

**ALMA MATER STUDIORUM-UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA**

SCUOLA DI INGEGNERIA E ARCHITETTURA
Corso di Laurea Magistrale in Ingegneria e Scienze Informatiche

TESI DI LAUREA
IN TECNICHE AVANZATE PER L'ANALISI DELLE IMMAGINI E VISIONE

**Ricostruzione 3D di oggetti ad alto potere diffusivo mediante
triangolazione laser**

Candidato:
Roberto Togni
Matricola 681165

Relatore:
Prof. Alessandro Bevilacqua

Correlatore:
Ing. Alessandro Gherardi

Indice

Introduzione	1
1 Tecniche di acquisizione 3D	3
1.1 Metodi ottici attivi	4
1.1.1 Triangolazione laser	4
1.1.2 Time-of-flight	6
1.1.3 Triangolazione a luce strutturata	8
1.2 Metodi ottici passivi	11
1.2.1 Triangolazione mediante coppia di telecamere stereo . .	12
2 Analisi del problema	15
2.1 Contesto del problema	15
2.2 Il problema della triangolazione a lama laser	17
2.2.1 Problematiche nel rilevamento del profilo laser	18
2.2.2 Geometria del sistema di triangolazione	18
3 Progettazione	23
3.1 Richiami teorici	23
3.1.1 Modello di proiezione prospettica della telecamera . . .	23
3.1.2 Distorsioni ottiche	24
3.2 Progettazione hardware	25
3.2.1 Telecamere utilizzate	26
3.2.2 Proiettore laser utilizzato	28
3.2.3 Sistema di attuazione	29
3.2.4 Configurazione geometrica del sistema di triangolazione	31

3.2.5	Architettura hardware del sistema di triangolazione . . .	36
3.3	Progettazione software	37
3.3.1	Calibrazione della telecamera	38
3.3.2	Triangolazione del profilo laser	42
3.3.3	Software e librerie utilizzate	44
3.3.4	Architettura della libreria software	45
4	Implementazione	49
4.1	Implementazione del sistema di trascinamento	49
4.2	Implementazione della libreria software	53
4.2.1	Implementazione tecniche di calibrazione della telecamera	53
4.2.2	Implementazione tecniche di rilevamento del profilo laser	57
5	Risultati sperimentali	63
5.1	Risultati della calibrazione delle telecamere	63
5.1.1	Risultati calibrazione mvBluefox-121C	63
5.1.2	Risultati calibrazione Sick Ranger E50	66
5.2	Risultati della calibrazione del sistema di trascinamento	66
5.3	Risultati della triangolazione	67
5.3.1	Test con pezzi non diffusivi	67
5.3.2	Test con pezzi diffusivi	73
	Conclusioni e sviluppi futuri	79
	Bibliografia	81

Introduzione

L'obiettivo principale che questa tesi intende perseguire è la progettazione di un sistema di acquisizione 3D a basso costo basato sul metodo di triangolazione a lama laser. A tale scopo si adotterà una telecamera *general-purpose*, e si provvederà sia alla realizzazione di un prototipo di sistema di attuazione per automatizzare la scansione dell'oggetto, sia all'implementazione di algoritmi per la calibrazione della geometria laser-telecamera, il rilevamento del profilo laser nell'immagine e la sua successiva ricostruzione 3D, tenendo anche conto delle complicazioni che l'uso di oggetti ad alto potere diffusivo apporta a tale problema, a causa dei marcati effetti di *subsurface scattering* della luce laser. Il sistema di triangolazione sarà validato caratterizzando, mediante il confronto con pezzi di dimensioni note, la precisione e l'accuratezza delle misurazioni, e valutando i tempi di esecuzione degli algoritmi. Inoltre, le prestazioni ed i risultati delle acquisizioni saranno messi a confronto con quelli ottenuti adottando una telecamera *high-end* dedicata. Lo studio preliminare svolto in questa sede è propedeutico per la futura realizzazione di un sistema per la ricostruzione in camera bianca di protesi d'osso su misura.

Capitolo 1

Tecniche di acquisizione 3D

Negli ultimi anni si è assistito ad un notevole incremento della diffusione delle tecniche di acquisizione 3D, in larga parte favorito dalla disponibilità di hardware con velocità di elaborazione sempre maggiore e di dimensioni contenute. Attualmente questi sistemi vengono impiegati in svariati settori applicativi, fra cui quello industriale (controllo della qualità, *reverse engineering*), medico (ricostruzione di tessuti e organi), della robotica (*scene understanding*) e della conservazione dei beni culturali (rilievi di sculture, monumenti e siti archeologici). La maggior parte delle tecniche di acquisizione ha come obiettivo la ricostruzione di un modello numerico tridimensionale della superficie dell'oggetto esaminato, o in alternativa l'estrazione di misure e caratteristiche utili [3]. Una tassonomia delle tecnologie impiegate è presente in [6]. Esse sono suddivise in due macrocategorie: a contatto e non a contatto. Un'importante sottoclasse della seconda è la famiglia dei metodi ottici, oggetto di questo capitolo, che sfruttano la radiazione ottica (in particolare lo spettro della luce visibile e infrarossa) per l'acquisizione di dati tridimensionali. Tali tecniche costituiscono la soluzione maggiormente adottata per via del costo più ridotto e della flessibilità d'uso. I metodi ottici si distinguono a loro volta in attivi e passivi.

1.1 Metodi ottici attivi

I sistemi di acquisizione basati su tecnologie a illuminazione attiva sono costituiti da un emettitore che illumina un punto, una sezione o l'intero oggetto di interesse, e da un sensore che riceve parte della luce riflessa dalla superficie dell'oggetto [6]. I tre metodi ottici attivi maggiormente utilizzati nel contesto dell'acquisizione 3D sono: triangolazione laser, *time-of-flight* e triangolazione a luce strutturata.

1.1.1 Triangolazione laser

Il principio di funzionamento di un sistema di acquisizione 3D basato su triangolazione laser è mostrato nella Figura 1.1. Un fascio di luce laser viene proiettato sull'oggetto da osservare da una determinata posizione, mentre il sensore (tipicamente CCD/CMOS) rileva la componente riflessa. Nota la geometria della coppia laser-sensore, si riesce così a triangolare la posizione di un punto della superficie nel sistema di riferimento adottato [7]. Nel caso della Figura 1.1, l'origine del sistema di riferimento è collocata nel laser, il quale proietta un fascio orientato di un angolo θ rispetto all'asse Z . Il sensore è posizionato ad una distanza d , detta *baseline*, e l'immagine del raggio laser riflesso dalla superficie dell'oggetto si forma ad una distanza p dal punto principale. La coordinata x del punto P si può ottenere come $x = z \tan \theta$, mentre, osservando che $\tan \beta = \frac{z}{d-x} = \frac{f_0}{p}$, la coordinata z può essere invece ricavata come $z = \frac{df_0}{p+f_0 \tan \theta}$.

L'illuminazione mediante laser risulta di notevole utilità per via delle sue proprietà: è infatti in grado di proiettare un fascio di luce coerente, monocromatica e collimata ad elevata luminosità. Inoltre, possono essere utilizzati filtri ottici passa-banda per isolare la luce laser in presenza di altre sorgenti luminose, che potrebbero disturbare l'acquisizione. Ciò semplifica la segmentazione del *laser spot*, ovvero il rilevamento dell'insieme dei pixel che costituiscono l'immagine del fascio riflesso dall'oggetto. L'illuminazione laser risulta anche estremamente versatile, in quanto il raggio può essere orientato in maniera controllata tramite l'uso di un sistema di specchi rotanti in modo da riuscire a scansionare una superficie. In alcuni scanner, per velocizzare

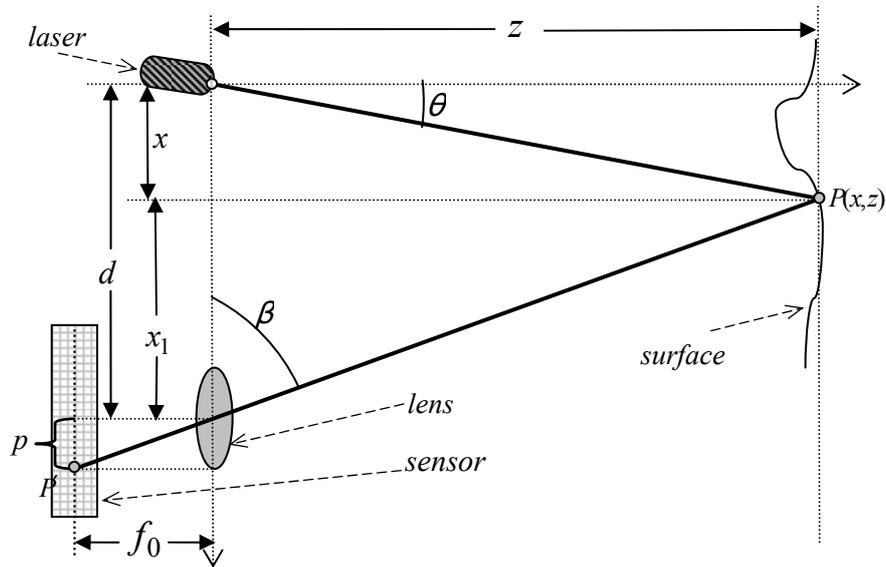


Figura 1.1: Un esempio di geometria laser-sensore adottata per la triangolazione. Figura tratta da [17].

il processo di acquisizione, la scansione viene effettuata proiettando una linea laser che identifica una sezione trasversale dell'oggetto, piuttosto che un singolo punto; in questo caso si parla di triangolazione a lama laser [7].

Fra gli svantaggi derivanti dall'uso di questo tipo di illuminazione vi sono i potenziali rischi per l'occhio umano dovuti all'elevata potenza del laser. È dunque importante valutare con attenzione quale classe di sicurezza adottare a seconda dei casi applicativi [2]. Inoltre, il grado di accuratezza ottenibile con la triangolazione laser dipende strettamente dal tipo di materiale illuminato: per un corretto rilevamento della posizione del laser sul sensore, l'utilizzo di questa tecnica è particolarmente indicato per superfici opache con un adeguato potere riflettente, mentre sono da evitare materiali trasparenti o traslucidi. L'acquisizione mediante tale metodo risulta adatta soprattutto per applicazioni *indoor*, con *range* di misurazione che spaziano dai 10 cm ai 5 m; con bassi valori di range, l'accuratezza raggiungibile è dell'ordine dei μm .

Alcuni esempi di scanner 3D *high-end* a triangolazione laser sono illustrati nella Figura 1.2. Il Minolta Vivid 910 è adottato in una vasta gamma di

ambiti applicativi e può raggiungere un'accuratezza di $10 \mu\text{m}$ in un range di misurazione fino a 2 m, scansionando circa 300000 punti in 2.5 s. La telecamera Sick Ranger E50, per via dei suoi elevati fattore di forma (3.0) e velocità di scansione (fino a 35000 profili/s), risulta invece particolarmente indicata per applicazioni industriali di controllo della qualità.



Figura 1.2: (a) Lo scanner 3D a triangolazione laser Minolta Vivid 910. (b) La telecamera Sick Ranger E50.

1.1.2 Time-of-flight

I sensori che sfruttano la tecnologia *time-of-flight* (tempo di volo) sono in grado di effettuare misurazioni di distanze emettendo un impulso di luce laser (anche infrarossa) da un illuminatore collocato nella stessa posizione del sensore, e calcolando il tempo trascorso dall'invio alla ricezione a seguito della riflessione del laser sulla superficie dell'oggetto [7]. Essendo nota la velocità della luce c ed il tempo t , è possibile ricavare la misura di distanza dal punto di riflessione come $\frac{c \cdot t}{2}$. L'accuratezza della misurazione effettuata è in questo caso limitata dalla risoluzione con cui il valore di t viene misurato, ed è tipicamente dell'ordine dei mm. Il *range* operativo di questi sensori è comunque elevato (dell'ordine di decine di metri), caratteristica che li rende adatti per acquisizioni *outdoor*, in particolare per scansioni di edifici ed

applicazioni di telerilevamento geografico. Il range massimo di misurazione è tuttavia limitato dalla durata della finestra temporale Δt che intercorre fra l'invio di impulsi successivi, ed è pari a $\frac{c \cdot \Delta t}{2}$; riflessioni da punti più distanti di questa quota portano quindi a misurazioni inconsistenti.

Una variante della tecnologia time-of-flight utilizza un segnale laser CW modulato in frequenza. In questo caso la distanza viene ricavata misurando la differenza di fase tra il segnale emesso e quello ricevuto a seguito della riflessione [7]. La massima differenza di fase misurabile in modo non ambiguo corrisponde al periodo completo, pertanto il range massimo è limitato dalla frequenza di modulazione adottata. Un esempio di sensore basato su questa tecnologia è il Sick LMS100 (Figura 1.3), pensato soprattutto per applicazioni *indoor* (range operativo: 0.5-20 m).



