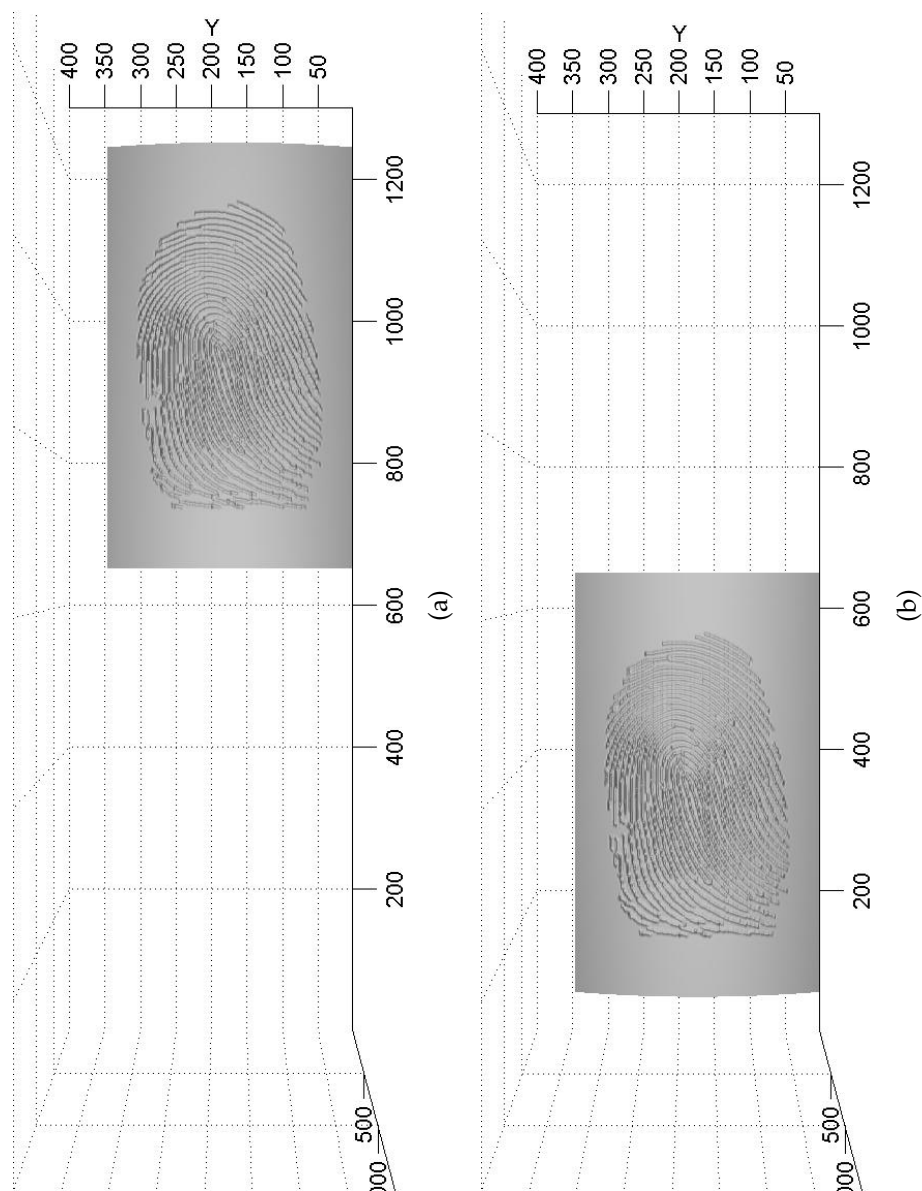


Figura 5.3: Esempi di immagini del dataset N.3: (a) immagine sinistra; (b) immagine destra



(b)

(a)

Figura 5.4: Esempi di immagini del dataset N.4: (a) immagine sinistra; (b) immagine destra

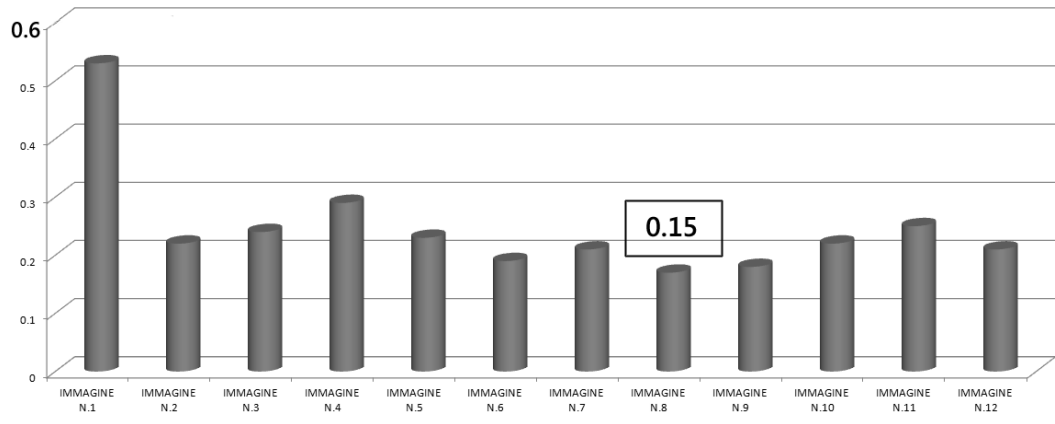


Figura 5.5: *Deviazione standard dell'errore del dataset N.1*

5.2 Analisi dell'errore di calibrazione

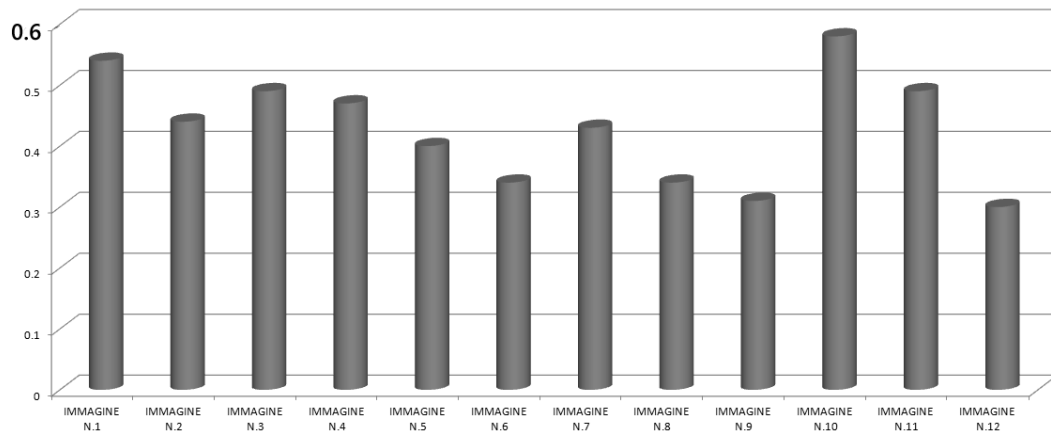
L'errore di calibrazione è calcolato come descritto in 4.1.2. In realtà, il parametro più significativo è la deviazione standard della distanza dei punti ricostruiti dal piano interpolante.

Tale parametro di errore ha però una gamma di valori diversa a seconda delle dimensioni fisiche della scacchiera. Al fine di permettere un confronto dei valori ottenuti dai diversi dataset, la deviazione standard è stata normalizzata dividendola per la lunghezza della diagonale della scacchiera. Per questo motivo i valori di errore sono adimensionali.

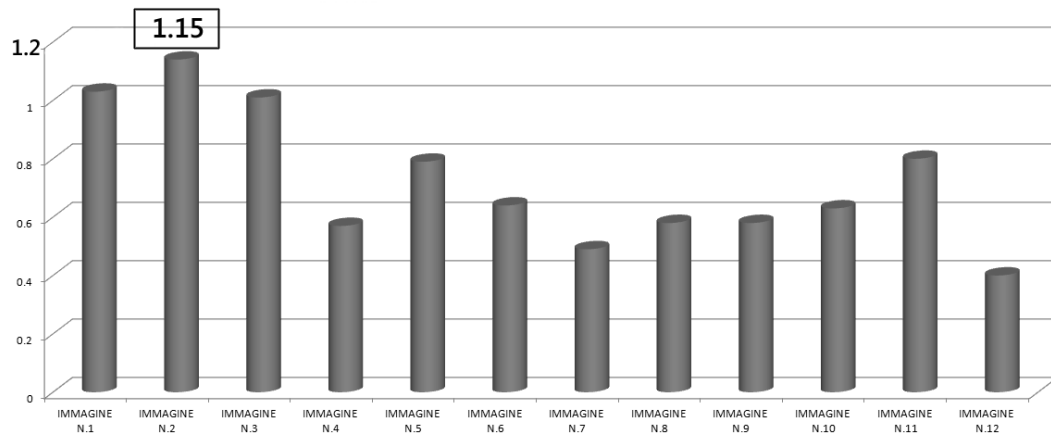
La calibrazione dei dataset è stata effettuata con il software OpenCV.

La procedura di calibrazione utilizzata durante l'acquisizione del dataset N.1 di immagini stereoscopiche acquisite senza contatto ha fornito i risultati mostrati in Figura 5.5, caratterizzati da un valore medio dell'errore di 0.245 ed una varianza di 0.0091. I risultati del dataset N.3, calcolato con i due diversi valori di spostamento tra le viste Δx_1 e Δx_2 , sono mostrati in Figura 5.6. Nel primo caso il valore medio dell'errore è stato pari a 0.43 mentre la varianza è stata di 0.0091; nel secondo caso invece il valore medio dell'errore è stato pari a 0.72 e la varianza è stata di 0.008.

Da una comparazione tra i risultati numerici e quelli visuali è stato



(a)



(b)

Figura 5.6: Deviazione standard dell'errore del dataset N.3 sintetico: (a) spostamento Δx_1 ; (b) spostamento Δx_2

stimato un valore di soglia dell'errore pari a 0.5 che rappresenta il massimo errore accettabile nel contesto applicativo. L'immagine N.8 del dataset N.1 ha presentato un'errore pari a 0.15. Come si può osservare nella Figura 5.7, la ricostruzione è visivamente accurata.

L'immagine N.2 del dataset N.3 acquisito con lo spostamento Δx_2 , invece, ha presentato un'errore pari a 1.15. Come mostrato in Figura 5.8, la ricostruzione è visivamente poco accurata e lo scostamento dei punti

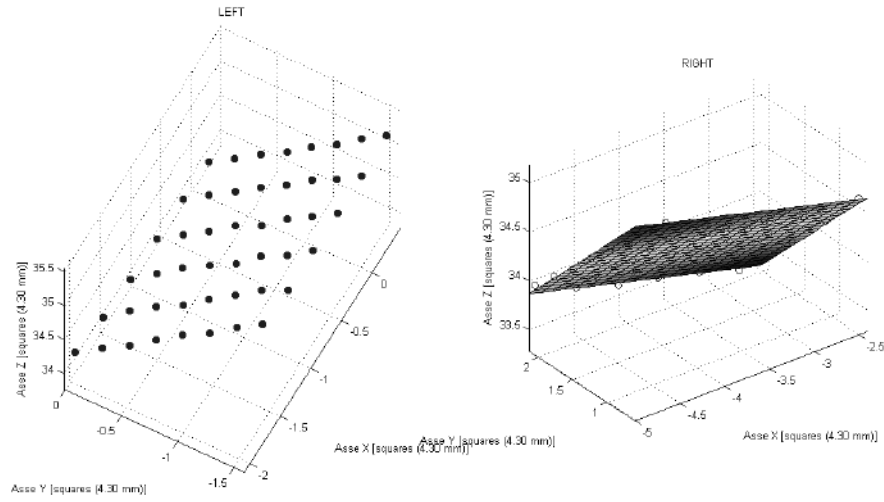


Figura 5.7: Ricostruzione dell'immagine N.8 del dataset N.1

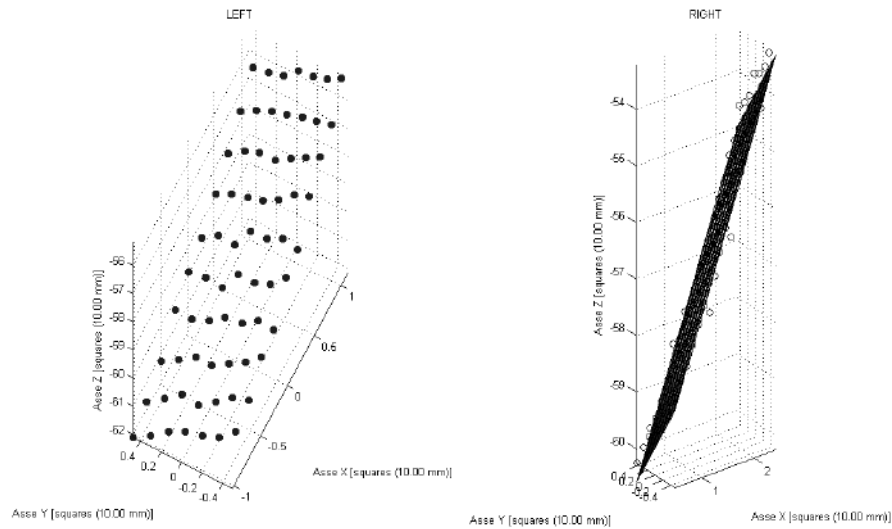


Figura 5.8: Ricostruzione dell'immagine N.2 del dataset N.3 con Δx_2

triangolati dal piano interpolante è chiaramente visibile.