Figura 1.3: Il sensore a variazione di di fase Sick LMS100.

La maggior parte dei sensori time-of-flight e a variazione di fase di uso industriale operano in un range compreso tra i 10 e i 100 m, con valori di accuratezza che variano dai 5 ai 10 mm, e acquisiscono un solo punto alla volta, quindi sono spesso dotati di sistemi di attuazione quali specchi rotanti per variare l'orientazione del laser in modo da scansionare l'intera superficie di interesse. Analogamente alla triangolazione laser, la qualità della misurazione è influenzata in buona parte anche dalle caratteristiche di riflettività del materiale da scansionare.

Recentemente sono state introdotte nel mercato anche telecamere basate su queste tecnologie, le quali consentono di acquisire simultaneamente valori di profondità per ciascun pixel (*depth map*) [7][5]. In tal caso si usa un *array* di diodi laser o LED (tipicamente nella banda dell'infrarosso) per illuminare l'intera porzione di interesse della scena osservata. Seppur siano dotate

di sensori di dimensioni piuttosto ridotte, sono in grado di raggiungere elevate frequenze di aggiornamento (30-60 fps), che le rendono particolarmente indicate per applicazioni *real-time* nell'ambito della robotica, logistica, video-sorveglianza e intrattenimento. Due esempi sono la telecamera SwissRanger 4000 (risoluzione sensore: 176 x 144 pixel, frequenza di refresh: 50 fps, range operativo: 0.8-8.0 m) e il sensore Kinect per Xbox One (risoluzione sensore: 512 x 424 pixel, frequenza di refresh: 30 fps, range operativo: 0.5-4.5 m), presentati nella Figura 1.4. Il primo è un prodotto di fascia alta, mentre il secondo è pensato per un'utenza generica.



Figura 1.4: (a) La telecamera SwissRanger 4000. (b) Il sensore Kinect per Xbox One.

1.1.3 Triangolazione a luce strutturata

I sistemi di acquisizione 3D basati su triangolazione a luce strutturata fanno uso di un illuminatore (solitamente un proiettore LCD) che proietta un particolare *pattern* di luce sulla superficie dell'oggetto, e di una telecamera che acquisisce l'immagine della scena. Il principio di funzionamento consiste nel riconoscimento del pattern all'interno della vista, da cui è possibile, sulla base della deformazione indotta dalla superficie dell'oggetto, calcolare le coordinate tridimensionali dei punti (Figura 1.5) [8].

Uno dei pattern di luce strutturata maggiormente utilizzati è quello costituito da strisce parallele, le quali hanno il vantaggio di poter essere seg-

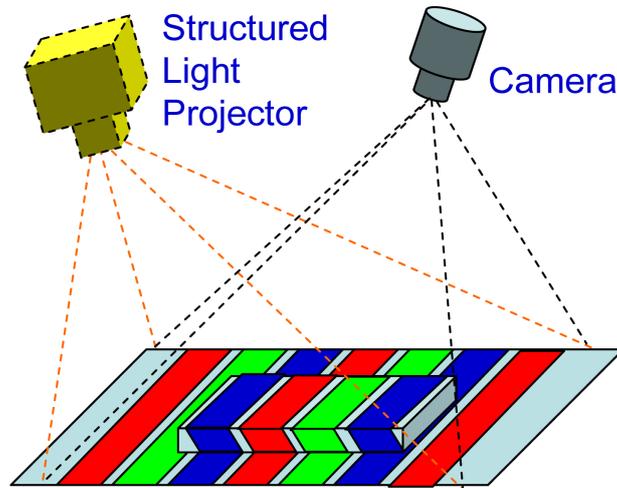


Figura 1.5: Illustrazione schematica di un sistema di acquisizione 3D basato su triangolazione a luce strutturata. Figura tratta da [8].

mentate in maniera semplice e accurata. Esse vengono codificate al fine di poter essere identificate univocamente nell'immagine. I metodi di codifica di questo determinato pattern sono molteplici, ed uno fra i più adottati è basato sulla codifica binaria di Gray, che consiste nella proiezione di una sequenza di M pattern caratterizzati da strisce a due livelli di intensità (bianco e nero) [8]. Mediante questa tecnica si possono codificare fino a 2^M strisce, in modo che i pixel appartenenti ad una determinata striscia possano essere distinti da tutti gli altri (Figura 1.6). L'uso della codifica di Gray garantisce una maggiore robustezza nel rilevamento, in quanto deve essere soddisfatta la condizione che la codifica di strisce adiacenti vari di un solo bit. Questa tecnica è di tipo *multi-shot* poiché fa uso di più acquisizioni successive, richiedendo dunque che l'oggetto rimanga completamente fermo durante l'intera durata della scansione. La risoluzione spaziale delle strisce cresce all'aumentare della dimensione della sequenza dei pattern proiettati, e questo può far sì che la durata dell'acquisizione possa essere troppo lunga per risultare utile nel contesto di applicazioni *real-time*. Per garantire acquisizioni a risoluzioni elevate riducendo al tempo stesso il numero di pattern, è possibile incrementare il numero di livelli di intensità utilizzati per la codifica (pattern a livelli di grigio).

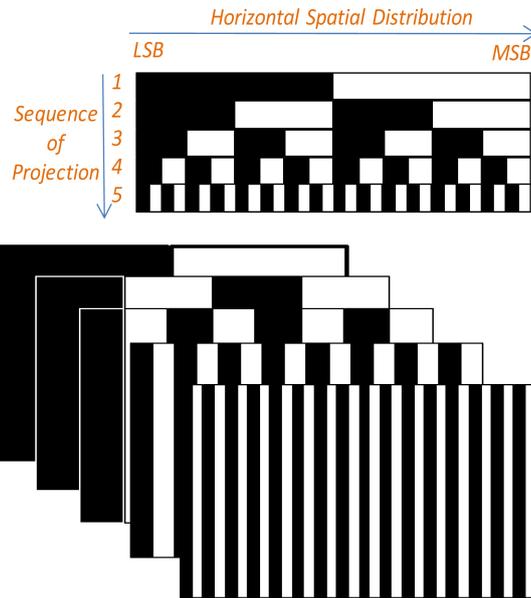


Figura 1.6: Codifica binaria di Gray delle strisce mediante una sequenza di 5 pattern. In questo caso è possibile codificare $2^5 = 32$ strisce. Figura tratta da [8].

Ulteriori metodi *single-shot* fanno invece uso di pattern a colori variabili per la codifica dei pixel, come il *rainbow pattern*, che raffigura tutti i colori dello spettro della luce visibile [8]. In tal caso, ai fini di un corretto rilevamento dei punti, è importante che le caratteristiche di riflettività del materiale dell'oggetto non alterino i colori del pattern luminoso.

Uno fra i principali vantaggi apportati da questa tecnica è l'elevata velocità di scansione. Infatti, invece di acquisire un punto o una linea alla volta (come avviene nel caso della triangolazione laser), gli scanner a luce strutturata scansionano l'intero campo visivo, triangolando, mediante l'utilizzo di una sola vista, tutti i punti della superficie visibile. Di conseguenza, non è necessario utilizzare sistemi di trascinamento per muovere l'oggetto.

Come la triangolazione laser, anch'esso è un metodo pensato per applicazioni *indoor*; a bassi *range* di misurazione può raggiungere un grado di accuratezza dell'ordine dei μm , prestandosi dunque come una valida alternativa all'uso di quest'ultima. Nel caso di utilizzo di illuminazione strutturata, la bontà dell'acquisizione può essere compromessa sia dal tipo di materiale

dell'oggetto da scansionare, il quale deve possedere le giuste caratteristiche di riflettività tali da consentire un rilevamento accurato del pattern, sia dall'interferenza della luce ambientale.

Un esempio di scanner 3D a luce strutturata è il Breuckmann triTOS-HE (Figura 1.7), una soluzione *high-end* progettata specificatamente per rilevamenti ad alta risoluzione nell'ambito della conservazione di beni culturali [1].



Figura 1.7: Lo scanner 3D Breuckmann triTOS-HE.

1.2 Metodi ottici passivi

I sistemi di acquisizione 3D basati su tecnologie a illuminazione passiva operano senza interferire con la scena, ossia si limitano a ricevere la luce proveniente dall'ambiente senza emetterne di propria [6]. Generalmente, si ottengono acquisizioni 3D meno robuste e dense rispetto a quelle prodotte da sistemi a illuminazione attiva. Uno fra i metodi appartenenti a questa categoria è la triangolazione mediante coppia di telecamere stereo.

1.2.1 Triangolazione mediante coppia di telecamere stereo

Questa tecnica di acquisizione 3D fa uso di due telecamere che inquadrano la medesima scena di interesse da due differenti punti di vista. In quanto passiva, sfrutta unicamente la luce ambientale, senza appoggiarsi a dispositivi di illuminazione aggiuntivi. Il principio che sta alla base delle applicazioni di visione stereo è sempre quello della triangolazione: avendo nota la geometria delle due telecamere, è possibile triangolare le coordinate tridimensionali di un punto della scena a partire dalla proiezione di tale punto sui due piani immagine [7]. Risulta chiaro, dunque, che il problema principale che questa tecnica presenta è quello di riuscire a trovare corrispondenze fra le due immagini che siano sufficientemente robuste ed affidabili.

La geometria generica di una coppia di telecamere stereo (qui semplificate secondo il modello *pinhole*) è presentata nella Figura 1.8 e prende il nome di geometria epipolare [11]. I due raggi proiettivi che congiungono P con i due centri di proiezione O_l e O_r intersecano i due piani immagine rispettivamente nei punti p_l e p_r . I due punti e_l e e_r sono detti epipoli, e sono rispettivamente la proiezione di O_r sul piano immagine della telecamera di sinistra e la proiezione di O_l sul piano immagine della telecamera di destra. Le rette che congiungono p_l a e_l e p_r a e_r sono dette rette epipolari. La retta epipolare della telecamera di sinistra è l'immagine del raggio proiettivo che collega P a O_r e viceversa; di conseguenza, data la proiezione di un punto su una delle due immagini, la corrispondenza sull'altra immagine è situata lungo la relativa retta epipolare. Questa proprietà consente di semplificare notevolmente il calcolo delle corrispondenze.

Prima di triangolare i punti è necessario ricondursi alla geometria ideale presentata nella Figura 1.9, che ha la particolarità di avere entrambi i piani immagine coplanari e con le righe allineate. Il sistema di riferimento della scena è centrato sul centro di proiezione della camera di sinistra, e i due assi principali sono paralleli all'asse Z , intersecando i piani immagine nei rispettivi punti principali. Si assume che la distanza f del piano immagine dal centro di proiezione sia la medesima per entrambe le telecamere. La differenza tra

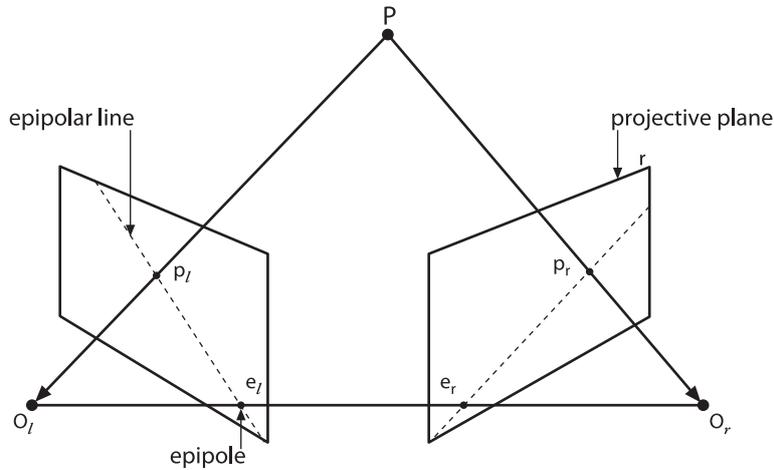


Figura 1.8: Geometria epipolare di una coppia di telecamere stereo (semplificazione con modello pinhole). Figura tratta da [4].

le due coordinate x dei punti p_l e p_r è definita disparità. Si noti che in questo caso le rette epipolari corrispondono alle righe dell'immagine, pertanto punti coniugati possiedono il medesimo indice di riga. Questa geometria ideale può essere ottenuta rettificando le due immagini a seguito della calibrazione del sistema stereo, tramite la quale vengono calcolati i parametri interni (che comprendono anche le distorsioni ottiche) delle due telecamere e le trasformazioni di rotazione e traslazione della prima rispetto alla seconda. Dopo aver individuato le corrispondenze nelle due immagini rettificate, è sufficiente riproiettare le coordinate dal piano immagine al sistema di riferimento della scena. In particolare, la quota di P lungo la direzione Z è inversamente proporzionale al valore di disparità e vale $Z = \frac{f \cdot T}{d}$, dove $d = x_l - x_r$, oppure $Z = \frac{f \cdot T}{(d - (c_x^l - c_x^r))}$, nel caso in cui le coordinate dei punti principali delle due immagini differiscano.