Al fine di valutare tramite un'unica rappresentazione l'accuratezza di tutte le immagini che compongono i dataset, è stata calcolata la frequenza cumulata della deviazione standard.

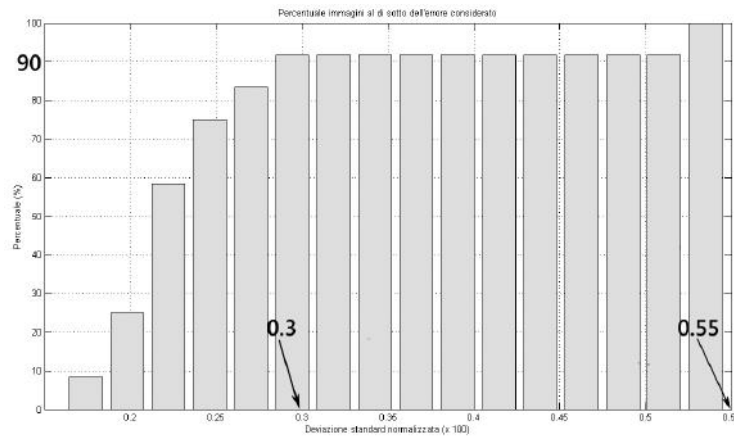


Figura 5.9: *Frequenza cumulata della deviazione standard del dataset N.1*

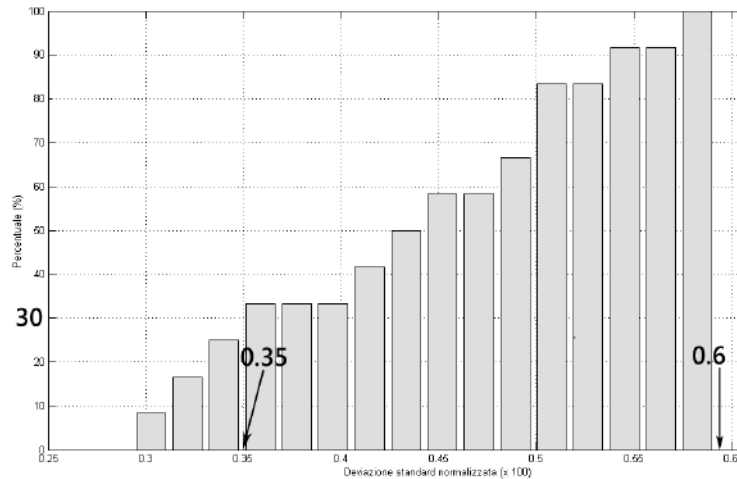


Figura 5.10: *Frequenza cumulata della deviazione standard del dataset N.3 con Δx_1*

Nelle Figure 5.9, 5.10, 5.11 sono mostrati i grafici della frequenza cumulata della deviazione standard del dataset N.1 e del dataset N.3 acquisito con spostamento Δx_1 e Δx_2 .

È possibile osservare come il dataset N.1 e il dataset N.3 con spostamento Δx_1 abbiano ottenuto un valore massimo di deviazione standard pari a 0.5, con il risultato che tutte le immagini del dataset sono state ricostruite con un errore minore o uguale alla soglia fissata. Il dataset N.1, inoltre, ha presentato il 90% delle immagini con un errore minore a 0.3,

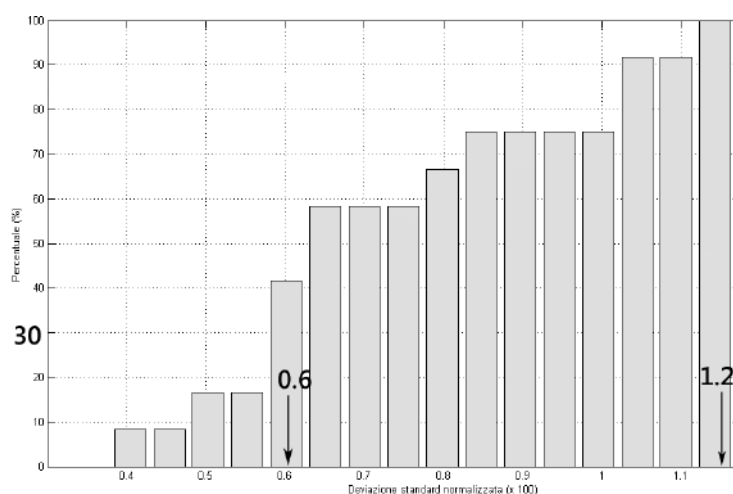


Figura 5.11: *Frequenza cumulata della deviazione standard del dataset N.3 con Δx_2*

corrispondente a circa metà del valore massimo della deviazione standard.

Il dataset N.3 con spostamento Δx_1 , invece, ha fornito un valore massimo di deviazione standard di 0.6, con la conseguenza che quasi tutte le immagini sono state ricostruite con un errore minore o uguale alla soglia fissata. Solo il 30% delle immagini, però, ha presentato un valore di deviazione standard inferiore alla metà del rispettivo valore massimo.

Il dataset N.3 con spostamento Δx_2 , invece, ha presentato un valore massimo di 1.2. Il 70% delle immagini, infatti, ha fornito un'errore di ricostruzione superiore alla soglia fissata.

Per questi motivi il dataset N.1, sottoposto a binarizzazione, ha fornito i migliori risultati di ricostruzione. Il dataset N.3 con spostamento Δx_1 , inoltre, ha fornito una buona ricostruzione visiva. Il dataset N.3 con spostamento Δx_2 , a causa dell'elevata percentuale di immagini che hanno presentato valori di errore troppo elevati, non fornito una ricostruzione tridimensionale accurata.

Tipo di filtraggio	Minuzie estratte	Match score
<i>Coppia n. 1</i>		
A	37	21
B	65	36
C	105	54
D	94	68
<i>Coppia n. 2</i>		
A	41	44
B	40	48
C	71	101
D	69	66

Tabella 5.1: Risultanti del matching tra l'immagine sinistra e l'immagine destra del dataset N.1

5.3 Confronto tra gli algoritmi di preprocessing

Il confronto tra le tecniche di preprocessing descritte in 4.2.3 è stato effettuato utilizzando i metodi descritti in 4.2.4, in modo da verificare il tasso di somiglianza tra i risultati dei filtri applicati a due viste stereoscopiche della stessa immagine. Dopo aver effettuato il calcolo della regione di interesse come descritto in 4.2.2, quindi, si è effettuato il procedimento di estrazione e matching delle minuzie descritto in 4.2.4. I test sono stati effettuati tra le coppie di immagini del dataset N.1.

I risultati così ottenuti hanno permesso di determinare quale degli algoritmi di preprocessing sia più adatto al calcolo delle corrispondenze stereoscopiche tra le immagini. I risultati visuali dell'applicazione dei metodi di preprocessing relativi ad una coppia di immagini sono mostrati in Figura 5.12, mentre nella Tabella 5.1 sono elencati per ciascuna coppia il numero di minuzie estratte, il match score ottenuto senza l'ausilio dell'algoritmo di raffinamento della posizione delle minuzie ed il match score raffinato.

Tra i vari metodi di preprocessing, il metodo *D*, basato su sottrazione dello sfondo e binarizzazione, è quello che presenta la minore variabilità dei risultati. Il valore di match score medio, inoltre, è paragonabile ai valori

di match score ottenuti tramite confronti genuini tra due acquisizioni della stessa impronta digitale. Per questi motivi, è stato utilizzato come metodo di riferimento.

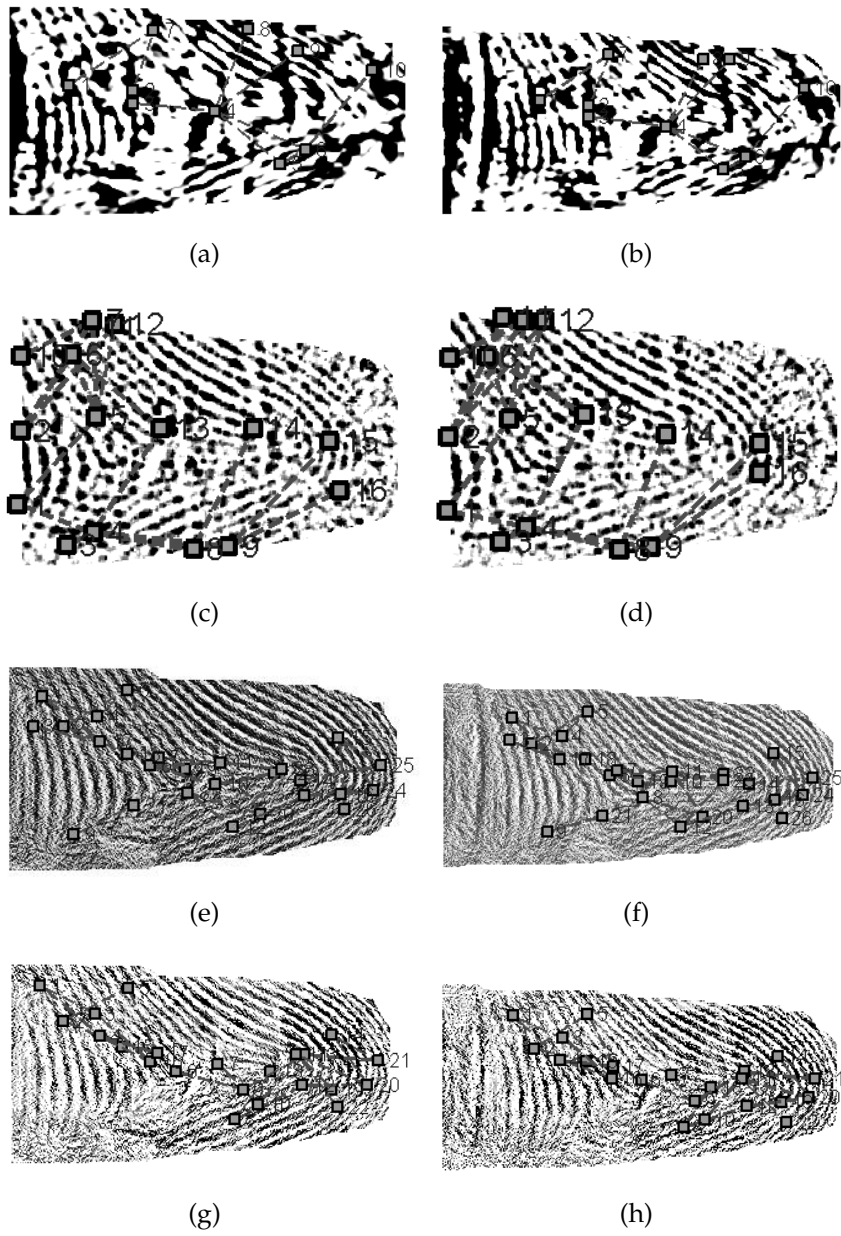


Figura 5.12: Confronto tra le tecniche di preprocessing: (a),(b) Metodo A; (c),(d) Metodo B; (e),(f) Metodo C; (g),(h) Metodo D

5.4 Confronto tra le immagini stereoscopiche

In questa sezione sono riportati i risultati dei confronti effettuati all'interno delle coppie di immagini stereoscopiche. Al fine di supportare efficacemente la ricostruzione tridimensionale basata sulle minuzie, infatti, il match score ottenuto dal confronto di due viste della stessa impronta digitale dovrebbe essere il più alto possibile.

L'accuratezza nel calcolo della posizione delle minuzie corrispondenti nelle due immagini, inoltre, deve essere sufficientemente elevata al fine di permettere una triangolazione corretta.

I test sono stati effettuati sul dataset N.3 utilizzando tre metodi di matching:

- **Metodo N.1:** matching basato su bozorth3;
- **Metodo N.2:** matching basato sulla vicinanza;
- **Metodo N.3:** matching basato su bozorth3 e raffinamento stereoscopico.

In Figura 5.13 sono mostrati i valori di match score ottenuti con il metodo di matching N.1, mentre nella Figura 5.14 è mostrata la frequenza cumulata dei valori. I valori di media e deviazione standard dei match score sono stati rispettivamente di 248 e 170.