L'accuratezza dell'acquisizione mediante questa tecnica risulta limitata se messa a confronto con altri metodi attivi quali la triangolazione laser o con luce strutturata, e necessita di un'adeguata illuminazione della scena. Per il calcolo delle corrispondenze generalmente si adottano due metodi di tipo locale [7]: il primo di questi è basato sulla ricerca di *features* robuste all'interno delle immagini (per esempio i *corner*, i quali hanno la proprietà di essere indipendenti dalle variazioni di illuminazione e del punto di vista). Una volta

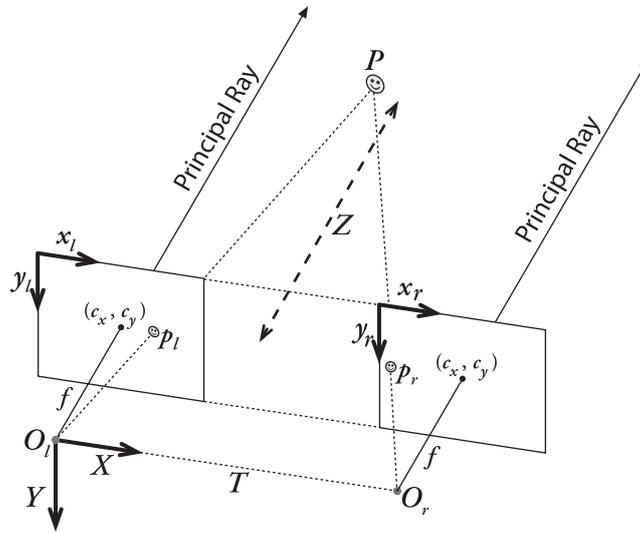


Figura 1.9: Geometria ideale della coppia di telecamere stereo ottenuta a seguito della rettificazione. Figura tratta da [4].

individuate, viene eseguito un *matching* basato sul confronto dei loro descrittori. Si noti che questo metodo produce un'acquisizione sparsa della scena e può risultare in un carico computazionale piuttosto elevato a causa dell'onerosità degli algoritmi di calcolo delle features e matching, fattore che non lo rende adatto per applicazioni *real-time*. Un secondo metodo consiste nel matching di piccole *patch* fra le due immagini secondo determinate misure di correlazione. Tale metodo è maggiormente soggetto ad errori rispetto al primo e necessita di una fase di *post-filtering* per eliminare i match spuri. A differenza del primo metodo si riesce ad ottenere un'acquisizione densa, e l'efficienza delle implementazioni attualmente disponibili ne consente l'utilizzo in contesti pratici.

Capitolo 2

Analisi del problema

In questo capitolo viene analizzato il problema oggetto della tesi, ovvero la ricostruzione 3D di oggetti ad alto potere diffusivo, introducendo innanzitutto il contesto e le necessità da cui esso scaturisce, motivando la scelta dell'utilizzo della triangolazione laser come metodo di acquisizione e mostrando l'impatto che tale tipologia di oggetti ha sul problema. Successivamente si procede ad elencare i vari problemi, da affrontare poi in fase di progetto, che la triangolazione laser presenta nell'ambito di tale caso di studio.

2.1 Contesto del problema

Il presente lavoro si inserisce nel più ampio contesto di un progetto di ricerca finanziato dalla regione Emilia-Romagna, e che coinvolge numerosi enti, fra cui l'*Istituto Ortopedico Rizzoli* di Bologna e il *Computer Vision Group* (CVG) dell'Università di Bologna (presso il quale è stata svolta la tesi). Tale progetto riguarda lo sviluppo di un sistema per la ricostruzione in camera bianca di protesi d'osso su misura. Lo studio svolto in questa tesi si focalizza sul sottosistema adibito all'acquisizione della superficie dell'osso e la successiva ricostruzione del modello 3D (CAD).

Il progetto è attualmente in fase di avvio, ed uno dei primi problemi esaminati è stato la scelta della tecnologia di acquisizione per rilevare la superficie dell'osso. I metodi appartenenti alla categoria delle tecniche ottiche, la cui

diffusione è in continua crescita, soprattutto grazie alla loro versatilità nel riuscire ad offrire soluzioni performanti anche a basso costo, costituiscono un'opzione appetibile e in linea con quanto richiesto dal problema; in particolare quelle basate su illuminazione attiva, le quali consentono di ottenere acquisizioni dense. Fra queste, le tecnologie di scansione di tipo *time-of-flight* o basate su misura della variazione di fase risultano poco adatte allo scopo, a causa della loro risoluzione limitata (come discusso nel Paragrafo 1.1.2). Un'ottima scelta può invece essere l'utilizzo di tecniche di triangolazione dei punti dell'oggetto mediante illuminazione controllata, come quella prodotta da un laser, atta ad identificare punti o sezioni della superficie. Essa si rivela particolarmente indicata per misurazioni limitate a *range* di pochi centimetri ed in grado di garantire un'elevata accuratezza. Nel caso di questo progetto si richiede che l'errore della ricostruzione sia compreso tra i 10 e i 50 μm , operando con pezzi delle dimensioni di pochi centimetri. Queste motivazioni hanno condotto quindi alla scelta della triangolazione laser come tecnica da adottare per l'acquisizione dell'osso; in particolare quella che fa uso di una linea proiettata, piuttosto che un singolo punto, in modo da velocizzare il processo di scansione.

Un problema di notevole rilevanza che emerge dall'uso della triangolazione laser, ma caratteristico di altre tecniche ad illuminazione attiva, è il fenomeno del *subsurface scattering*. Il tessuto osseo è infatti un materiale ad elevato potere diffusivo; ciò significa che gran parte della luce incidente penetra nell'oggetto, diffondendosi in maniera disordinata al di sotto della superficie, per poi fuoriuscire da zone differenti da quella di ingresso, producendo nell'immagine acquisita un alone luminoso di disturbo intorno al segnale utile (che è costituito dalla componente della luce riflessa in maniera diffusa). Nella Figura 2.1 vengono presentate due immagini che mostrano gli effetti di interazione della luce laser con un frammento di osso, sia nel caso di tessuto esterno compatto, che di tessuto interno dalla struttura trabecolare. Sempre nel contesto della triangolazione laser, in [9] è stato condotto uno studio dell'errore nel caso di scattering all'interno del marmo.

Gli effetti di scattering incidono significativamente sull'accuratezza e la precisione dell'acquisizione, pertanto uno dei temi trattati in questa tesi

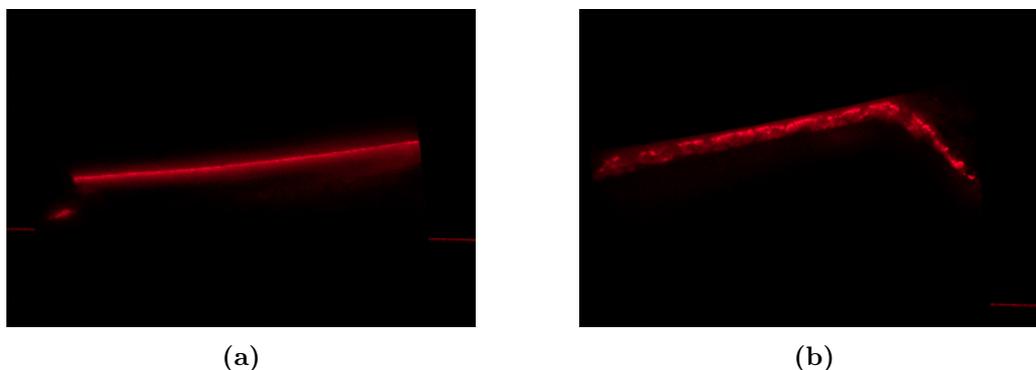


Figura 2.1: (a) Subsurface scattering del laser sul tessuto osseo compatto. (b) Subsurface scattering del laser sul tessuto osseo spongioso.

riguarda lo studio preliminare sulla triangolazione laser di superfici diffuse.

2.2 Il problema della triangolazione a lama laser

Dopo aver introdotto il problema di carattere generale e valutata la tecnologia di acquisizione più adatta allo scopo, in questo paragrafo viene presentata una panoramica delle principali problematiche legate all'utilizzo della triangolazione a lama laser, di cui è necessario tenere conto per la successiva progettazione della soluzione. La triangolazione a lama laser, come suggerisce il nome, lavora proiettando una linea che identifica un profilo su una sezione trasversale dell'oggetto. Per riuscire a scansionare una superficie tridimensionale si richiede quindi un sistema di attuazione, in modo da acquisire più profili al variare della sezione dell'oggetto colpito dal laser, per esempio ruotandolo o traslandolo attraverso la lama. Sarà importante in questo caso valutare l'adeguata risoluzione del passo di scansione, al fine di garantire l'accuratezza voluta. Per quanto concerne l'accuratezza della ricostruzione dei profili, vi sono diversi fattori che la influenzano, tra cui una corretta calibrazione del sistema, l'accuratezza nel rilevamento della linea laser all'interno dell'immagine e la risoluzione della misurazione.

2.2.1 Problematiche nel rilevamento del profilo laser

Rilevare in modo quanto più possibile accurato il profilo laser all'interno dell'immagine è uno dei problemi cruciali da risolvere. Oltre all'effetto di *scattering* menzionato in precedenza, il corretto rilevamento può essere complicato anche da altri fenomeni, quali rumore del sensore, riflessioni secondarie e occlusione del laser. L'occlusione, in particolare, si verifica quando parte del profilo non risulta visibile dal sensore, e si distingue in due tipi (Figura 2.2) [16]:

- Occlusione della telecamera. Una porzione della linea laser è nascosta, poiché occlusa dall'oggetto stesso. Gli effetti di questo tipo di occlusione variano a seconda della configurazione geometrica della coppia laser-telecamera.
- Occlusione del laser. Il laser non riesce fisicamente ad illuminare una parte dell'oggetto. In questo caso, se si vuole ottenere un'acquisizione completa della superficie del pezzo, è necessario eseguire più scansioni variandone l'orientazione.

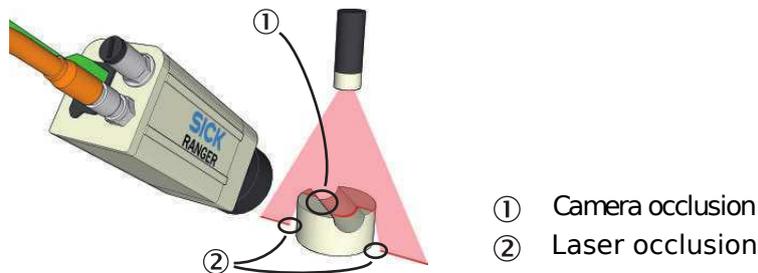


Figura 2.2: Un esempio di occlusione della telecamera e occlusione laser. Figura tratta da [16].

2.2.2 Geometria del sistema di triangolazione

Fissato un sistema di riferimento cartesiano tridimensionale, si possono individuare due angoli che risultano determinanti per la caratterizzazione della geometria della coppia laser-telecamera (Figura 2.3) [16]:

- L'angolo α , che definisce l'orientazione dell'asse ottico della telecamera rispetto alla normale della direzione Y .
- L'angolo β , che definisce di orientazione del fascio laser rispetto alla normale della direzione Y .

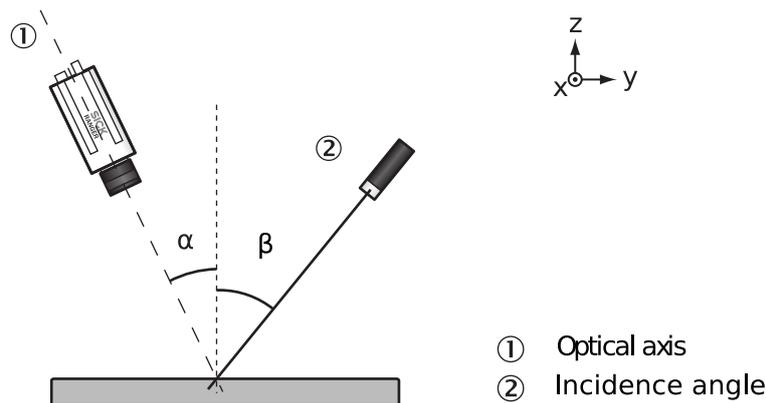


Figura 2.3: Angoli formati dall'asse ottico della telecamera e dal fascio laser con la normale alla direzione Y . Figura tratta da [16].