In Figura 5.15 sono mostrati invece i valori di match score ottenuti con il metodo di matching N.2, mentre nella Figura 5.16 è mostrata la frequenza cumulata dei valori. I valori di media e deviazione standard dei match score sono stati rispettivamente di 39 e 15.

Dalle Figure si può osservare come le distribuzioni dei valori di match score siano risultate simili. Il metodo di matching N.1 ha prodotto però valori di match score elevati nella quasi totalità dei casi, mentre il metodo N.2 ha prodotto valori di match score inferiori. Il confronto effettuato con il metodo N.1, però, ha prodotto un valore di somiglianza tra le immagini,

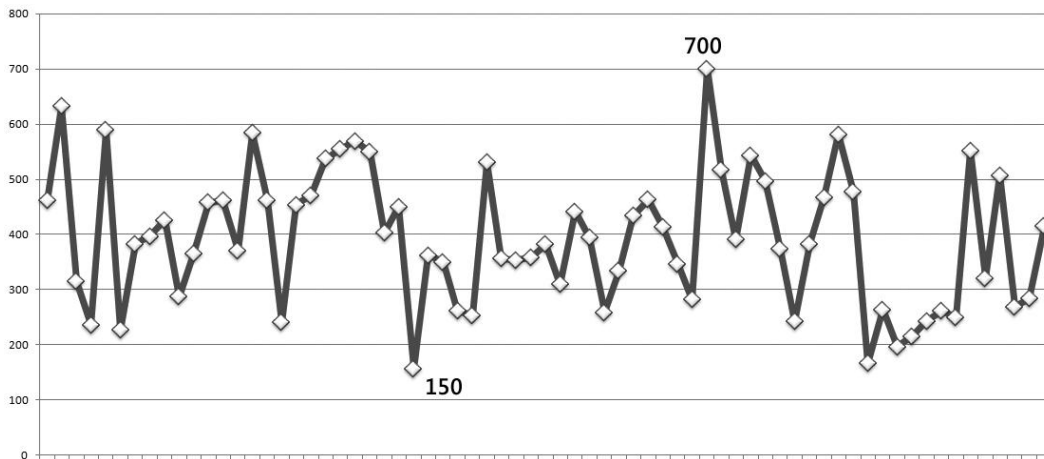


Figura 5.13: Valori di match score ottenuti con il metodo N.1

mentre il confronto basato sul metodo N.2 ha fornito il numero di minuzie corrispondenti nelle due immagini.

Per valutare l'accuratezza della corrispondenza stereoscopica tra le minuzie è stato necessario confrontare i risultati del matching effettuato con il metodo N.2 con i risultati forniti dal metodo di matching N.3. Da alcuni test effettuati il numero di minuzie corrispondenti fornite dal metodo N.2 è risultato analogo a quello fornito dal metodo N.3.

Per questi motivi, il metodo N.1 è risultato più adatto per effettuare il riconoscimento biometrico delle impronte digitali, mentre i metodi N.2 e N.3 sono risultati quasi equivalenti per ricerca delle corrispondenze stereoscopiche tra le minuzie.

In Figura 5.17 sono mostrati due esempi di matching stereoscopico effettuati con il metodo N.2 e con il metodo N.3.

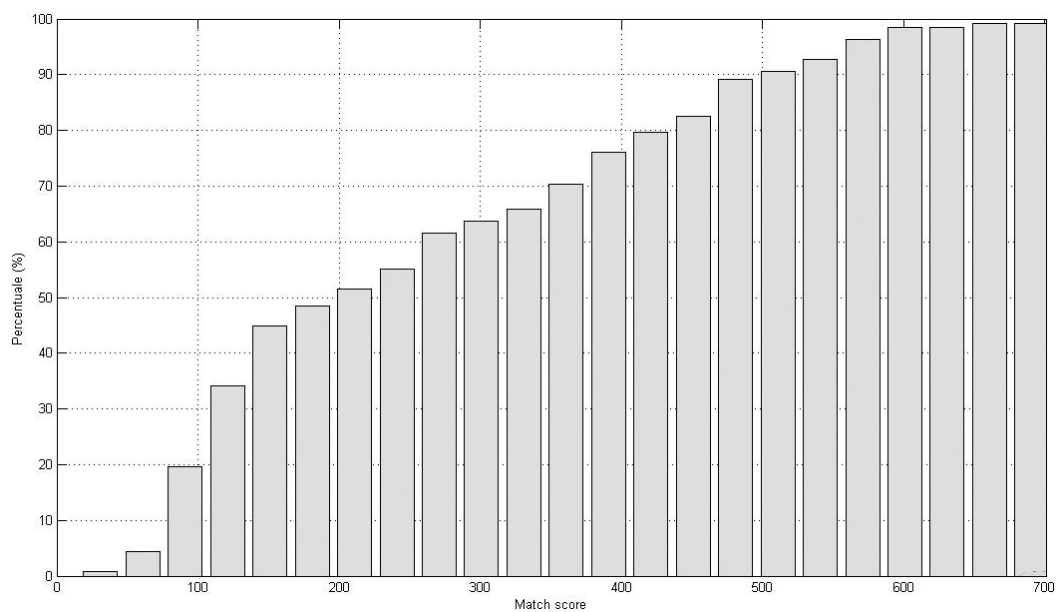


Figura 5.14: *Frequenza cumulata dei match score ottenuti con il metodo N.1*

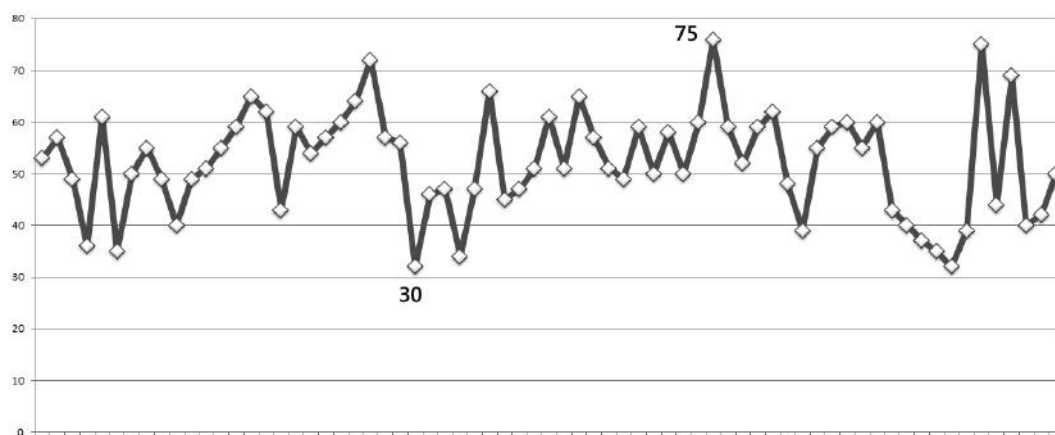


Figura 5.15: *Valori di match score ottenuti con il metodo N.2*

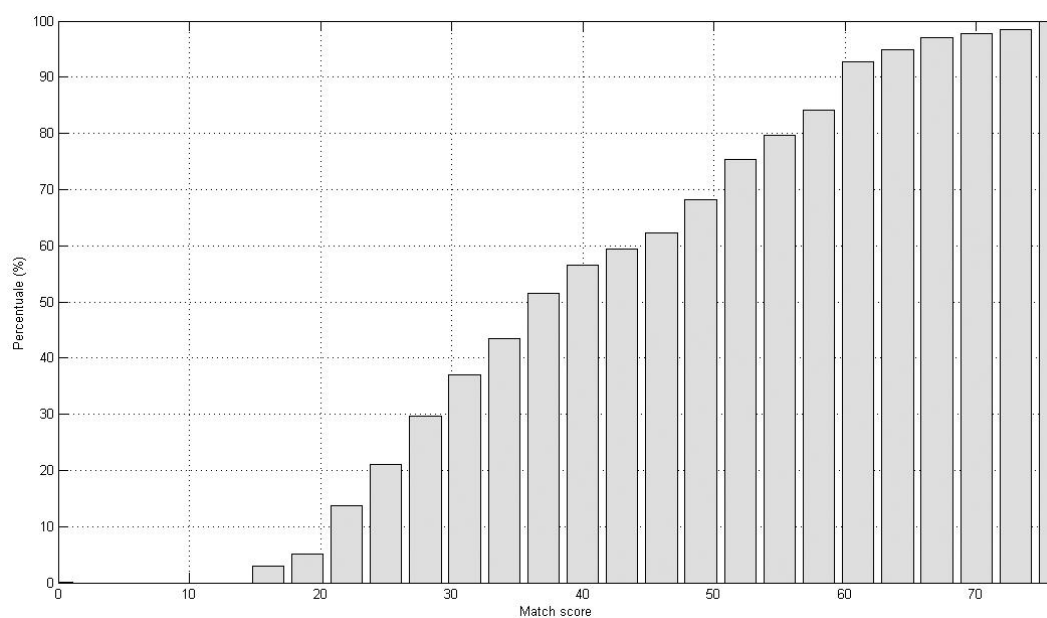


Figura 5.16: *Frequenza cumulata dei match score ottenuti con il metodo N.2*

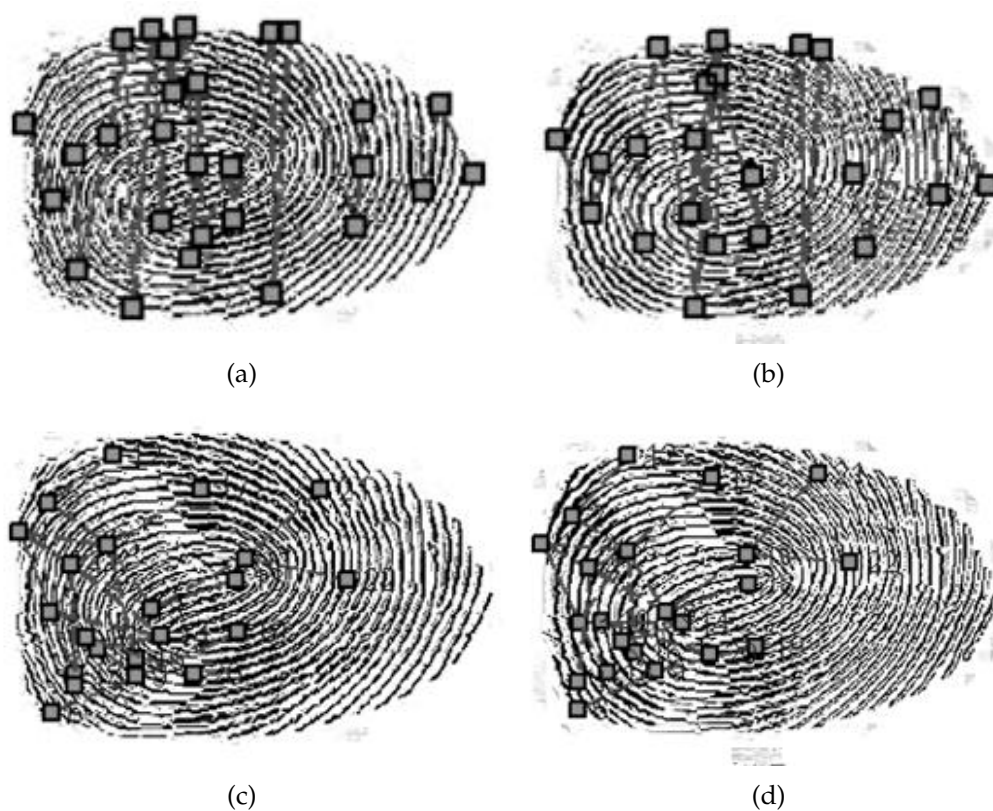


Figura 5.17: *Matching stereoscopico di una coppia di immagini del Dataset N.3: (a),(b) metodo N.2; (c),(d) metodo N.3*

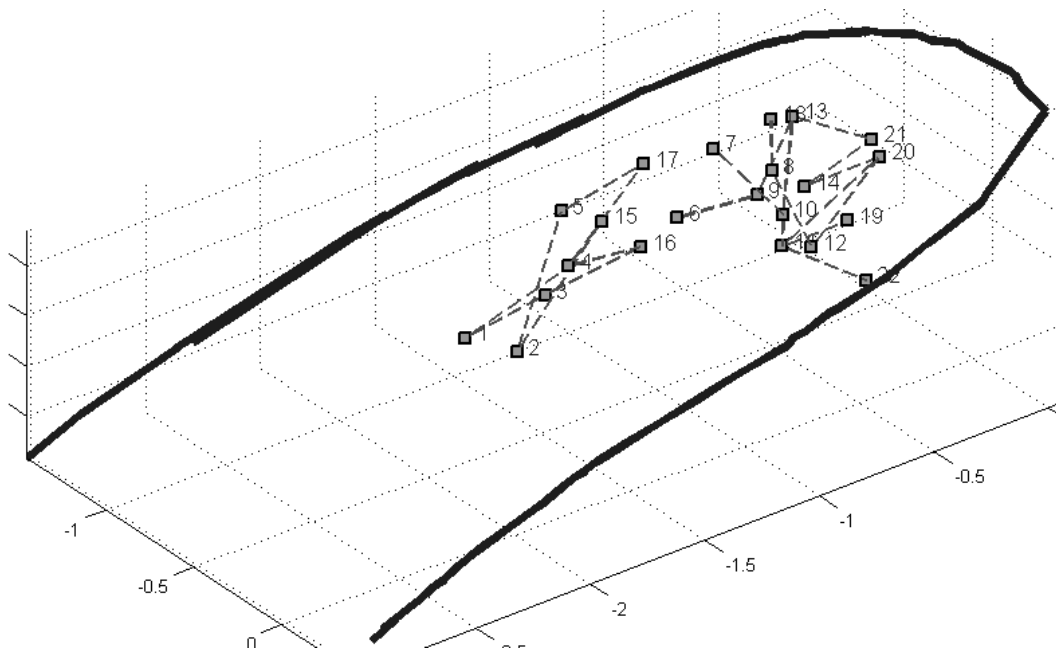


Figura 5.18: *Triangolazione delle minuzie di una coppia di immagini del Dataset N.1*

5.5 Modelli visuali

In questa sezione sono riportati alcuni esempi di modelli visuali ottenibili utilizzando le tecniche di ricostruzione tridimensionale descritte in 4.2 e 4.3. In particolare sono proposti i risultati applicati al dataset N.1 di immagini stereoscopiche acquisite senza contatto ed al dataset N.3 di immagini stereoscopiche sintetiche.

In Figura 5.18 è mostrato un esempio di triangolazione delle minuzie appartenenti alla coppia di immagini N.2 del Dataset N.1. Come già accennato in 4.2.5, la ricostruzione basata sulle minuzie genera un modello tridimensionale in cui lo spazio occupato dal dito è verosimile, ma in cui la profondità delle minuzie non è corretta a causa di errori nel calcolo della corrispondenza tra minuzie.

In Figura 5.19 è mostrato un esempio ricostruzione tridimensionale basata sulle minuzie relativa ad una coppia di immagini del Dataset N.1. Anche in questo caso, i valori di profondità delle minuzie nello spazio

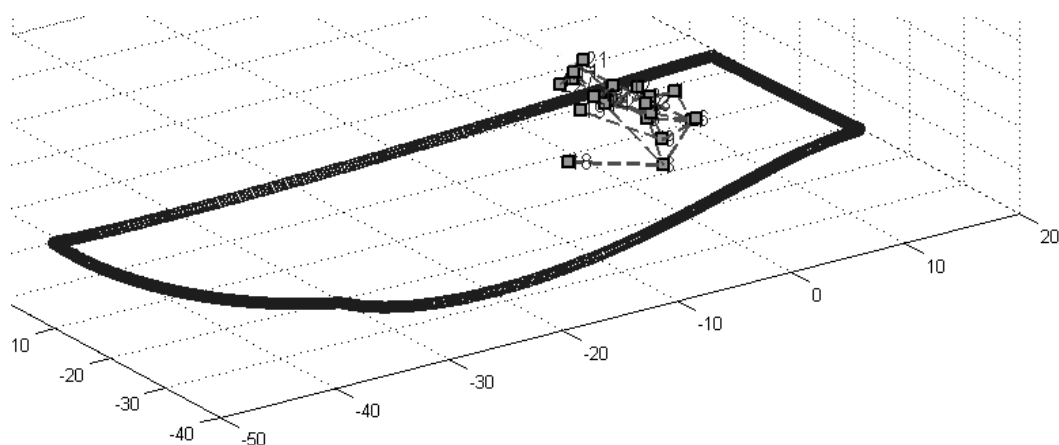
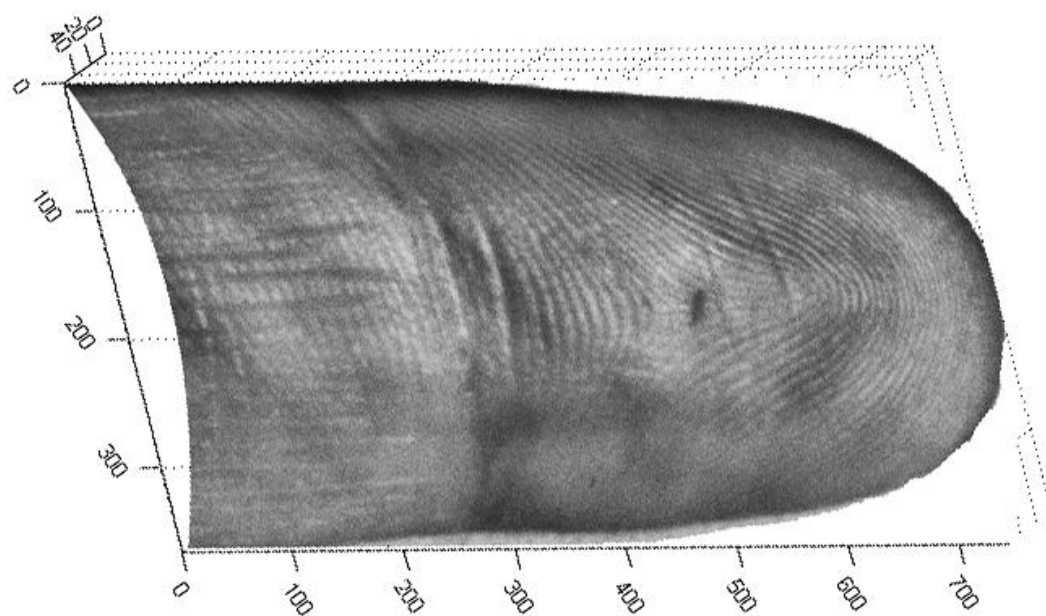


Figura 5.19: *Triangolazione delle minuzie di una coppia di immagini del Dataset N.3*

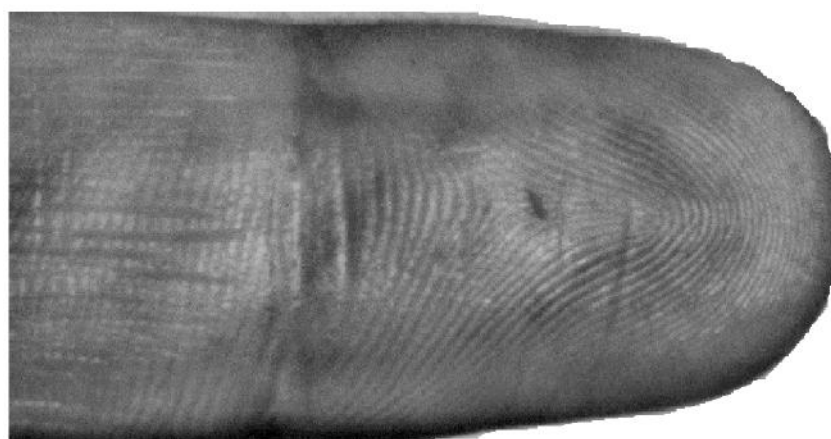
tridimensionale non rispecchiano correttamente la fisionomia originale del dito.

Nella Figura 5.20, invece, è mostrata la procedura di ricostruzione tramite modello cilindrico applicata ad una coppia di immagini del dataset N.1. Il modello così ricostruito è stato successivamente completato con il wrapping delle immagini originali ed infine sottoposto alla fase di unwrapping.

In Figura 5.21 sono mostrati i risultati delle medesime elaborazioni applicate ad un'immagine del dataset N.3. In quest'immagine si può notare come la fase di unwrapping abbia avuto effetto sulle creste posizionate nelle zone periferiche dell'impronta digitale. Non si osservano, infatti, gli effetti di aumento della frequenza apparente delle creste, tipici delle acquisizioni senza contatto.



(a)



(b)

Figura 5.20: Ricostruzione tridimensionale con modello cilindrico ed *unwrapping* del modello (immagine del Dataset N.1): (a) modello dopo la fase di *wrapping*; (b) immagine dopo la fase di *unwrapping*

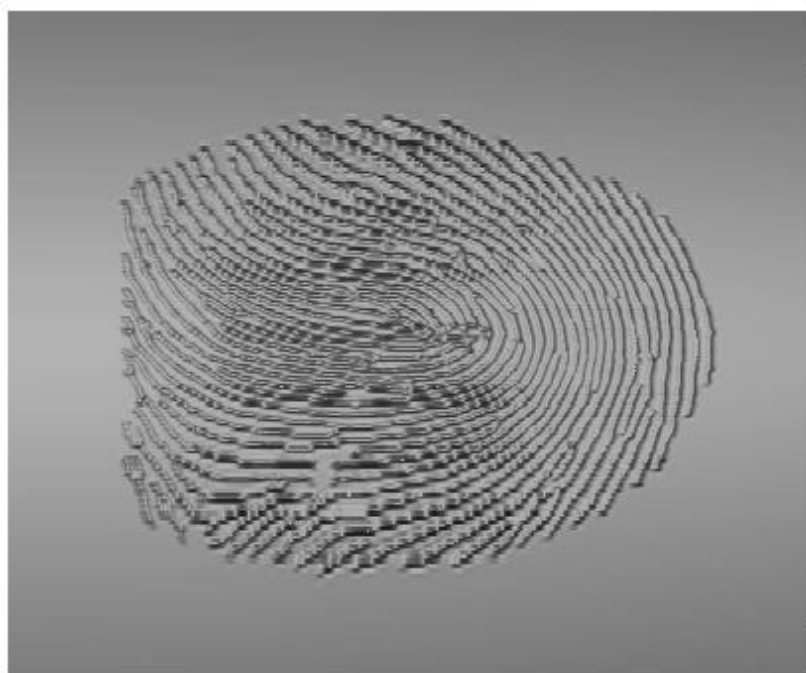
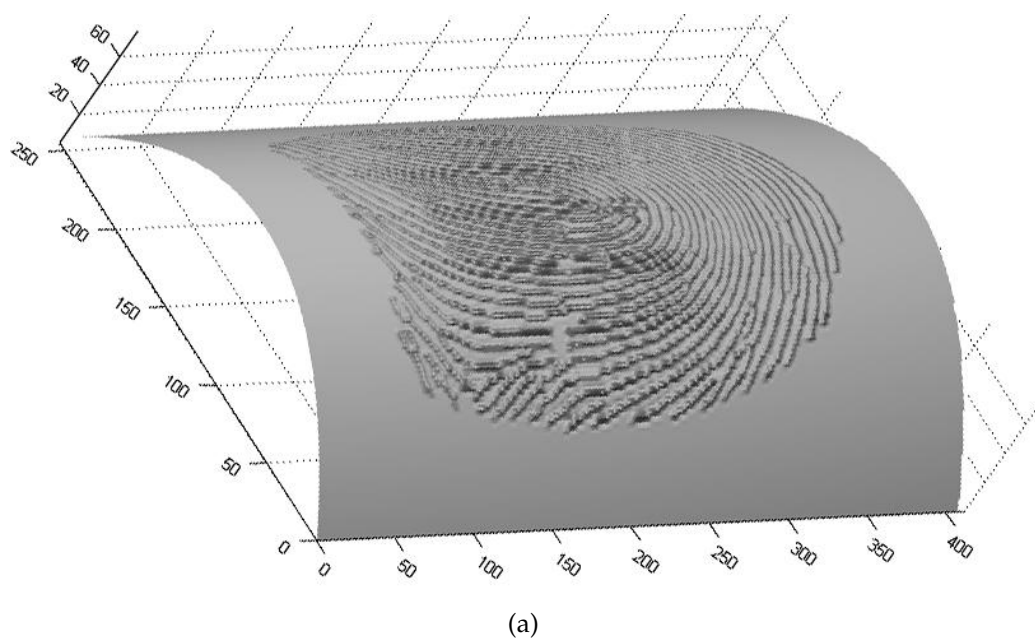


Figura 5.21: Ricostruzione tridimensionale con modello cilindrico ed unwrapping del modello (immagine del Dataset N.3): (a) modello dopo la fase di wrapping; (b) immagine dopo la fase di unwrapping

5.6 Accuratezza biometrica ottenuta

Im questa sezione sono riportati i risultati in termini di accuratezza dei dataset N.3 e N.4. I valori calcolati sono quindi confrontati con le prestazioni del sistema di riferimento, basato sul dataset N.2. Il dataset N.1, a causa del ridotto numero di immagini disponibile, non è stato preso in considerazione.

In Figura 5.22 sono mostrati i risultati dei test effettuati sul dataset N.3 di immagini sintetiche. L'EER in questo caso è pari all'11%.