La risoluzione lungo gli assi X e Z è correlata al *range* di copertura della misurazione, determinato a sua volta dall'estensione del *Field Of View* (FOV) adottato per la telecamera. Inoltre, come si evince dalla Figura 2.4, il range nella direzione Z dipende anche dall'angolo tra l'asse ottico della telecamera e il fascio laser. Infatti, la diminuzione di tale angolo comporta un incremento del range, ma allo stesso tempo anche un peggioramento della risoluzione. Viceversa, una maggiore apertura consente di ottenere risoluzioni più elevate, limitando però il range e aumentando la possibilità di effetti di occlusione della telecamera [16].

Sulla base di quanto appena detto, si possono individuare quattro principali geometrie di configurazione della coppia laser-telecamera, ottenibili variando gli angoli α e β (Figura 2.5) [16] [14]:

- *Ordinary*. Questo tipo di *setup* ha la particolarità di avere la telecamera posizionata ortogonalmente rispetto alla direzione Y e consente di

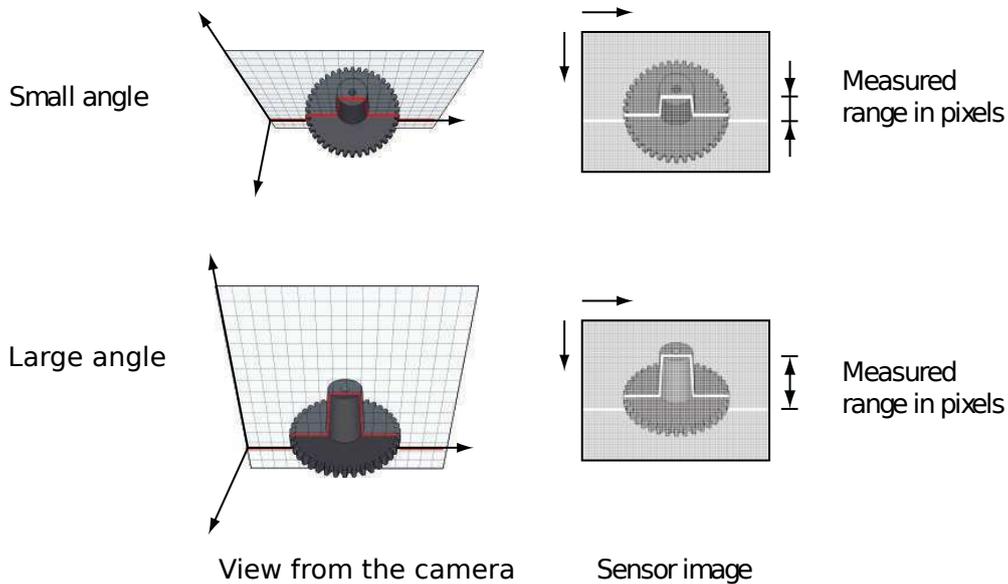


Figura 2.4: Illustrazione schematica della variazione della risoluzione e range lungo la direzione Z in funzione dell'angolo tra l'asse ottico della telecamera e il fascio laser. Figura tratta da [16].

ottenere la massima risoluzione nella direzione Z . Per via dell'orientazione del laser, il sistema dipende unicamente dalla componente di luce dovuta alla riflessione diffusa.

- *Reversed Ordinary.* Come suggerisce il nome, questa geometria si ottiene invertendo le posizioni di telecamera e laser nella configurazione Ordinary. Anch'essa è pensata per superfici in grado di riflettere la luce in maniera diffusa e garantisce una buona risoluzione nella direzione Z .
- *Specular.* Telecamera e laser sono montati alla stessa angolazione, ma lungo direzioni opposte. Questa geometria richiede una minore potenza del laser e risulta di utilità sia nel caso di oggetti scuri che lucidi, in quanto la telecamera in questo caso è in grado di rilevare anche le riflessioni speculari. Tuttavia, un grosso svantaggio che rende questo setup poco pratico per l'acquisizione 3D è la forte presenza di occlusione della parte frontale e posteriore dell'oggetto.

- *Look-Away*. Telecamera e laser sono montati sullo stesso lato rispetto alla normale. Per via di tale configurazione, questa è la geometria che richiede maggiore potenza del laser e con cui si ottiene la risoluzione peggiore. Un ulteriore svantaggio è dato dall'occlusione della parte posteriore dell'oggetto.

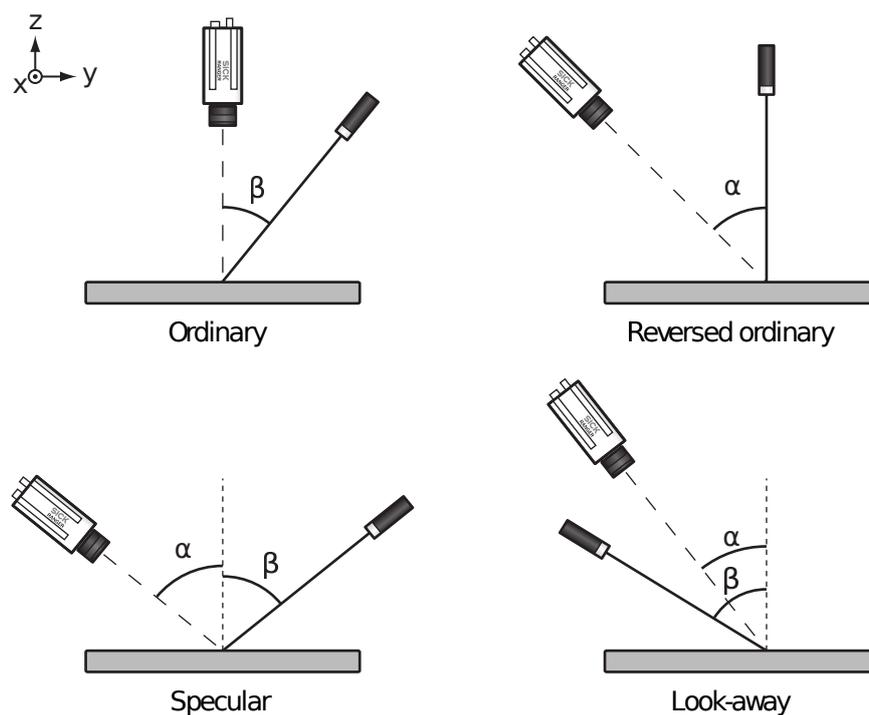


Figura 2.5: Principali configurazioni geometriche di un sistema di triangolazione a lama laser. Figura tratta da [16].

Capitolo 3

Progettazione

3.1 Richiami teorici

In questo paragrafo si richiamano alcuni concetti di teoria riguardanti il modello di proiezione prospettica della telecamera e le distorsioni delle lenti, i quali saranno in seguito utilizzati per la risoluzione del problema della triangolazione.

3.1.1 Modello di proiezione prospettica della telecamera

Una telecamera definisce un *mapping*, nello specifico una proiezione prospettica, tra una scena tridimensionale e un'immagine 2D. La tipologia di proiezione prospettica di interesse per il caso applicativo trattato è la *finite projective camera*, di seguito descritta tenendo conto dell'approssimazione della telecamera secondo il modello *pinhole* [11]. Tale mapping è esprimibile mediante una matrice \mathbf{P} di dimensioni 3×4

$$\mathbf{P} = \mathbf{K}[\mathbf{R} | -\mathbf{R}\tilde{\mathbf{C}}]$$

che trasforma un punto \mathbf{X} del sistema di riferimento della scena in un punto \mathbf{x} dell'immagine. Questa e le successive trasformazioni presentate mappano punti espressi in coordinate omogenee, dunque l'uguaglianza $\mathbf{x} = \mathbf{P}\mathbf{X}$ è de-

finita a meno di un fattore di scala. \mathbf{K} è una matrice 3×3 contenente i parametri intrinseci della telecamera

$$\mathbf{K} = \begin{bmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

dove α_u e α_v rappresentano la distanza tra il centro di proiezione \mathbf{C} e il piano del sensore, espressa in pixel rispettivamente per le direzioni u e v del sensore; u_0 e v_0 sono le coordinate in pixel del punto principale; mentre s è il parametro di *skew*, che di norma vale 0 nella maggior parte delle telecamere. Noto \mathbf{K} , si può specificare la proiezione di un punto \mathbf{X}_{cam} del sistema di riferimento della telecamera (con origine in \mathbf{C}) in un punto \mathbf{x} dell'immagine

$$\mathbf{x} = \mathbf{K}[\mathbf{I}|\mathbf{0}]\mathbf{X}_{\text{cam}}$$

\mathbf{R} è una matrice 3×3 che definisce l'orientazione degli assi del sistema di riferimento della telecamera, mentre $\tilde{\mathbf{C}}$ è il vettore che rappresenta la posizione del centro di proiezione nel sistema di riferimento della scena. La matrice di rotazione \mathbf{R} ed il vettore di traslazione $\mathbf{t} = -\mathbf{R}\tilde{\mathbf{C}}$ sono i parametri estrinseci che consentono di descrivere la trasformazione euclidea di un punto \mathbf{X} del sistema di riferimento della scena in un punto \mathbf{X}_{cam} del sistema di riferimento della telecamera

$$\mathbf{X}_{\text{cam}} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ 0 & 1 \end{bmatrix} \mathbf{X}$$

Combinando le due trasformazioni viste sopra si ottiene così la proiezione $\mathbf{x} = \mathbf{P}\mathbf{X}$.

3.1.2 Distorsioni ottiche

Finora si è assunto che la telecamera fosse approssimata in accordo con il modello *pinhole*, rendendo così la proiezione prospettica da scena a immagine un *mapping* lineare. Nella pratica tale assunzione non risulta valida, in quanto

il gruppo lenti all'interno dell'obiettivo della telecamera introduce delle distorsioni non-lineari [11]. Dato un punto \mathbf{x} dell'immagine, la distorsione delle sue coordinate (u, v) nelle coordinate (u_d, v_d) può essere descritta mediante le seguenti relazioni [4]:

$$\begin{cases} u_d = u + u_0(k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 u_0 v_0 + p_2(r^2 + 2u_0^2) \\ v_d = v + v_0(k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_2 u_0 v_0 + p_1(r^2 + 2v_0^2) \end{cases}$$

in cui:

- k_1 , k_2 e k_3 sono i coefficienti di distorsione radiale, effetto causato dall'utilizzo di lenti sferiche.
- p_1 e p_2 sono i coefficienti di distorsione tangenziale, dovuta principalmente al disallineamento delle lenti nell'obiettivo.
- u_0 e v_0 valgono rispettivamente $(u - u_c)$ e $(v - v_c)$, dove (u_c, v_c) è il centro di distorsione (spesso approssimato al punto principale).
- $r = \sqrt{u_0^2 + v_0^2}$.

Noti i coefficienti k_1 , k_2 , k_3 , p_1 e p_2 del modello, è quindi possibile stimare le coordinate ideali dei punti dell'immagine ottenuti a seguito della trasformazione prospettica secondo il modello pinhole, a partire da quelle distorte realmente osservate nell'immagine.

3.2 Progettazione hardware

In questa prima fase di progettazione vengono trattati gli aspetti riguardanti la scelta dell'hardware adottato per la realizzazione del sistema di triangolazione a lama laser (telecamere, proiettore laser, sistema di attuazione) e la loro configurazione.

3.2.1 Telecamere utilizzate

L'apparato prevede l'uso di due telecamere. La prima di esse è una telecamera a basso costo e *general-purpose*, le cui immagini acquisite dovranno essere in seguito elaborate mediante i metodi di triangolazione sviluppati in questa tesi. La seconda, invece, è una *range camera* di fascia alta con algoritmi dedicati per la scansione 3D mediante triangolazione a lama laser. In questo modo sarà quindi possibile raffrontare le misurazioni ottenute nei due casi. La scelta per le due tipologie è ricaduta rispettivamente sulle telecamere mvBluefox-121C e Sick Ranger E50.

Telecamera mvBluefox-121C

La mvBluefox-121C (Figura 3.1), prodotta da Matrix Vision, è una telecamera industriale pensata per applicazioni nell'ambito della visione artificiale, e del costo di circa 400 €. Essa monta un sensore CCD con risoluzione di 1024×768 , con pixel di dimensioni $4.65 \times 4.65 \mu\text{m}$, ed è in grado di raggiungere un *frame rate* massimo di 39 fps. La telecamera si interfaccia al PC tramite porta USB 2.0 (da cui riceve anche l'alimentazione), ed è inoltre dotata di un connettore D-sub maschio a 9 pin, utile per pilotare l'acquisizione delle immagini tramite l'invio di impulsi di *trigger*, e di un FPGA per l'elaborazione *on-board*. Oltre al software di controllo, la casa costruttrice mette a disposizione un kit di sviluppo per l'interfacciamento del dispositivo in applicazioni *custom*.

Telecamera Sick Ranger E50

La Sick Ranger E50 (Figura 1.2b) è una *range camera* prodotta da Sick AG e progettata specificatamente per l'utilizzo in applicazioni di acquisizione 3D con triangolazione a lama laser. Si tratta di un prodotto *high-end* il cui prezzo si assesta sulla fascia 9000-10000 €. La telecamera monta un sensore CMOS con risoluzione di 1536×512 , con pixel di dimensioni $9.5 \times 9.5 \mu\text{m}$. Uno dei principali pregi di questo dispositivo è l'elevata velocità di scansione: esso è infatti in grado di processare fino a 35000 profili laser al secondo; per questo motivo risulta largamente impiegato in ambito industriale per l'ispezione di



Figura 3.1: La telecamera industriale mvBluefox-121C.

prodotti in catena di montaggio. La telecamera si interfaccia al PC mediante Ethernet, e con essa sono forniti il software adibito alla calibrazione (3D Coordinator) e quello per il controllo della telecamera (Ranger Studio), oltre al kit di sviluppo per il programmatore (iCon API).

A seguito della calibrazione, durante la fase di ricostruzione del profilo le coordinate dell'immagine in pixel vengono trasformate nelle coordinate del piano individuato dalla lama laser, le quali sono espresse in unità assolute come millimetri o pollici. Mediante l'accessorio Power-I/O terminal è possibile alimentare il dispositivo e collegare segnali esterni provenienti da un *encoder*, al fine di sincronizzare l'acquisizione alla ricezione di impulsi di *trigger*. Se accoppiata con un meccanismo di attuazione, la telecamera è così in grado di effettuare una scansione del pezzo in movimento scandita dagli impulsi di trigger inviati dall'encoder. Associando un indice di posizione a ciascun profilo laser, la corretta sequenza di questi ultimi può essere ricostruita.

Per rilevare in modo efficiente ed accurato la linea laser all'interno dell'immagine, la telecamera implementa diverse tecniche, tra cui una basata sul calcolo del *center-of-gravity* (Hi3D COG) [16]. Essa consiste in una fase di sogliatura globale (con valore di soglia impostato manualmente), al fine di segmentare la componente riflessa del laser dal *background*, per poi procedere alla stima, colonna per colonna, della posizione del picco del laser mediante il calcolo della media degli indici di riga pesata sui rispettivi valori di inten-

sità (Figura 3.2). Le coordinate dei punti laser rilevati nell'immagine sono determinate quindi da un indice intero di colonna e da un indice di riga a *sub-pixel* (nello specifico, una risoluzione dichiarata di 1/16 di pixel [16]).

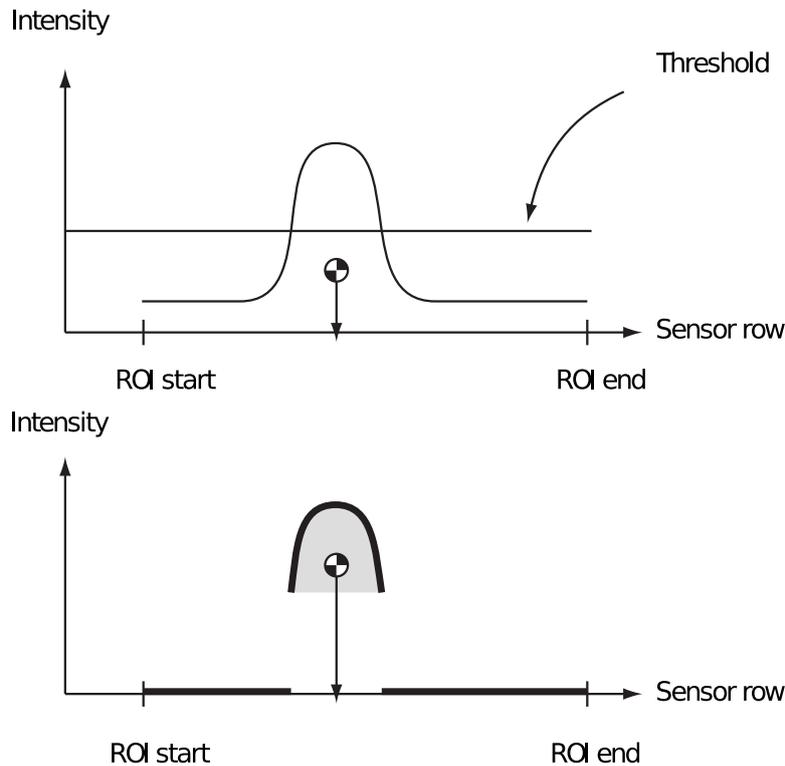


Figura 3.2: Illustrazione schematica del funzionamento dell'algoritmo Hi3D COG, implementato nella telecamera Sick Ranger E50. Figura tratta da [16].

3.2.2 Proiettore laser utilizzato

Si è deciso di utilizzare un proiettore laser a bassa potenza, soprattutto per minimizzare il rischio per l'operatore umano, per evitare di saturare il segnale e per limitare gli effetti di *subsurface scattering*, in quanto, all'aumentare della potenza, la luce laser penetra maggiormente al di sotto della superficie dell'oggetto illuminato. Il dispositivo adottato fa parte della serie ZM18S3 prodotta da Z-Laser (Figura 3.3). In particolare si tratta di un laser di classe 2M [2], con lunghezza d'onda pari a 640 nm; la potenza nominale è di 15 mW, ma può essere modulata applicando una tensione in corrente continua

che va da 0 V (<10% della potenza nominale) a 1 V (100% della potenza nominale). In questo caso la potenza è stata regolata a circa 1 mW.



Figura 3.3: Il proiettore laser ZM18S3.

3.2.3 Sistema di attuazione

Il sistema di triangolazione necessita di un apparato per l'automatizzazione della scansione, muovendo l'oggetto per variarne la sezione colpita dal laser e pilotando l'acquisizione sincronizzata delle due telecamere. Nell'ambito di questa tesi, si è scelto di realizzare un semplice prototipo di sistema di trascinamento lineare costituito da una struttura di sostegno in alluminio, con una base montata su un carrello con cremagliera che scorre lungo una guida. La rotazione dell'albero di un motore passo-passo sulla cremagliera consente lo spostamento del carrello.

L'utilizzo di un motore passo-passo costituisce la scelta più adatta per questo tipo di problema, nel quale si richiede uno spostamento tracciabile ed accurato. Pilotandolo ad un basso numero di giri, come nel nostro caso, è in grado di erogare una coppia elevata con un limitato consumo di corrente, ed è possibile simulare una velocità di rotazione costante temporizzando equamente l'invio di impulsi di step alle fasi del motore. Inoltre, a rotore fermo, garantisce una coppia statica che consente di mantenere il carico saldamente in posizione.

Il prototipo è stato realizzato utilizzando un motore passo-passo bipolare Sonceboz 6540 con le seguenti specifiche:

- impiega 200 step per compiere una rotazione completa dell'albero;
- l'angolo di rotazione di un singolo step è pari a 1.8° ;

- la massima corrente assorbita per fase è di 0.52 A.

Per pilotarlo ci si è serviti del *driver* Big Easy Driver¹ abbinato alla scheda con microcontrollore Arduino Uno², con l'ausilio della libreria AccelStepper³. La scelta di questo particolare driver deriva dalla necessità di dover pilotare il motore con la tecnica del *micro-stepping*, ovvero a frazioni di uno step completo. Infatti, si era inizialmente notato che la rotazione *full-step* dell'albero produceva notevoli vibrazioni che si propagavano sul carrello e sull'intera struttura di sostegno. La soluzione adottata è stata dunque quella di pilotarlo a frazioni di step, producendo così una rotazione più fluida e riducendo sensibilmente le vibrazioni. Inoltre, una maggiore risoluzione nella traslazione del piano incide positivamente sull'accuratezza della scansione. La scheda Big Easy Driver, economica e versatile, consente di raggiungere una risoluzione di micro-step pari ad 1/16 di uno step completo, che è quella adottata in questo caso. Il termine *step*, da qui in avanti, starà ad indicare 1/16 di full-step. È stato inoltre montato un pannello di MDF alla base della struttura di sostegno, al fine di smorzare ulteriormente le vibrazioni.

Non avendo a disposizione un *encoder* per scandire l'acquisizione delle due telecamere, si è deciso di emularlo all'interno della logica Arduino, facendo in modo che ogni numero prefissato di step percorsi dal motore passo-passo (passo di scansione) vengano inviati due impulsi digitali in quadratura A e B (come mostrato in Figura 3.4), con i rispettivi segnali negati /A e /B per controllare che non siano avvenuti errori durante la trasmissione. L'utilizzo di due segnali consente di codificare la rotazione del motore in senso orario e antiorario; infatti, il rilevamento del fronte di salita di B prima di quello di A corrisponde ad una rotazione in senso orario e viceversa. Questi quattro segnali in output dalla scheda Arduino sono inviati alla telecamera Sick Ranger E50, nella quale è possibile impostare il verso di scansione desiderato, mentre alla mvBluefox-121C è sufficiente inviare solo A o B, poiché in questa telecamera il pilotaggio mediante impulsi di *trigger* non è pensato per essere eseguito da un encoder.

¹rev. 1.6, <http://www.schmalzhaus.com/BigEasyDriver/>

²rev. 3, <https://www.arduino.cc/en/Main/arduinoBoardUno/>

³ver. 1.46, <http://www.airspayce.com/mikem/arduino/AccelStepper/>

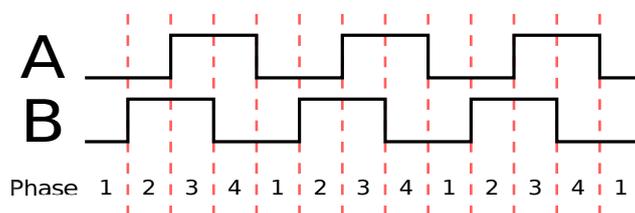


Figura 3.4: Illustrazione schematica dei due treni d'impulsi in quadratura A e B emessi dall'encoder (rotazione in senso orario).

Il prototipo di sistema di trascinamento, la cui implementazione verrà trattata approfonditamente nel Paragrafo 4.1, è mostrato nella Figura 3.5.

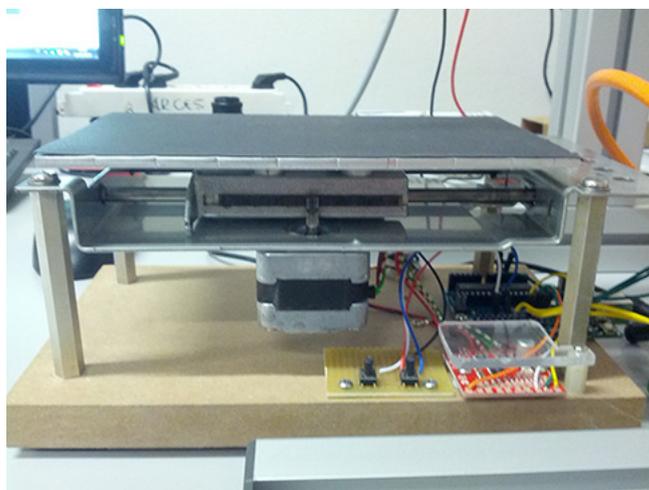


Figura 3.5: Il prototipo di sistema di trascinamento.

3.2.4 Configurazione geometrica del sistema di triangolazione

Si passerà ora alla descrizione della configurazione geometrica dell'apparato laser-telecamere. Data la particolare scelta di progettazione del sistema di attuazione, la geometria che si adatta di più in tale situazione è quella di tipo *Reversed Ordinary*. Infatti, posizionando la lama laser ortogonalmente al piano mobile, è possibile introdurre un sistema di riferimento cartesiano della scena in cui gli assi X e Z giacciono sul piano laser, mentre la direzione

Y corrisponde a quella di trascinamento attraverso la lama, come mostrato nella Figura 3.6.