In Figura 5.23 sono mostrati i risultati dei test effettuati sul dataset N.4 di immagini ottenute tramite unwrapping. I risultati sono simili a quelli ottenuti con il dataset N.3, il valore di EER è leggermente inferiore e pari al 9%.

In Figura 5.24 sono mostrati i risultati dei test effettuati sul dataset N.2 di immagini acquisite tramite contatto. Il valore di EER in questo caso è pari all'1.6%.

In Figura 5.25 sono mostrate le curve ROC dei tre dataset.

Da questi risultati si può osservare come il sistema biometrico basato sul dataset N.3 abbia prestazioni quasi equivalenti a quello basato sulle immagini del dataset N.4, seppur inferiori al sistema basato sul dataset N.2. I valori di errore permettono comunque di dire che il riconoscimento effettuato utilizzando le immagini dei due dataset è realizzabile. Entrambi i dataset però, presentano due problemi:

1. sono stati creati a partire dalle immagini del dataset N.2. Non sono presenti, quindi, le creste della regione laterale del dito. Tale regione è quella che maggiormente viene influenzata dall'operazione di unwrapping. Il dataset N.4 rappresenta solo in forma ridotta, quindi, un miglioramento rispetto al dataset N.3;
2. il contrasto tra creste e valli del dataset N.3 e del dataset N.4 è superiore a quello di un'immagine reale acquisita senza contatto. L'ac-

curatezza del confronto biometrico realizzato è quindi solo in parte rapportabile ad una situazione reale.

Per questi motivi risulta necessario integrare questi risultati con quelli ottenuti misurando l'accuratezza di immagini reali acquisite senza contatto.

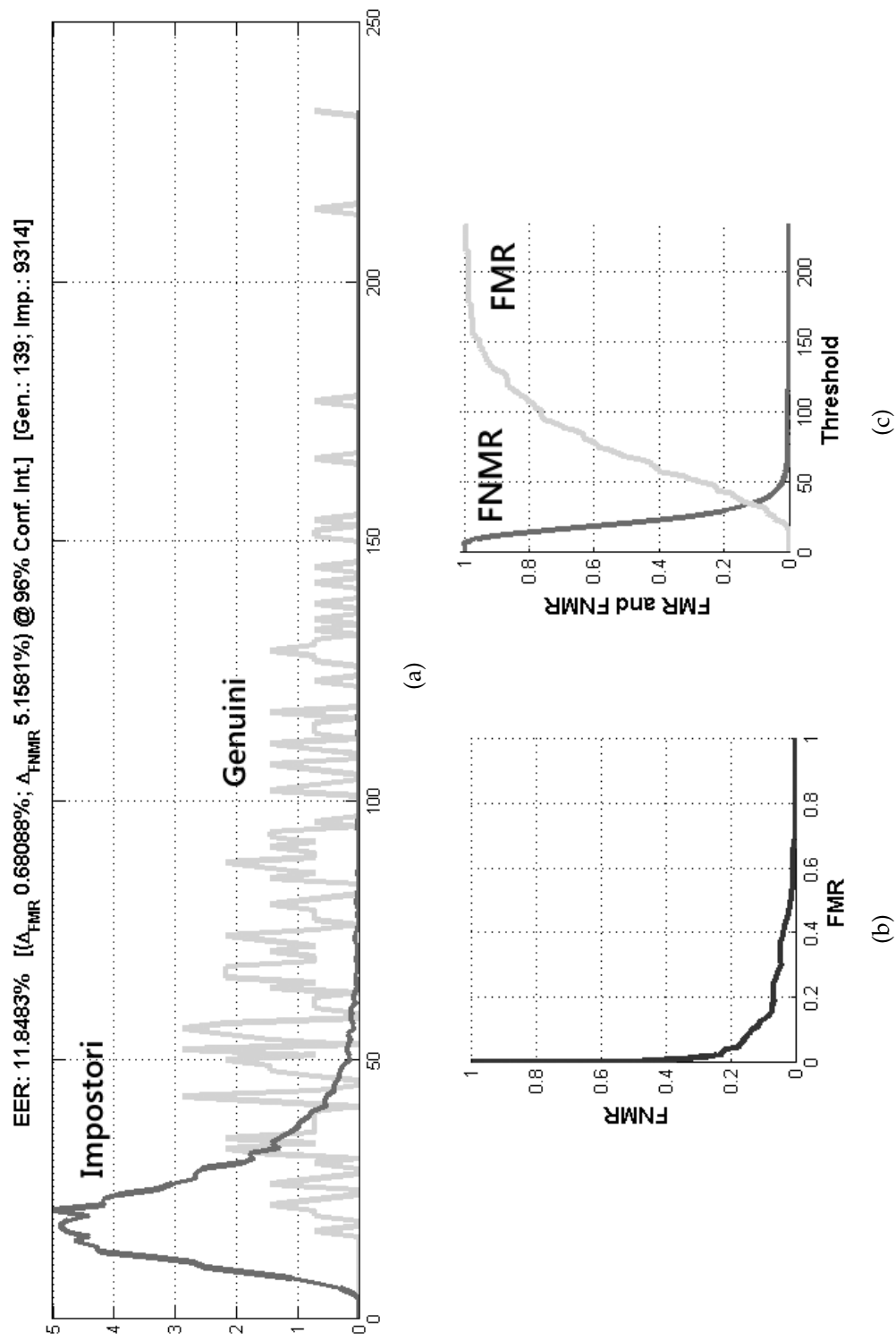


Figura 5.22: *Grafici del dataset N.3: (a) distribuzione degli impostori e dei genuini; (b) FNMR e FMR; (c) FNMR e FMR in funzione della soglia*

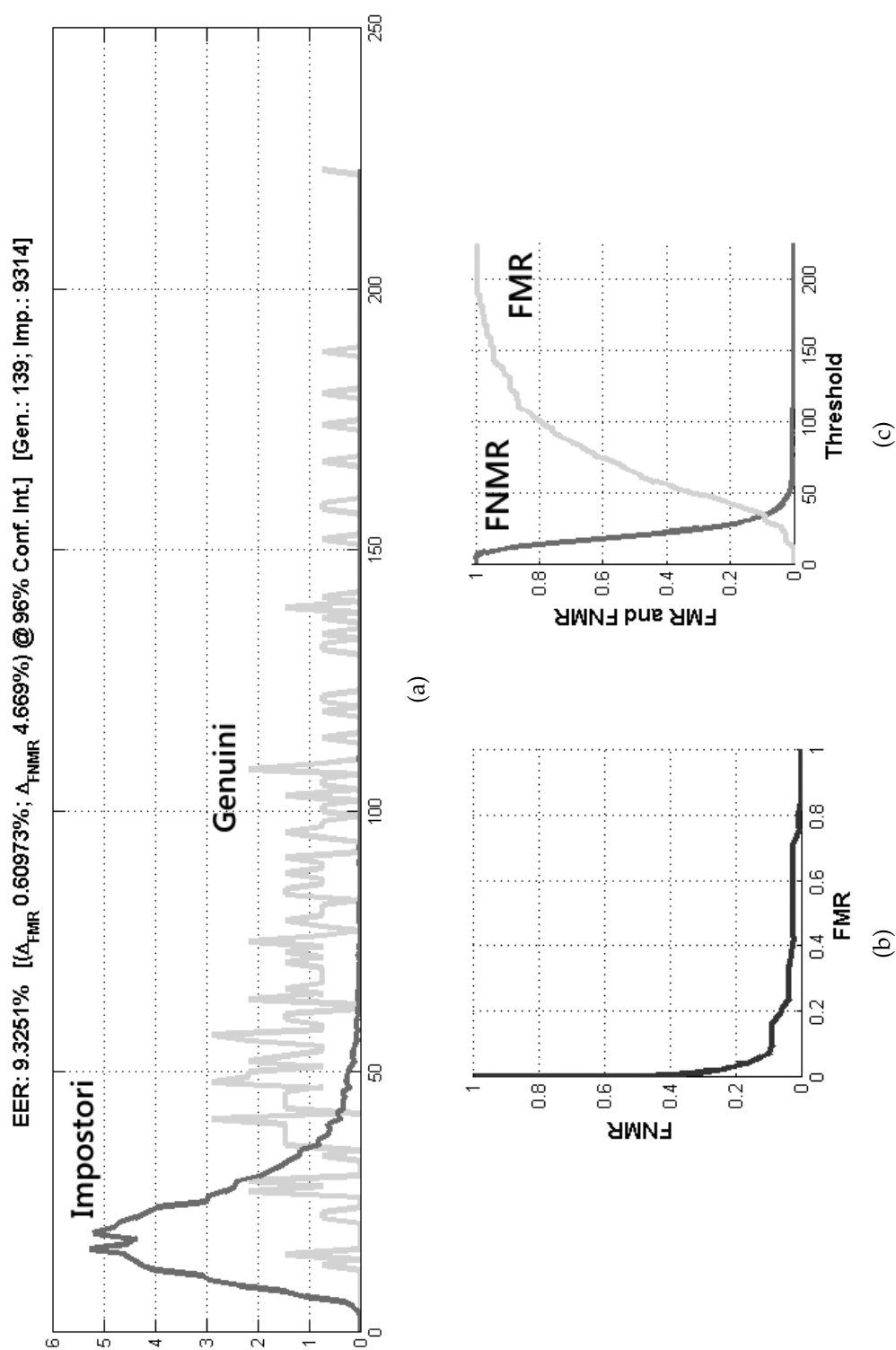


Figura 5.23: *Grafici del dataset N.4: (a) distribuzione degli impostori e dei genuini; (b) FNMR e FMR; (c) FNMR e FMR in funzione della soglia*

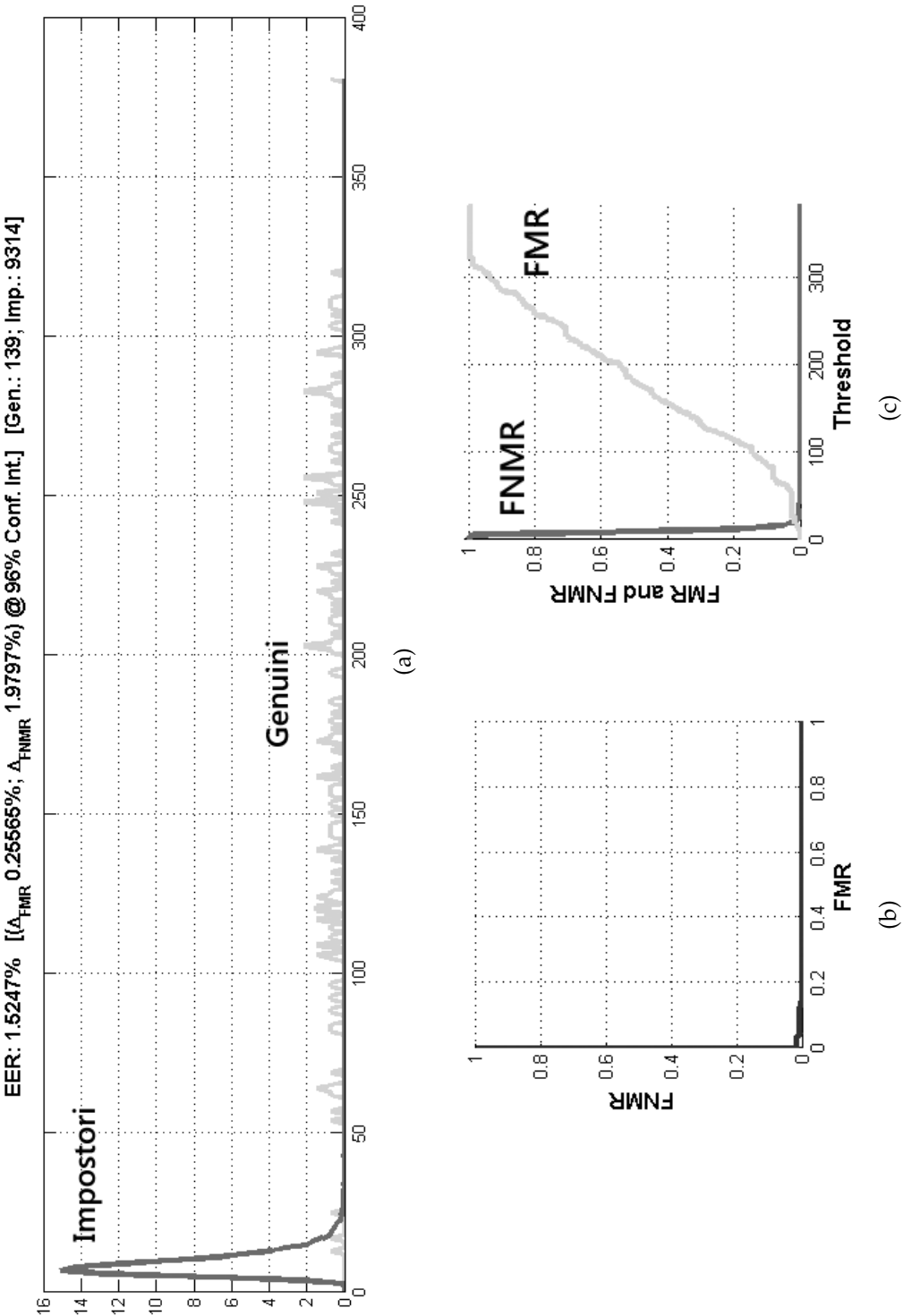


Figura 5.24: *Grafici del dataset N.2: (a) distribuzione degli impostori e dei genuini; (b) FNMR e FMR; (c) FNMR e FMR in funzione della soglia*