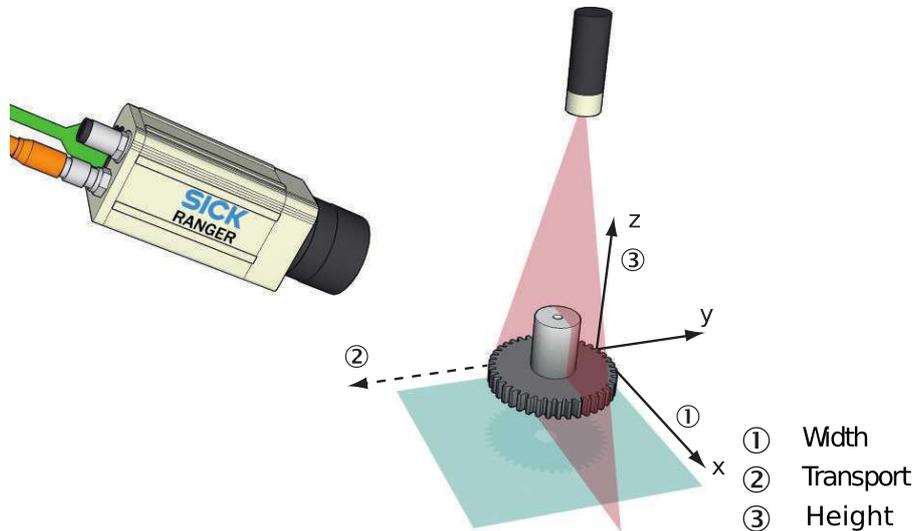


Figura 3.6: Sistema di riferimento adottato per la geometria Reversed Ordinary. Figura tratta da [16].

In questo modo, dunque, l’acquisizione consiste in una sequenza di profili nel piano X - Z , in cui tutti i punti di uno stesso profilo possiedono la medesima coordinata di Y , che è un multiplo del passo di scansione adottato. Questa scelta semplifica quindi il calcolo della triangolazione dei punti da immagine a sistema di riferimento della scena.

Un adeguato passo di scansione, dato dal numero di step percorsi tra due profili acquisiti, dovrà essere selezionato in maniera tale da garantire una risoluzione di spostamento lungo la direzione Y compresa tra i 10 e 50 μm . Ovviamente sarà prima necessario caratterizzare la velocità di trascinamento in modo assoluto (mm/s), a partire da quella espressa in step/s , per quantificare lo spostamento lineare prodotto da un singolo step e di conseguenza quello di un singolo passo di scansione, come verrà mostrato nel Paragrafo 5.2. La velocità di spostamento lineare del piano con cui si effettuerà la scansione dovrà essere di circa 0.5-1.5 mm/s , in modo da riuscire ad ottenere la risoluzione del passo di scansione desiderata senza che le telecamere “saltino” acquisizioni. Infatti, nel caso in cui il *cycle time* risultasse maggiore dell’in-

tervallo di tempo tra due segnali di *trigger*, alcuni di essi potrebbero essere ignorati. Questa valutazione è particolarmente importante per la telecamera mvBluefox-121C, il cui massimo *frame rate* raggiungibile è pari a 39 fps (che corrisponde ad un *cycle time* di 25.64 ms). Nel caso di questa telecamera è importante limitare il più possibile il tempo di esposizione, in modo da non abbassare ulteriormente il *frame rate*; per questo motivo si è scelto di lavorare con un tempo di esposizione di 2500 μs , con il quale si è osservato un *cycle time* pari a 27.58 ms. La telecamera Sick Ranger E50, adatta per operare ad elevate velocità di scansione, è invece in grado di raggiungere valori di *cycle time* di pochi microsecondi. Per quest'ultima è stato impostato un tempo di esposizione di 5000 μs , tale da consentire un adeguato livello di intensità del laser, mantenendo un basso *cycle time* (10 ms).

La *baseline* B , ovvero la distanza del centro di proiezione dal piano laser, l'angolo di orientazione θ dell'asse ottico rispetto alla *baseline*, ed il FOV sono stati regolati nelle due telecamere in modo da inquadrare la stessa porzione di scena e avere un adeguato *range* (mm) e risoluzione (mm/pixel) in X e Z , cercando di allineare al meglio le righe del sensore con la linea laser proiettata sul piano d'appoggio. Entrambe montano un obiettivo con focale da 25 mm con un anello distanziatore spesso 2.5 mm nel caso della Sick Ranger E50, e 1 mm nel caso della mvBluefox-121C, per allontanare ulteriormente il piano del sensore dal centro ottico dell'obiettivo e consentire una corretta messa a fuoco della lama laser. È essenziale che nell'immagine la linea laser risulti spessa almeno dai 3 ai 5 pixel, in modo tale che la posizione dei punti del profilo possa essere stimata con risoluzione di *sub-pixel*, come avviene per il metodo Hi3D COG e quelli che verranno successivamente illustrati. La Figura 3.7 mostra la struttura di supporto con il laser, le due telecamere e il piano mobile.

Partendo dalla seguente stima dei valori di B e θ per le due telecamere:

- mvBluefox-121C: $B = 330$ mm, $\theta = 24^\circ$;
- Sick Ranger E50: $B = 240$ mm, $\theta = 36^\circ$;

sono stati calcolati i valori di *range* e risoluzione nelle direzioni X e Z mediante relazioni trigonometriche. Si noti che tale valutazione preliminare risulta

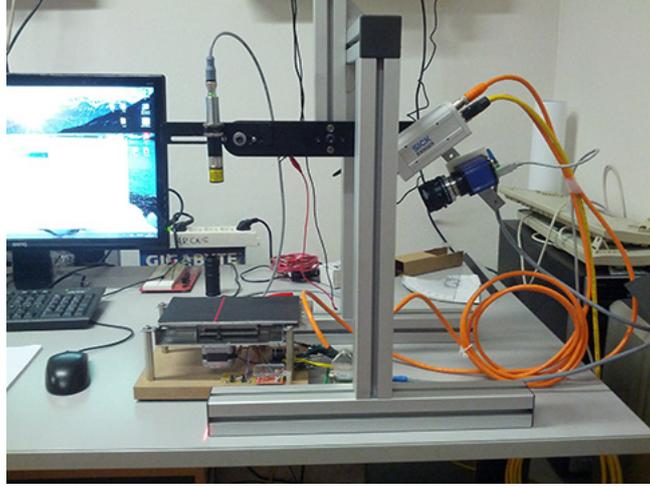


Figura 3.7: La struttura di supporto con il laser, le due telecamere e il piano mobile.

valida solo nell'ipotesi in cui la telecamera sia disposta di fronte alla lama laser e la direzione dell'asse X venga scelta in modo da risultare parallela alle righe del sensore.

Nota la distanza d del centro ottico dal sensore (~ 26 mm nella mvBluefox-121C e ~ 27.5 mm nella Sick Ranger E50) e le dimensioni di quest'ultimo (4.76×3.57 mm nella mvBluefox-121C e 14.59×4.86 mm nella Sick Ranger E50), è possibile calcolare gli angoli di campo, rispettivamente secondo la larghezza e altezza del sensore, FOV_w e FOV_h

$$FOV_w = 2 \arctan \frac{sensor_w}{2d} \quad \text{e} \quad FOV_h = 2 \arctan \frac{sensor_h}{2d}$$

Il valore di range lungo la direzione Z è ricavabile mediante la seguente relazione

$$Z_{range} = B \cdot \tan\left(\theta + \frac{FOV_h}{2}\right) - B \cdot \tan\left(\theta - \frac{FOV_h}{2}\right)$$

Il valore di range di X , invece, varia lungo Z a causa della distorsione indotta dalla proiezione prospettica. Si introducono due segmenti r_1 ed r_2 , coplanari con l'asse ottico, che collegano C con il piano laser rispettivamente all'estre-

mità superiore ed inferiore del range di Z , la cui lunghezza può essere così ricavata

$$r_1 = \frac{B}{\cos(\theta - \frac{FOV_h}{2})} \quad \text{e} \quad r_2 = \frac{B}{\cos(\theta + \frac{FOV_h}{2})}$$

Arrivati a questo punto, si può calcolare il range di X all'estremità superiore ed inferiore del range di Z come

$$X_{range-top} = 2 \cdot r_1 \cdot \tan(\frac{FOV_w}{2}) \quad \text{e} \quad X_{range-bot} = 2 \cdot r_2 \cdot \tan(\frac{FOV_w}{2})$$

Quindi, la risoluzione di X nei due casi si ottiene dividendo il range per la larghezza del sensore in pixel, mentre la risoluzione di Z può essere approssimata con la formula [16][14]

$$Z_{res} \approx \frac{X_{res}}{\sin(90 - \theta)}$$

I valori di FOV, range e risoluzione stimati per le due telecamere sono raccolti nella Tabella 3.1. Dato che i valori $X_{range-top}$ e $X_{range-bot}$ della telecamera Sick Ranger E50 risultano fin troppo elevati per i nostri scopi, si è deciso di lavorare con una ROI centrata sull'immagine di dimensioni 512×512 .

	mvBluefox-121C	Sick Ranger E50
FOV_w	10.88°	32.54°
FOV_h	8.17°	11.11°
$X_{range-top}$ (mm)	64.34	148.49
$X_{range-bot}$ (mm)	68.40	168.88
Z_{range} (mm)	54.36	65.13
$X_{res-top}$ (mm/pixel)	0.06	0.10
$X_{res-bot}$ (mm/pixel)	0.07	0.11
$Z_{res-top}$ (mm/pixel)	0.07	0.12
$Z_{res-bot}$ (mm/pixel)	0.07	0.14

Tabella 3.1: Stime di FOV orizzontale e verticale, e di range e risoluzione nelle direzioni X e Z per le due telecamere.

3.2.5 Architettura hardware del sistema di triangolazione

Nella Figura 3.8 è rappresentato lo schema a blocchi dell'architettura del sistema. Come si può osservare, il PC è collegato alla scheda Arduino mediante interfaccia USB, in modo da poter inviare stringhe di comando in comunicazione seriale per il pilotaggio del motore passo-passo. In uscita dalla scheda Arduino al Power-I/O terminal si hanno i quattro segnali A, B, /A e /B (corrispondenti alle due fasi dell'encoder ed i rispettivi negati), che vengono poi rediretti alla telecamera Sick Ranger E50 insieme all'alimentazione, mentre verso la telecamera mvBluefox-121C viene rediretto solamente il segnale B, da utilizzare come impulso di *trigger*.

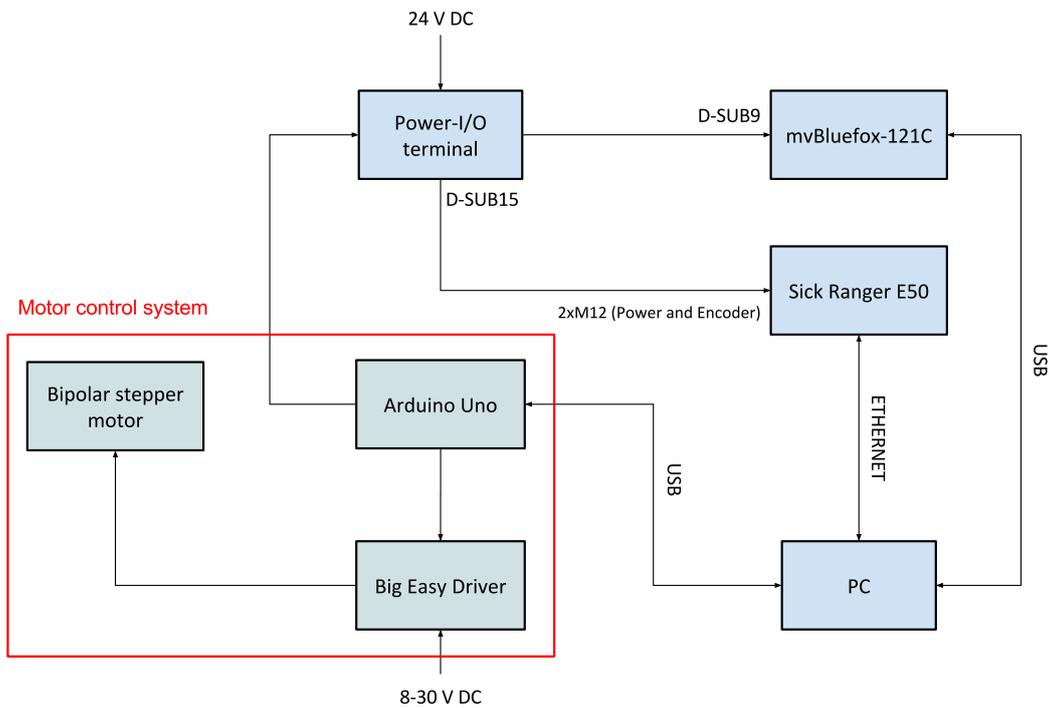


Figura 3.8: Schema a blocchi dell'architettura hardware del sistema di triangolazione.