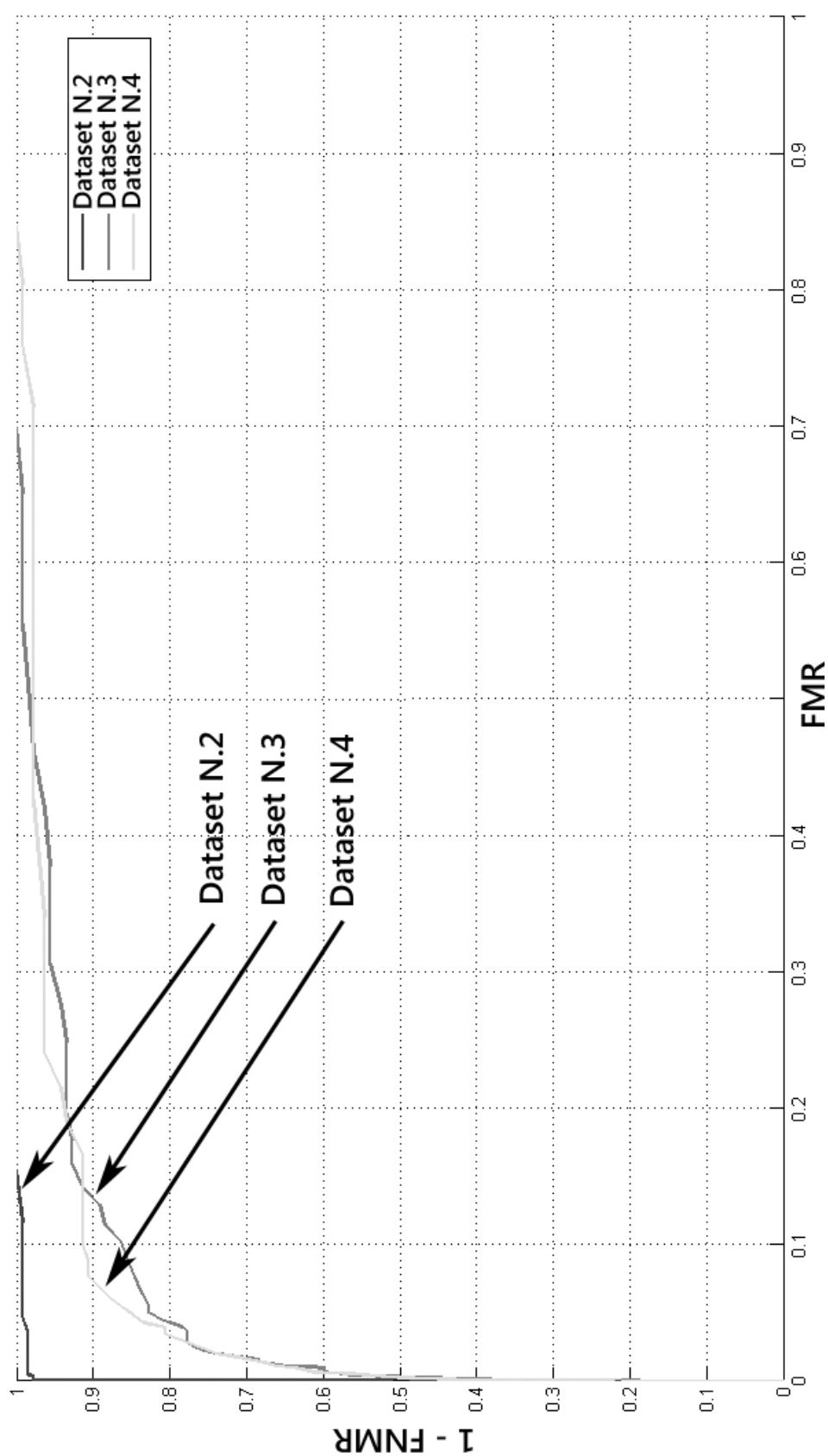


Figura 5.25: Curve ROC relative ai tre dataset

CONCLUSIONI E SVILUPPI FUTURI

Il calcolo di modelli tridimensionali del dito e dell'impronta digitale è un argomento di ricerca innovativo e le tecnologie ad esso correlate sono attualmente poco diffuse. Nonostante le tecniche di ricostruzione tridimensionale siano ormai presenti da anni, la loro applicazione ai sistemi biometrici presenta alcune problematiche, di cui la principale consiste nell'elevato livello di dettaglio richiesto. La maggior parte dei sistemi di ricostruzione tridimensionale, infatti, non è adatta all'elaborazione dei tratti biometrici. Risulta quindi necessario studiare ed impiegare tecniche innovative al fine di riuscire a rappresentare accuratamente i dettagli dell'impronta digitale in uno spazio tridimensionale.

In questo contesto si colloca il lavoro di tesi, che è consistito in uno studio preliminare dello stato dell'arte riguardante sia le tecniche di ricostruzione tridimensionale che le tecniche di acquisizione attuali di impronte digitali senza contatto. Ciò ha permesso di definire gli obiettivi del lavoro e di effettuare lo studio, la progettazione e l'implementazione di nuove tecniche di ricostruzione tridimensionale, applicabili a immagini di impronte digitali acquisite senza contatto.

In particolare, è stato realizzato un metodo di ricostruzione tridimensionale basato sulla creazione di un modello cilindrico. L'uso di un mod-

ello geometrico ha permesso di realizzare una sua mappatura su un piano bidimensionale, con la conseguenza che l'immagine risultante è stata confrontabile con immagini acquisite tramite un sensore dedicato.

I risultati dei confronti biometrici effettuati sulle immagini ottenute utilizzando questo metodo hanno fornito un'EER pari a 9%, superiore al valore di EER pari a 1.6% ottenuto utilizzando immagini tradizionali. Le immagini utilizzate, però, presentano caratteristiche differenti dalle immagini acquisite tramite sensore dedicato: hanno un contrasto tra creste e valli molto inferiore e necessitano dell'applicazione di tecniche di enhancement dell'immagine per poterle confrontare i dettagli biometrici. Considerando queste problematiche, i risultati sono comunque incoraggianti: è possibile effettuare il riconoscimento utilizzando queste immagini, e l'accuratezza presenta margini di miglioramento.

Il metodo di ricostruzione basato sulle minuzie, però, non è in grado di ottenere un modello visuale convincente del dito. Ciò è dovuto in prima istanza ai metodi di preprocessing dell'immagine, ancora non ottimizzati per la specifica applicazione studiata in questo lavoro di tesi. Secondariamente, l'elevata accuratezza richiesta dal calcolo della posizione tridimensionale richiede lo sviluppo di tecniche di corrispondenza stereoscopica maggiormente accurate.

L'accuratezza nel calcolo della posizione può essere aumentata effettuando più acquisizioni del dito nella stessa posizione e calcolando un'immagine media. In questo modo risulterebbe possibile compensare lo spostamento, anche minimo, causato dal movimento involontario della mano.

La corrispondenza stereoscopica, inoltre, deve essere resa più accurata anche ai fini della ricostruzione del modello parametrico: è necessario introdurre metodi in grado di determinare con precisione ed univocità la profondità dei punti del bordo e del punto di massima altezza del dito.

I metodi di ricostruzione studiati, inoltre, non sono in grado di realizzare una rappresentazione tridimensionale completa dell'andamento delle creste e delle valli. Tuttavia, il raggiungimento di questo obiettivo permet-

terebbe un'elaborazione più spinta ed introdurrebbe la possibilità di nuovi metodi di analisi delle impronte digitali e di riconoscimento biometrico.

Un'ulteriore possibile miglioramento è rappresentato dalla possibilità di integrare più acquisizioni senza contatto del dito, effettuate da diverse angolazioni. In questo modo la costruzione dei modelli tridimensionali potrebbe avvalorarsi di informazioni su un'area più estesa del dito. Integrando, per esempio, un'acquisizione del dito effettuata da una posizione centrale con due acquisizioni laterali, risulterebbe possibile compensare la diminuzione di qualità dell'immagine causata dalla curvatura del dito.

Una possibile prospettiva di ricerca è rappresentata, inoltre, dagli algoritmi di matching delle impronte digitali che operino nativamente sui modelli tridimensionali. Un sistema biometrico così progettato potrebbe effettuare il riconoscimento con un'accuratezza maggiore, ma non risulterebbe interfacciabile con i database esistenti. La sua applicazione rimarrebbe circoscritta, quindi ai soli sistemi impieganti questo tipo di rappresentazione biometrica.

Per quanto riguarda la parte di lavoro sulla generazione sintetica di immagini multiple di impronte digitali, è necessario studiare nuovi algoritmi per il miglioramento dell'operazione di calibrazione delle immagini stereoscopiche sintetiche e definire con maggiore accuratezza il punto di vista ed i parametri della fotocamera virtuale. Inoltre, ulteriore lavoro è necessario per riprodurre in modo più accurato le proprietà di riflettanza della pelle umana ed il contrasto tra creste e valli tipico di un'acquisizione reale.

Bibliografia

- [1] G. Bradski and A. Kaehler, *Learning OpenCV*. O'Reilly Media, September 2008.
- [2] S. Mattoccia, *Introduzione alla visione stereo*, Università di Bologna.
- [3] R. Y. Tsai, "A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE Journal of Robotics and Automation*, no. 3, pp. 323–344, 1987.
- [4] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330–1334, 2000.
- [5] H. Zollner and R. Sablatnig, "Comparison of methods for geometric camera calibration using planar calibration targets," *Pattern and Image Processing Group (PRIP), Vienna University of Technology*.
- [6] L. Bornaz, "L'analisi ed il trattamento dei dati laser scanner terrestri," Ph.D. dissertation, Geodesia e Geomatica.
- [7] C. Nitschke, *3D Reconstruction - Real-time Volumetric Scene Reconstruction from Multiple Views*. VDM Verlag Dr. Muller, April 2007.

- [8] A. Laurentini, "The visual hull concept for silhouette-based image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 150–162, 1994.
- [9] K. Shanmukh and A. Pujari, "Volume intersection with optimal set of directions," *Pattern Recognition Lett.*, pp. 165–170, 1991.
- [10] A. Laurentini, "How many 2d silhouettes does it take to reconstruct a 3d object?" *Computer Vision Image Understanding*, pp. 81–87, 1997.
- [11] H. Lin and J. Wu, "3d reconstruction by combining shape from silhouette with stereo," 2003, lin@ee.ccu.edu.tw, jr@image.ee.ccu.edu.tw.
- [12] S. Sinha and M. Pollefeys, "Multi-view reconstruction using photo-consistency and exact silhouette constraints: A maximum-flow formulation," *Department of Computer Science, University of North Carolina at Chapel Hill, USA*.
- [13] L. Shapiro and G. Stockman, *Computer Vision*. Prentice Hall, 2001.
- [14] F. Dellaert, S. Seitz, C. Thorpe, and S. Thrun, "Structure from motion without correspondence," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2000.
- [15] H. Wallach and D. O'Connell, "The kinetic depth effect," *Journal of Experimental Psychology*, no. 45, pp. 205–217, 1953.
- [16] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge University Press, August 2003, vol. 3, no. 7.
- [17] I. Point Grey Research, "Bumblebee®2."
- [18] J. Bouguet, "Camera calibration toolbox for matlab."
- [19] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1-2-3, pp. 7–42, April-June 2002.

- [20] F. Tombari, S. Mattoccia, and L. D. Stefano, "Segmentation-based adaptive support for accurate stereo correspondence," *IEEE Pacific-Rim Symposium on Image and Video Technology*, 2007.
- [21] D. Brown, "Close-range camera calibration," *Photogrammetric Engineering*, no. 37, pp. 855–866, 1971.
- [22] C. Harris and M. Stephens, "A combined corner and edge detector," *Proceedings of the 4th Alvey Vision Conference*, pp. 147–151, 1988.
- [23] K. Konolige, "Small vision system: Hardware and implementation," *Proceedings of the International Symposium on Robotics Research*, pp. 111–116, 1997.
- [24] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of artificial "gummy" fingers on fingerprint systems," *Proceedings of SPIE*, vol. 4677.
- [25] G. Parziale, "Touchless fingerprinting technology," *Advances in Biometrics*, pp. 39–62, 2007.
- [26] American National Standard for Information Systems, "Data format for the interchange of fingerprint facial, and other biometric information - part 1."
- [27] E. Sano, T. Maeda, T. Nakamura, M. Shikai, K. Sakata, M. Matsushita, and K. Sasakawa, "Fingerprint authentication device based on optical characteristics inside a finger," in *CVPRW '06: Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop*. Washington, DC, USA: IEEE Computer Society, 2006, p. 27.
- [28] V. Piuri and F. Scotti, "Fingerprint biometrics via low-cost sensors and webcams," *IEEE*, 2008.
- [29] B. Hiew, A. Teoh, and D. Ngo, "Preprocessing of fingerprint images captured with a digital camera."

- [30] Y. Chen, G. Parziale, E. Diaz-Santana, and A. Jain, "3d touchless fingerprints: Compatibility with legacy rolled images."
- [31] G. Parziale, E. Diaz-Santana, and R. Hauke1, "The Surround ImagerTM: A multi-camera touchless device to acquire 3d rolled-equivalent fingerprints," *Lecture Notes in Computer Science*, vol. Volume 3832, pp. 244–250, 2005.
- [32] G. Parziale and A. Niel, "Fingerprint matching using minutiae triangulation," *Institute of Digital Image Processing, Joanneum Research*.
- [33] Y. Wang, Q. Hao, A. Fatehpuria, L. Hassebrook, and D. Lau, "Data acquisition and quality analysis of 3-dimensional fingerprints," *IEEE*.
- [34] R. Gonzalez and R. Woods, *Digital Image Processing*, 3rd ed. Prentice Hall, 2008.
- [35] R. Guerchouche and F. Coldefy, "Camera calibration methods evaluation procedure for images rectification and 3d reconstruction," *Orange Labs, France Telecom R & D*.
- [36] A. Z. G. Loy, "Fast radial symmetry for detecting points of interest," *IEEE PAMI*, vol. 25, no. 8, pp. 959–973, 2003.
- [37] Nist Biometric Image Software, "<http://fingerprint.nist.gov/NFIS/>."
- [38] C. I. Watson, M. D. Garriss, E. Tabassi, C. L. Wilson, R. M. McCabe, S. Janet, and K. Ko, *User's Guide to Export Controlled Distribution of NIST Biometric Image Software (NBIS-EC)*, National Institute of Standards and Technology.
- [39] P. D. Kovesi, "MATLAB and Octave functions for computer vision and image processing," School of Computer Science & Software Engineering, The University of Western Australia, available from: <http://www.csse.uwa.edu.au/~pk/research/matlabfns/>.

- [40] G. Taubin, "Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations, with applications to edge and range image segmentation," *IEEE Trans. PAMI*, vol. 13, pp. 1115–1138, 1991.
- [41] J. D'Errico, "Gridfit," 2006.
- [42] P. D. Wasserman, "Solid-state fingerprint scanners - a survey of technologies," 2005.
- [43] R. Gonzalez, R. Woods, and S. Eddins, *Digital Image Processing Using MATLAB*, 2nd ed. Gatesmark Publishing, 2009.

Elenco delle figure

2.1	Pinhole model	7
2.2	Pinhole model semplificato	7
2.3	Distorsione radiale	10
2.4	Distorsione tangenziale	11
2.5	Esempio di pattern geometrico di calibrazione	12
2.6	Funzionamento della luce strutturata	13
2.7	Utilizzo di un pattern più complesso per l'analisi di superfici . .	14
2.8	Setup di base per l'utilizzo della luce strutturata	15
2.9	Un distanziometro ad impulsi. La distanza D dell'oggetto dal trasmettitore è proporzionale al tempo di percorrenza degli im- pulsi	16
2.10	Esempio di scanner laser: (a) scanner laser; (b) nuvola di punti; (c) modello tridimensionale	17
2.11	Intersezione di coni generalizzati	19
2.12	Visual Hull	20
2.13	Perdita di informazione relativa alla profondità	24
2.14	Usando due fotocamere l'informazione relativa alla profondità non viene persa	25
2.15	Vincolo epipolare	26
2.16	Esempio di sistema per l'acquisizione di immagini stereoscopiche	27

2.17	La relazione tra la distanza dell'oggetto e la disparity: (a) oggetti posizionati a varie distanze dalle fotocamere; (b) relazione tra distanza e disparity	28
2.18	Field of View. La forma dei campi visivi dipende dalla lunghezza focale	29
2.19	Area utile per il calcolo delle disparità	30
2.20	Rotazione R e traslazione T del sistema di riferimento dell'oggetto (con origine O) rispetto al sistema di riferimento della fotocamera (con origine C)	31
2.21	Sistema stereo (virtuale) dopo la rettificazione	33
2.22	Cambiamenti apportati dalla rettificazione: (a),(b) immagini prima della rettificazione; (c),(d) immagini dopo la rettificazione . .	34
2.23	Triangolazione a partire dal valore di disparità	37
2.24	Esempio di rappresentazione di una scena tramite level-set: (a),(b),(c) scena che varia nel tempo; (d),(e),(f) variazione dell'altezza del piano intersecante il volume tridimensionale	39
2.25	Mesh di triangoli	40
2.26	Esempio di calcolo di una disparity map partendo da immagini acquisite da un sistema stereoscopico: (a) immagine sinistra; (b) immagine destra; (c) disparity map: i pixel più chiari corrispondono a punti più vicini	41
2.27	Mapping come descritto da un'omografia	43
2.28	Ricerca delle corrispondenze	53
3.1	Confronto tra fingerprint in caso di dito secco: (a) fingerprint acquisita senza contatto; (b) fingerprint acquisita a contatto . . .	57
3.2	Immagini della medesima impronta: (a) immagine acquisita a distanza (b) Immagine acquisita tramite sensore a contatto	58
3.3	Struttura generale di un sistema biometrico basato su impronte digitali acquisite tramite contatto	58
3.4	Struttura generale di un sistema biometrico basato su impronte digitali acquisite senza contatto	59

3.5	Sistemi senza contatto basati su riflettanza	60
3.6	Spostamento delle creste in caso di luce non perpendicolare . . .	61
3.7	Rappresentazione semplificata degli strati della pelle	61
3.8	Sistemi basati su trasmittanza	63
3.9	Immagine prodotta da un sistema basato su trasmittanza	63
3.10	Processo di elaborazione di un impronta digitale acquisita senza contatto	64
3.11	Processo di elaborazione basato su filtro passa banda	66
3.12	Processo di elaborazione basato su trasformata di Fourier	67
3.13	Esempio di estrazione delle minuzie da immagine acquisita a distanza (a) Immagine originale (b) Minuzie rilevate e sovrain- poste all'immagine elaborata	69
3.14	Struttura di un sistema di riconoscimento basato su modello tridimensionale dell'impronta digitale	70
3.15	Unwrapping parametrico usando un cono: (a) il punto (x,y,z) è mappato sul punto (θ,z) ; (b) la distanza tra i punti sulla superfi- cie del dito viene modificata in modo non proporzionale	71
3.16	Confronto tra modello tridimensionale dopo unwrapping non parametrico e immagine ottenuta da un sensore a contatto: (a) modello tridimensionale dopo unwrapping; (b) immagine ot- tenuta da un sensore a contatto	72
3.17	Il sistema a specchio circolare	73
3.18	Sistema a scansione circolare	74
3.19	Il Surround Imager TM della TBS Holding AG	76
3.20	Esempio di ricostruzione tridimensionale	77
3.21	Rappresentazione tridimensionale di una minuzia con coordi- nate (x,y,z) e orientamento (ϕ,θ)	77
3.22	Confronto tra un'immagine classica e un'immagine ottenuta da un modello tridimensionale: (a) immagine acquisita con tec- niche classiche; (b) modello tridimensionale; (c) immagine ot- tenuta da unwrapping e binarizzazione	79

3.23	Un'immagine dell'impronta digitale in scala di grigi è rappresentabile come modello tridimensionale	80
4.1	Confronto tra due acquisizioni della stessa scacchiera: (a) tutti i corner sono riconosciuti e evidenziati dalle linee che li collegano; (b) un corner (evidenziato in basso a destra) non è stato riconosciuto	83
4.2	Schema a blocchi dell'algoritmo di binarizzazione adottato per le immagini delle scacchiere di calibrazione	84
4.3	Istogramma dell'immagine di una scacchiera	85
4.4	Confronto tra il riconoscimento degli angoli sull'immagine originale e binarizzata: (a) non tutti i corner sono riconosciuti; (b) tutti i corner sono riconosciuti correttamente e evidenziati dalle linee che li collegano	86
4.5	Insieme di corner della scacchiera di calibrazione ricostruiti nello spazio tridimensionale e il corrispondente piano interpolante: (a) i punti nello spazio tridimensionale; (b) il piano interpolante .	88
4.6	Schema a blocchi della ricostruzione tramite triangolazione delle minuzie	92
4.7	Schema a blocchi del procedimento di segmentazione dell'immagine	94
4.8	Estrazione del bordo e della maschera binaria di segmentazione: (a) immagine originale; (b) bordo del dito; (c) maschera binaria di segmentazione; (d) immagine segmentata	95
4.9	Schema a blocchi del calcolo della regione di interesse	97
4.10	Posizione dell'ultima giunzione del dito	98
4.11	Limitazione della ROI dell'immagine: (a) maschera di segmentazione originale; (b) maschera di segmentazione limitata alla punta del dito; (c) maschera di segmentazione limitata ed erosa .	99
4.12	Differenti tipi di preprocessing delle immagini: (a) Metodo A; (b) Metodo B; (c) Metodo C; (d) Metodo D	101
4.13	I principali tipi di minuzie	102

4.14 Risultato dell'estrazione delle minuzie	103
4.15 Le coordinate delle minuzie ottenute tramite il software <code>mindtct</code> non corrispondono sempre alla loro posizione effettiva: (a) po- sizione originale; (b) posizione raffinata	104
4.16 Matching delle minuzie e raffinamento dei risultati: (a) minuzie nell'immagine sinistra; (b) minuzie nell'immagine destra; (c) minuzie raffinate nell'immagine sinistra; (d) minuzie raffinate nell'immagine destra;	109
4.17 Matching delle minuzie basato sulla vicinanza: (a) minuzie nel- l'immagine sinistra; (b) minuzie nell'immagine destra	110
4.18 Triangolazione del bordo del dito e delle minuzie corrisponden- ti: (a) vista dall'alto; (b) vista laterale	111
4.19 La posizione ed il tipo delle minuzie risultano modificate tra le due immagini: (a) due minuzie nell'immagine sinistra; (b) le due minuzie corrispondenti nell'immagine destra	112
4.20 Schema a blocchi dell'algoritmo per il calcolo di un modello tridimensionale cilindrico	114
4.21 Mappa degli orientamenti delle creste: (a) variazione dell'orien- tamento delle creste; (b) zona di massima curvatura	115
4.22 Posizione del punto di massima curvatura	117
4.23 Calcolo del modello cilindrico del dito: (a) i tre punti princi- pali; (b) la posizione dei tre punti nel dito; (c) interpolazione con spline e centro della circonferenza interpolante; (d) interpo- lazione della superficie del modello esteso	121
4.24 Spiegazione del procedimento di wrapping: (a) proiezione del- l'oggetto sull'immagine; (b) proiezione dell'immagine sul mod- ello	122
4.25 Tecnica di wrapping tramite analisi monodimensionale	122
4.26 Wrapping dei dettagli sul modello cilindrico: (a) wrapping delle minuzie; (b) wrapping del pattern di creste	123
4.27 Zone dell'immagine escluse dall'operazione di wrapping	123

4.28 Esempio di risultato dell'operazione di wrapping dell'immagine sul modello cilindrico	124
4.29 Calcolo delle coordinate polari del modello cilindrico	126
4.30 Modello cilindrico delle minuzie dopo la fase di unwrapping	127
4.31 Unwrapping di immagini applicate al modello cilindrico del dito: (a) l'immagine prima dell'interpolazione; (b) l'immagine dopo l'interpolazione	128
4.32 Schema a blocchi del metodo di calcolo di un modello tridimensionale sintetico	130
4.33 Calcolo dell'altezza delle creste	132
4.34 Modello tridimensionale sintetico del dito: (a) immagine acquisita con un sensore tradizionale; (b) modello tridimensionale sintetico basato sull'immagine (a)	133
4.35 Confronto tra setup stereoscopico reale e setup stereoscopico virtuale: (a) situazione reale; (b) situazione virtuale	134
4.36 Setup stereoscopico virtuale	135
4.37 Simulazione di acquisizione stereoscopica del modello sintetico del dito: (a) immagine sinistra; (b) immagine destra	136
4.38 Modello tridimensionale sintetico del pattern di calibrazione	137
4.39 Rotazione del pattern di calibrazione	140
4.40 Acquisizione stereoscopica del pattern di calibrazione: (a) immagine sinistra; (b) immagine destra	141
4.41 Segmentazione dell'immagine ottenuta dal modello sintetico: (a) immagine originale; (b) maschera di segmentazione; (c) maschera della regione di interesse	142
5.1 Due coppie di immagini del dataset N.1: (a),(c) immagini acquisite tramite la vista di sinistra; (b),(d) immagini acquisite tramite la vista di destra	145
5.2 Esempi di immagini del dataset N.2: le immagini (a),(b),(c),(d) appartengono a persone diverse	146