3.3 Progettazione software

Dopo aver affrontato lo studio del *setup* di laser, telecamere e sistema di trascinamento, si passa alla trattazione dei passi procedurali che compongono la risoluzione al problema della triangolazione, i quali verranno poi concretizzati nello sviluppo di una libreria software per la triangolazione 3D dei punti a partire dalle immagini dei profili laser acquisite dalla telecamera mvBluefox-121C.

La fase preliminare consiste nella calibrazione della telecamera, in modo da stimare i suoi parametri interni, le distorsioni ottiche introdotte dall'obiettivo ed i parametri estrinseci che definiscono la posa della telecamera rispetto alla lama laser. A partire da tali parametri è possibile ricavare la trasformazione proiettiva (omografia) tra il piano del sensore (u, v) con coordinate in pixel e il piano laser $X-Z$ con coordinate in millimetri.

Una volta effettuata la calibrazione è possibile procedere con la triangolazione dei profili, che consta delle seguenti fasi:

1. rilevamento, all'interno dell'immagine, delle coordinate distorte dei punti del profilo laser riflesso sul sensore;
2. stima delle coordinate non distorte dei punti, a partire da quelle distorte;
3. riproiezione dei punti dell'immagine non distorti sul piano laser $X-Z$.

Noto il passo di scansione, che definisce il valore della coordinata Y , la posizione di ciascun punto nella scena può così essere determinata. I punti acquisiti vengono immagazzinati in una *point cloud*, che può essere in seguito ulteriormente elaborata (pulizia, *subsampling*, calcolo delle normali) per ricostruire il modello 3D della superficie scansionata (CAD o *mesh* di triangoli). Le fasi di elaborazione descritte sono schematizzate nella Figura 3.9.

Durante l'acquisizione, l'interfacciamento con le due telecamere avviene utilizzando i relativi software di controllo forniti dai produttori. Nel caso della telecamera mvBluefox-121C, per una maggiore versatilità del collaudo degli algoritmi, viene salvata la sequenza di immagini che compone la scansione,

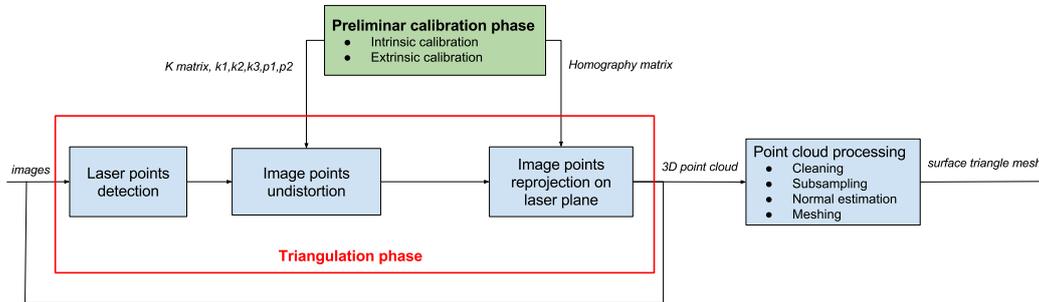


Figura 3.9: Pipeline di elaborazione software.

così da poter essere elaborata in un secondo momento utilizzando le tecniche implementate; nel caso della telecamera Sick Ranger E50 vengono invece salvati i profili ricostruiti in tempo reale mediante elaborazione *on-board*, i cui punti potranno essere in seguito estratti e rielaborati (tramite le API della libreria iCon) per ottenere una point cloud.

3.3.1 Calibrazione della telecamera

Lo scopo della calibrazione della telecamera è quello di stimare la matrice dei parametri intrinseci \mathbf{K} , i coefficienti del modello di distorsione ed i parametri estrinseci (matrice \mathbf{R} e vettore \mathbf{t}), per riuscire a caratterizzare la proiezione prospettica \mathbf{P} dei punti della lama laser sull'immagine. I parametri estrinseci sono di interesse anche per stimare più accuratamente la *baseline* e l'angolo θ di *setup*. Dato che per tutti i punti giacenti sul piano laser vale che $Y = 0$, \mathbf{P} diventa una matrice 3×3 non singolare

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix}$$

dove \mathbf{r}_1 ed \mathbf{r}_3 sono rispettivamente la prima e terza colonna di \mathbf{R} . La matrice così ottenuta descrive l'omografia che mappa i punti del piano laser sul piano del sensore; invertendola si ottiene invece il *mapping* che consente di riproiettare i punti dell'immagine sul piano laser. Alternativamente, se sono noti i coefficienti di distorsione, tale omografia può essere stimata direttamente a partire da un insieme di corrispondenze di punti dell'immagine non distorti e punti del piano laser [11].

Prima di valutare la metodologia che si vuole adottare per la stima dei parametri di calibrazione, si descrive brevemente la procedura di calibrazione della telecamera Sick Ranger E50 implementata nel software 3D Coordinator. Per ulteriori dettagli in merito si rimanda a [15]. La calibrazione di questa telecamera richiede un *pattern* a dente di sega di un materiale con un adeguato potere riflettente e di dimensioni tali da ricoprire l'intero FOV della telecamera con almeno quattro denti. Nel nostro caso è stato fatto realizzare un calibratore a dente di sega in alluminio (Figura 3.10) con le seguenti specifiche:

- spessore: 8 mm;
- lunghezza: 300 mm;
- altezza: 24.29 mm;
- angolo di picco: 90° ;
- raggiatura picco: 0.1 mm;
- raggiatura valle: 0.5 mm;
- numero di denti: 30;
- larghezza del singolo dente: 10 mm;
- altezza del singolo dente (tenendo conto della raggiatura): 4.75 mm;
- altezza del singolo dente (ignorando la raggiatura): 5.00 mm.

Dopo aver impostato la ROI di calibrazione ed i valori in millimetri di altezza del dente (ignorando la raggiatura) e larghezza del dente, si procede con una prima fase che consiste nel posizionare il lato piatto del calibratore sotto il raggio laser, muovendolo in modo da ricoprire ogni zona della ROI e variandone l'orientazione. Mediante questa operazione il software stima le distorsioni ottiche. La seconda fase è analoga, ma viene eseguita posizionando il calibratore nel verso del profilo a dente di sega cercando di mantenerlo il più possibile allineato con la lama laser, e serve a stimare l'omografia tra



Figura 3.10: Il calibratore a dente di sega.

piano del sensore e piano laser a partire dai picchi e le valli individuati nelle immagini acquisite. Al termine della calibrazione viene visualizzato l'errore di riproiezione nei vari settori della ROI e prodotta in output una *Look-Up Table* da utilizzare in fase di misurazione.

Si passa ora alla trattazione delle tecniche scelte per la calibrazione della telecamera, la cui implementazione verrà analizzata nel dettaglio nel Paragrafo 4.2.1.

Metodo per la stima dei parametri intrinseci e coefficienti di distorsione

Si è deciso di adottare una soluzione basata sul metodo descritto in [18], in quanto consente una procedura di calibrazione di semplice esecuzione. Inoltre, di tale metodo sono disponibili implementazioni consolidate in numerose librerie *open-source*, tra cui OpenCV [4]. Date n viste (almeno due) di un pattern di calibrazione planare disposto in diverse orientazioni, in ciascuna di esse vengono estratte m *features* messe in corrispondenza con i rispettivi punti del piano della scena. A partire da queste corrispondenze, il metodo stima così la matrice \mathbf{K} , i coefficienti di distorsione e le n coppie (\mathbf{R}, \mathbf{t}) che descrivono le pose della telecamera rispetto al pattern, attraverso la minimizzazione iterativa dell'errore geometrico

$$\sum_{i=1}^n \sum_{j=1}^m \|\mathbf{m}_{ij} - \hat{\mathbf{m}}_{ij}\|^2$$

dove \mathbf{m}_{ij} è il j -esimo punto osservato nell'immagine i , mentre $\hat{\mathbf{m}}_{ij}$ è il corrispondente punto della scena riproiettato usando i parametri stimati.

Nel nostro caso è stato utilizzato un pattern 3×9 di tipo *asymmetric circles* (Figura 3.11), le cui features sono i centri dei cerchi. Il pattern è stato stampato a risoluzione di 600 dpi e successivamente incollato su una lastra di vetro. Le dimensioni sono di circa 5×3 cm, in modo da risultare interamente coperto dal campo visivo della telecamera.

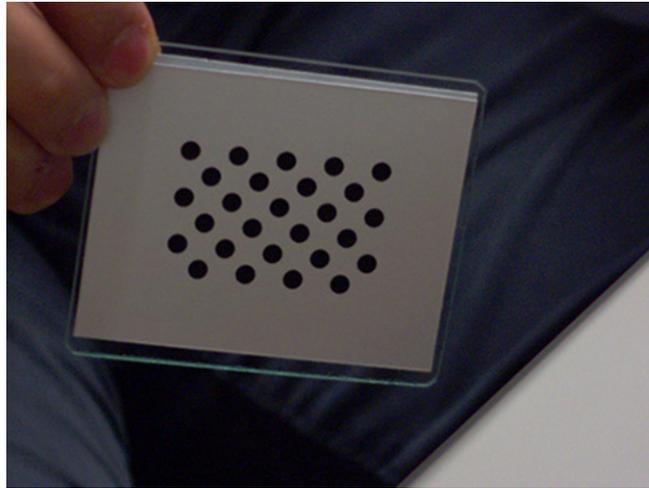


Figura 3.11: Il pattern planare di tipo asymmetric circles.

Metodo per la stima dei parametri estrinseci

Nota la matrice \mathbf{K} ed i coefficienti di distorsione, è possibile stimare \mathbf{R} e \mathbf{t} a partire da corrispondenze di punti di un'immagine e punti del piano laser mediante minimizzazione iterativa dell'errore di riproiezione precedentemente definito. In questo caso il *pattern* planare risulta di scarsa praticità, data la difficoltà nel riuscire ad allinearli accuratamente con il piano laser; pertanto si è scelto di utilizzare viste del calibratore a dente di sega colpito dalla lama laser, estraendo il profilo mediante le tecniche di rilevamento descritte nel paragrafo successivo e rilevando i picchi e le valli da adottare come *features* per

le corrispondenze, proprio come avviene nel software di calibrazione del sistema Sick. Le coordinate dei punti corrispondenti sul piano laser $X-Z$ possono essere assegnate avendo note la larghezza e l'altezza in millimetri del dente di sega. L'utilizzo di molteplici viste del calibratore, acquisite variandone le posizione, permette di ottenere una stima più robusta.

3.3.2 Triangolazione del profilo laser

Per quanto concerne la triangolazione, l'aspetto più importante da trattare è quello del rilevamento accurato dei punti della linea laser nell'immagine. Avendo la striscia proiettata uno spessore finito di circa 1 mm, è necessario stimarne il centro. La procedura consiste dunque nell'analizzare, colonna per colonna, il profilo di intensità dei pixel illuminati dal laser lungo le righe dell'immagine. Come si può notare dalla Figura 3.12, tale profilo di intensità presenta generalmente un andamento approssimabile ad una curva gaussiana, in cui il picco centrale di massima intensità rappresenta la posizione del punto di impatto del laser nella determinata colonna [12]. Rilevare l'indice di riga che corrisponde al picco stimato a risoluzione di pixel interi non sarebbe sufficiente per garantire misurazioni dell'accuratezza voluta lungo Z , per cui è necessario calcolarlo a livello di *sub-pixel*. Si noti, invece, che la risoluzione di misurazione nella direzione X è limitata dal fatto che viene rilevato un solo punto laser per ciascuna colonna.

Per la stima dell'indice di picco a sub-pixel si intende sviluppare tre differenti metodi:

- Calcolo del *center-of-gravity* del profilo d'intensità, che è la tecnica analoga a quella implementata nell'algoritmo Hi3D COG della telecamera Sick Ranger E50. La posizione del picco viene stimata come media degli indici di riga dei pixel che compongono il profilo pesata rispetto ai valori d'intensità.
- *Fitting* di una funzione gaussiana sui pixel del profilo di intensità e ricerca del picco.