5.3	Esempi di immagini del dataset N.3: (a) immagine sinistra; (b) immagine destra	148
5.4	Esempi di immagini del dataset N.4: (a) immagine sinistra; (b) immagine destra	149
5.5	Deviazione standard dell'errore del dataset N.1	150
5.6	Deviazione standard dell'errore del dataset N.3 sintetico: (a) spostamento Δx_1 ; (b) spostamento Δx_2	151
5.7	Ricostruzione dell'immagine N.8 del dataset N.1	152
5.8	Ricostruzione dell'immagine N.2 del dataset N.3 con Δx_2	152
5.9	Frequenza cumulata della deviazione standard del dataset N.1	153
5.10	Frequenza cumulata della deviazione standard del dataset N.3 con Δx_1	153
5.11	Frequenza cumulata della deviazione standard del dataset N.3 con Δx_2	154
5.12	Confronto tra le tecniche di preprocessing: (a),(b) Metodo A; (c),(d) Metodo B; (e),(f) Metodo C; (g),(h) Metodo D	157
5.13	Valori di match score ottenuti con il metodo N.1	159
5.14	Frequenza cumulata dei match score ottenuti con il metodo N.1	160
5.15	Valori di match score ottenuti con il metodo N.2	160
5.16	Frequenza cumulata dei match score ottenuti con il metodo N.2	161
5.17	Matching stereoscopico di una coppia di immagini del Dataset N.3: (a),(b) metodo N.2; (c),(d) metodo N.3	162
5.18	Triangolazione delle minuzie di una coppia di immagini del Dataset N.1	163
5.19	Triangolazione delle minuzie di una coppia di immagini del Dataset N.3	164
5.20	Ricostruzione tridimensionale con modello cilindrico ed unwrapping del modello (immagine del Dataset N.1): (a) modello dopo la fase di wrapping; (b) immagine dopo la fase di unwrapping	165

5.21 Ricostruzione tridimensionale con modello cilindrico ed unwrapping del modello (immagine del Dataset N.3): (a) modello dopo la fase di wrapping; (b) immagine dopo la fase di unwrapping .	166
5.22 Grafici del dataset N.3: (a) distribuzione degli impostori e dei genuini; (b) FNMR e FMR; (c) FNMR e FMR in funzione della soglia	169
5.23 Grafici del dataset N.4: (a) distribuzione degli impostori e dei genuini; (b) FNMR e FMR; (c) FNMR e FMR in funzione della soglia	170
5.24 Grafici del dataset N.2: (a) distribuzione degli impostori e dei genuini; (b) FNMR e FMR; (c) FNMR e FMR in funzione della soglia	171
5.25 Curve ROC relative ai tre dataset	172

Elenco delle tabelle

4.1	Esempio di righe della tabella di compatibilità analizzate dal software bozorth3	106
5.1	Risultanti del matching tra l'immagine sinistra e l'immagine destra del dataset N.1	155

Indice analitico

- acquisizione, 2, 4, 6, 9, 11, 14, 16,
26–29, 55, 56, 58, 59, 62, 73,
75, 79, 82, 83, 91, 129, 130,
132, 134, 136, 138, 141, 144,
145, 150, 156, 164, 173–175
- asse ottico, 8
- baseline, 27, 28, 133, 144
- binarizzata, 65, 67, 78, 82, 85, 86,
93, 95, 96, 100, 139
- binarizzazione, 65, 79, 83–85, 99,
101, 154, 155
- biometria, 1, 2, 55, 57–59, 62, 68,
143, 159, 167, 175
- bordo, 3, 9, 21, 51, 52, 61, 65, 75,
85, 91–93, 95, 102, 107, 110,
111, 113, 116, 117, 174
- Bouguet, 33, 40
- bozorth3, 68, 103, 106–108, 158
- Brown, 40
- calibrazione, 3, 4, 6, 10–12, 22, 23,
29, 34, 39, 40, 42, 45, 46, 49,
73, 81, 82, 84, 86, 88, 129,
137, 140, 141, 144, 150, 175
- capacitivo, 57
- centro ottico, 24
- coni generalizzati, 18, 19
- contatto, 2–4, 55–60, 62, 64, 67, 68,
72, 79, 127, 129–131, 139,
144, 150, 163, 175
- contrasto, 52, 59, 74, 102, 167, 174,
175
- coordinate cartesiane, 125, 131
- coordinate omogenee, 8, 45, 54
- coordinate polari, 125, 126, 131
- corner, 11, 40, 42–44, 46, 50, 82, 83,
86–88, 144
- correlazione, 22, 23, 35, 75, 77, 91,
103
- corrispondenza, 21, 23, 26, 35, 36,
41, 46, 51–53, 75, 89, 91, 96,
103, 105, 106, 108–110, 113,
119, 125, 155, 159, 163, 174
- curvatura, 3, 61, 62, 68, 111, 113–

- 118, 125, 175
- dataset, 3, 129, 137, 144–146, 148–155, 158, 162–167, 169–172
- deformazione, 2, 13, 14, 56, 68, 78, 127
- denso, 22, 35
- Depth of Focus, 59, 61
- dermatoglifo, 55, 68, 130
- deviazione standard, 89, 139, 150, 151, 153, 154, 158
- Direct Linear Transformation, 11
- disparity, 27, 28, 35, 38, 39, 41, 53, 54
- disparity map, 39, 41, 51
- distanziometro, 16
- distorsione, 9, 10, 12, 33, 40, 42, 46, 48, 49, 51, 72, 75
- distorsione radiale, 9, 10, 12
- distorsione tangenziale, 10
- distribuzione, 83, 169–171
- elemento strutturante, 96, 116
- enhancement, 63, 65, 66, 174
- epipolare, 25, 26, 33
- estrinseci, 12, 33, 40, 42, 46, 48
- fast radial, 98
- Field of View, 28, 60
- filtro mediano, 92
- FMR, 169–171
- FNMR, 169–171
- focale, 6, 8, 29, 54, 110, 144
- fonte di luce, 132, 134
- fonti di luce, 59, 60
- Fourier, 66, 67, 97
- frequenza, 3, 15, 60, 61, 65, 67, 97, 100, 107, 152–154, 158, 160, 161, 164
- genuini, 156, 169–171
- giunzione, 93, 94, 96, 120, 139
- gradiente, 21, 43, 94, 98, 115, 116
- impostori, 169–171
- impronta digitale, 1–4, 55–59, 62, 64, 68–71, 74, 76, 79–81, 91, 93, 100, 103, 104, 112, 113, 129, 131, 132, 139, 144, 145, 158, 159, 173, 175
- impronta latente, 56
- interpolazione, 51, 87–89, 117, 121, 125, 126, 128, 131, 152
- intrinseca, 9, 10, 12, 33, 40, 42, 45, 46, 48, 53
- iper-stereo, 27
- Kinetic Depth Effect, 21
- lama laser, 5, 13, 15, 16
- larghezza, 29, 65, 96, 113, 116, 120, 126, 129, 137
- laser scanner, 15, 16
- level-sets, 38
- linescan, 73
- logaritmico-polare, 72

- luce strutturata, 2, 5, 13, 15, 78
Lucy-Richardson, 65
lunghezza d'onda, 60, 62

mesh, 5, 17, 38, 40, 77, 118
mindtct, 68, 100, 102, 104
minuzia, 3, 56, 68, 69, 71, 76–78,
81, 91, 92, 100, 102–113, 118,
119, 123, 126, 127, 155, 158,
159, 163, 164, 174
Mitsubishi, 62
modello tridimensionale, 4, 5, 13,
17, 20, 21, 53, 69, 70, 72, 74,
76, 78–82, 89, 91, 110, 113,
114, 116, 125, 127, 129–134,
136, 137, 139, 146, 163
morfologico, 85, 93, 96, 112, 116

NBIS, 68, 100, 104
NIST, 100, 103
nuvola di punti, 5, 16, 17, 76, 118,
119, 131

omografia, 40, 42, 43, 45, 46
OpenCV, 40, 42, 44, 49–51, 89, 143
operatore di Roberts, 92, 93
orientamento, 3, 11, 64, 65, 77, 98,
100–102, 104, 105, 108, 114–
116
ottico, 8, 24, 57, 145
otturatore, 60
parametrico, 70–72, 76, 78, 174
passa-banda, 65, 96
pattern, 3, 10, 12–14, 62, 73, 78,
102, 104, 118, 123, 129, 130,
137, 140, 141
Phase Measuring Philometry, 78
Photo Hull, 21
piano dell'immagine, 6, 8–10, 24
pinhole, 6–9, 48
preprocessing, 62, 65, 66, 91, 96,
98, 99, 101, 102, 112, 143,
155, 157, 174
proiettiva, 8
punto di massima curvatura, 3, 113–
117
punto notevole, 3
punto principale, 8, 54
punto singolare, 74, 100

raffinamento, 12, 107, 109, 155, 158
regione di interesse, 3, 64, 91, 93,
96, 97, 142, 155
registrazione, 64, 65
rettificazione, 23, 33, 34, 40, 50–52,
54, 75
RGB, 65, 66
ricostruzione tridimensionale, 2, 5,
16, 20, 23, 40, 77, 87–89,
129, 139, 146, 154, 158, 163,
165, 166, 173
riflessione, 59
risoluzione, 62, 75, 100, 131, 144,
145

- ROC, 172
- rotazione, 30–32, 34, 45, 46, 49, 50, 103, 109, 138, 140
- segmentazione, 18, 64, 65, 67, 81, 91–96, 99, 107, 116, 139, 142
- senza contatto, 3, 55–57, 59, 60, 62, 64, 67, 68, 79, 127, 129, 144, 150, 163, 175
- Shape from photo-consistency, 20, 21, 23
- shape from silhouette, 18–21, 23
- sintetico, 5, 82, 129, 130, 133, 135–139, 142, 146, 151
- skeleton, 103, 112
- smoothness assumption, 36
- software di matching, 68, 91, 103
- soglia, 44, 67, 83, 84, 95, 107, 109, 139, 151, 153, 154, 169–171
- spazio virtuale, 37, 132, 134, 136
- specchio ad anello, 72
- spline, 117, 121
- stato dell'arte, 4, 173
- stereoscopia, 24, 26, 29, 33, 41, 50, 54, 74, 82, 86, 107, 132, 134, 135, 138, 146, 159, 162
- Structure from motion, 21, 23
- Surround Imager, 74, 76
- tabella di compatibilità, 104–106
- TBS, 73, 74, 76
- tempo di esposizione, 62
- thinning, 93
- tradizionale, 1, 4, 57, 127, 129–131, 133, 138, 143, 144
- traslazione, 30–32, 34, 45, 49, 50, 53, 103, 109, 110
- trasmissione, 62, 63
- triangolazione, 3, 5, 22, 24, 36–38, 76, 92, 110, 111, 116, 119, 158, 163, 164
- Tsai, 11, 12, 34
- Unsharp masking, 65
- unsharp masking, 65
- Unwrapping, 69–72, 75, 76, 78, 79, 125–128, 139, 143, 144, 164–166
- unwrapping, 69–72, 75, 76, 78, 79, 125–128, 139, 143, 144, 164–166
- viste multiple, 3, 4, 18, 69, 72
- Visual Hull, 18, 20
- voxel, 38, 131
- Wiener, 65
- wrapping, 113, 118, 119, 122–125, 130, 164–166
- YUV, 65
- Zhang, 11, 12, 34, 40

Ringraziamenti

Ringrazio tutte le persone che mi in questi anni mi hanno aiutato e mi hanno dato il loro supporto, in particolare i miei genitori, i miei amici e e tutti gli altri che hanno contribuito alla realizzazione di questa tesi. Ringrazio il mio relatore e il correlatore per l'aiuto che mi hanno sempre offerto.

Grazie a tutti quelli che mi hanno aiutato a concentrarmi nello studio, a quelli che hanno sempre provato a fare il contrario, e a chi mi ha dato un obiettivo da superare. Un grazie a tutti quelli che mi hanno aiutato a passare il tempo in questi anni di università.

Una menzione anche, infine, per chiunque stia leggendo questa tesi.