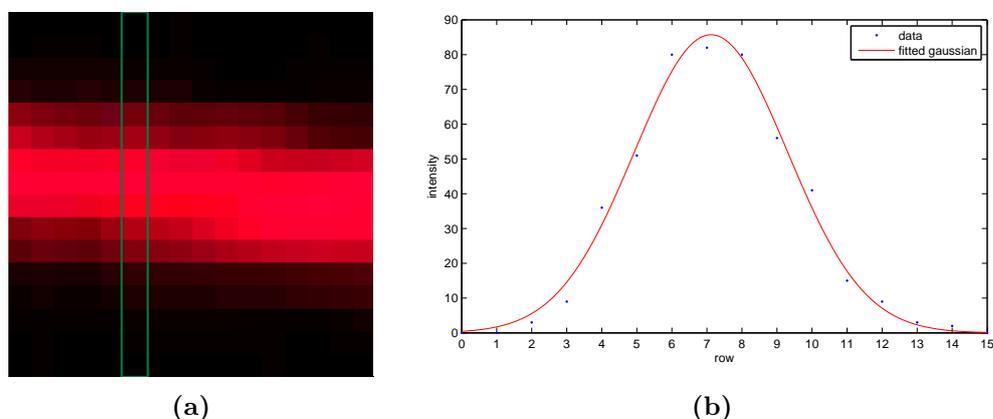


Figura 3.12: (a) Un dettaglio dell'immagine di un profilo laser acquisita con la telecamera mvBluefox-121C. (b) Profilo di intensità trasversale della colonna evidenziata.

- Interpolazione dei pixel del profilo di intensità mediante spline cubica naturale e ricerca del picco.

Nei metodi basati su center-of-gravity e fitting di una gaussiana, la presenza di illuminazione di disturbo intorno alla linea laser dovuta a *scattering*, riflessioni secondarie o luce ambientale (nel nostro caso si cercherà di minimizzare quest'ultima in fase di acquisizione), può comportare una sensibile deviazione della posizione del picco stimato da quello ideale. Per questo motivo è necessaria un'operazione di sogliatura preliminare (sperimentando diverse tecniche), allo scopo di eliminare tali componenti di disturbo senza peggiorare il segnale utile. L'uso di valori bassi di potenza del laser e tempo d'esposizione evita di saturare il segnale, mantenendo una significativa differenza d'intensità tra profilo laser e illuminazione di disturbo, in modo da rendere possibile un'efficace segmentazione. La sogliatura risulta utile anche per azzerare pixel di *background* nelle zone di occlusione del laser, evitando il rilevamento di falsi positivi.

La linea laser può inoltre presentare delle irregolarità (come visibile nella Figura 3.13), dovute in minima parte al rumore del sensore, ma soprattutto alla riflettività non uniforme della superficie scansionata. Ciò porta alla

formazione di picchi di massima intensità scostati anche di qualche pixel rispetto al centro ideale della striscia. Per tali ragioni si è scelto di applicare un filtro di *smoothing* a valle della fase di sogliaatura, in modo da rilevare un profilo meno affetto da rumore.

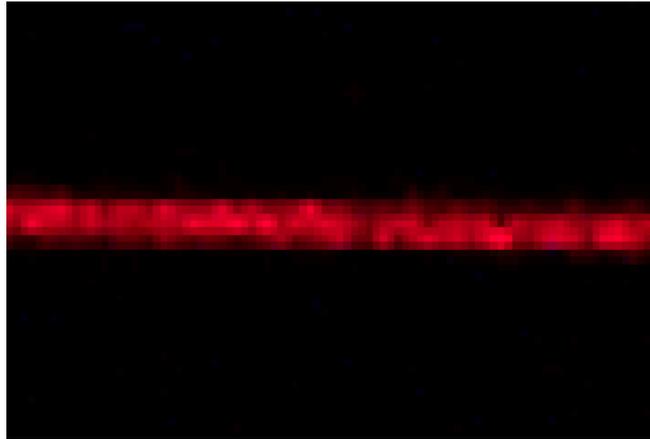


Figura 3.13: Un dettaglio dell'immagine di un profilo laser con rumore acquisita con la telecamera mvBluefox-121C.

Dopo aver corretto le distorsioni, i punti del profilo laser possono essere riproiettati sul piano laser mediante l'omografia stimata.

3.3.3 Software e librerie utilizzate

Il linguaggio di programmazione adottato è il C++. In particolare ci si appoggia alle librerie *open-source* OpenCV⁴, che fornisce strutture dati e metodi per l'elaborazione delle immagini e la calibrazione della telecamera, e PCL (Point Cloud Library)⁵, per la creazione ed esportazione di *point cloud*. Entrambe le librerie sono distribuite sotto licenza BSD. Per la visualizzazione e il *processing* di point cloud è stato invece utilizzato il software CloudCompare⁶.

⁴ver. 2.4.9, <http://www.opencv.org/>

⁵ver. 1.7.2, <http://www.pointclouds.org/>

⁶ver. 2.6.1, <http://www.cloudcompare.org/>

3.3.4 Architettura della libreria software

La Figura 3.14 mostra il diagramma UML di struttura della libreria per la calibrazione e triangolazione dei profili laser sviluppata in questa tesi (CvMeasLib), la quale è organizzata in tre moduli che si appoggiano a strutture dati e funzioni di OpenCV. All'interno del modulo `calib.h` vi è definita la classe `CameraCalibrator`, che mantiene nei suoi campi i vari dati di calibrazione. Il metodo `computeIntrinsicParams` effettua la stima dei parametri intrinseci e dei coefficienti di distorsione, ricevendo in input il vettore dei *paths* delle immagini del *pattern* planare, la struttura `PatternData` contenente le caratteristiche del pattern, e un *flag* per regolare quali coefficienti di distorsione stimare. Il metodo `computeExtrinsicParamsSawtooth` consente invece di stimare i parametri estrinseci a partire da un vettore di paths delle immagini del profilo laser del calibratore a dente di sega, i dati di altezza e larghezza del dente, la struttura contenente i parametri per il *pre-processing* sull'immagine, ed il tipo di metodo con cui si vuole rilevare la linea laser, in quanto la sua implementazione si appoggia alle funzioni definite nel modulo `laser.h`. La classe mette inoltre a disposizione due metodi per importare rispettivamente i soli parametri intrinseci (`importIntrinsicParams`) o i dati della calibrazione completa (`importCameraCalibrationParams`) da file XML. Ulteriori metodi di interesse sono `undistortImagePoints`, che effettua la correzione della distorsione su un vettore di punti immagine passato in input, e `undistortView`, che corregge la distorsione su un'immagine. Nella Figura 3.15 è riportato il diagramma UML di stato della classe.

All'interno del modulo `laser.h` sono definite le funzioni `preprocessView`, per eseguire il pre-processing sull'immagine passata in input, e `detectLaserStripe` per effettuare il rilevamento del profilo laser sull'immagine utilizzando uno dei tre metodi a disposizione; tale funzione restituisce in output un vettore contenente gli indici di riga a *sub-pixel* rilevati in ogni colonna dell'immagine e ordinati a partire dalla quella di sinistra.

Nel modulo `transform.h` è definita la funzione `transform2DPoints`, che proietta i punti contenuti nel vettore di input utilizzando la matrice di

omografia 3×3 data (trasformazione da piano immagine a piano laser).

Nella Figura 3.16 è mostrato il diagramma UML di interazione che schematizza la sequenza di chiamate alla libreria CvMeasLib per effettuare la calibrazione della telecamera e la triangolazione di un profilo laser.

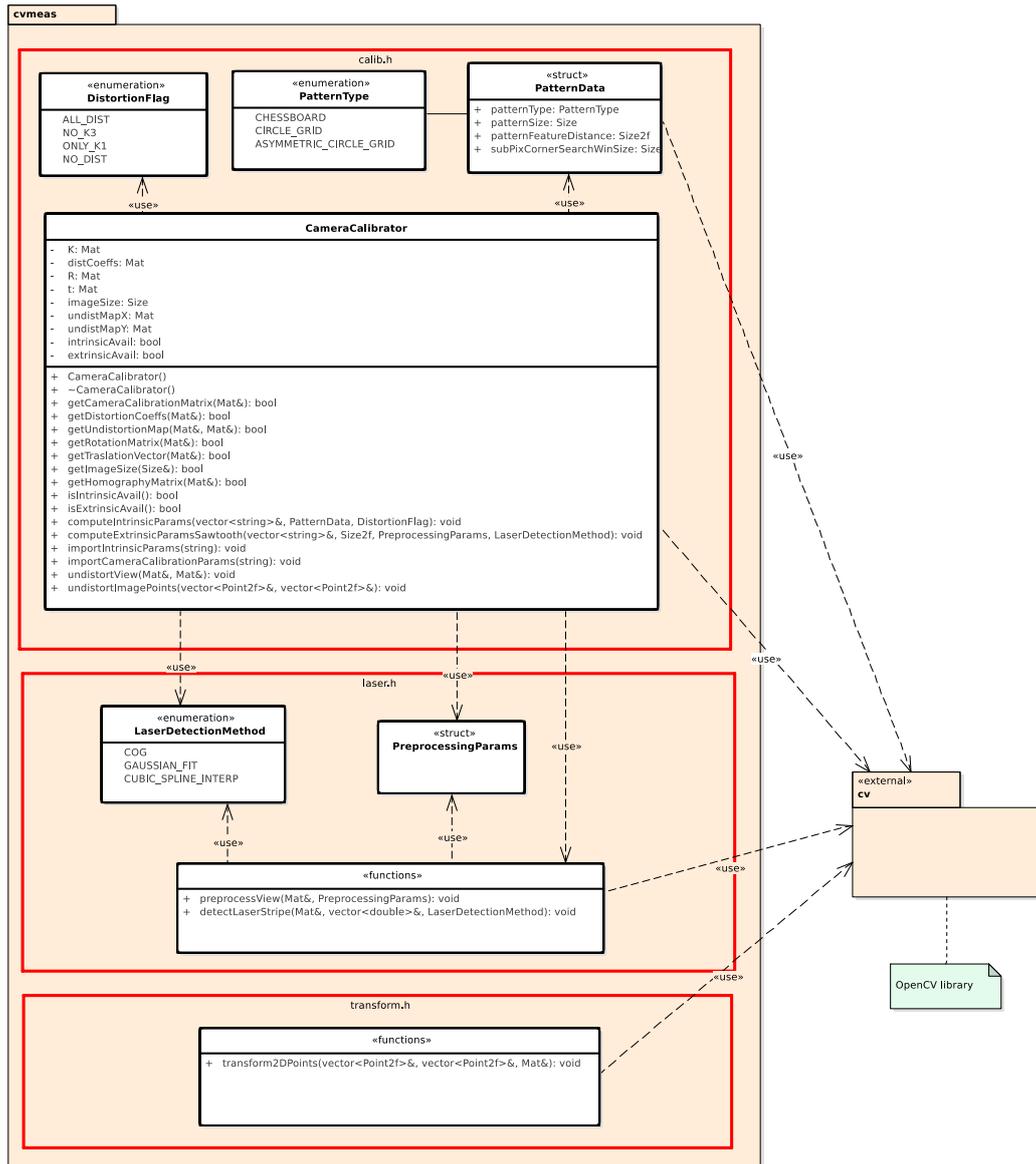


Figura 3.14: Diagramma UML di struttura della libreria CvMeasLib.

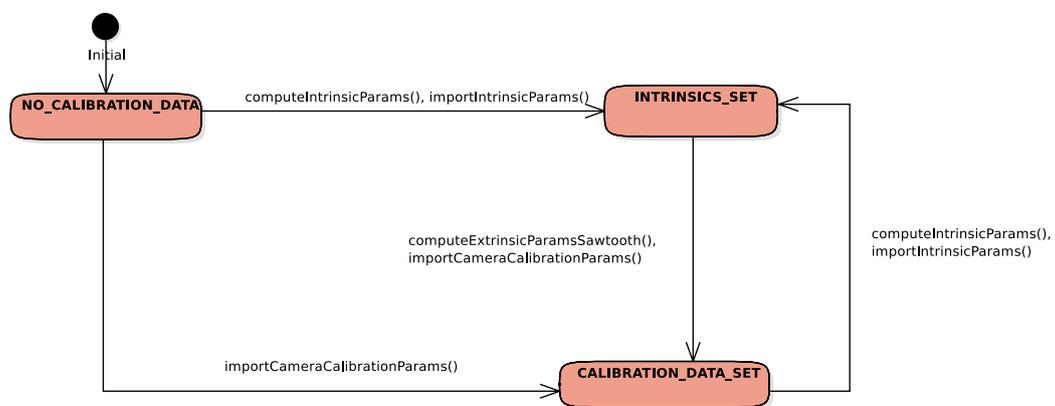


Figura 3.15: Diagramma UML di stato della classe `CameraCalibrator`. I parametri dei metodi sono stati omessi per una migliore leggibilità. Si osservi che nel caso ci si trovi nello stato `CALIBRATION_DATA_SET` e venga effettuato un nuovo calcolo dei parametri interni, gli estrinseci precedentemente calcolati saranno invalidati.

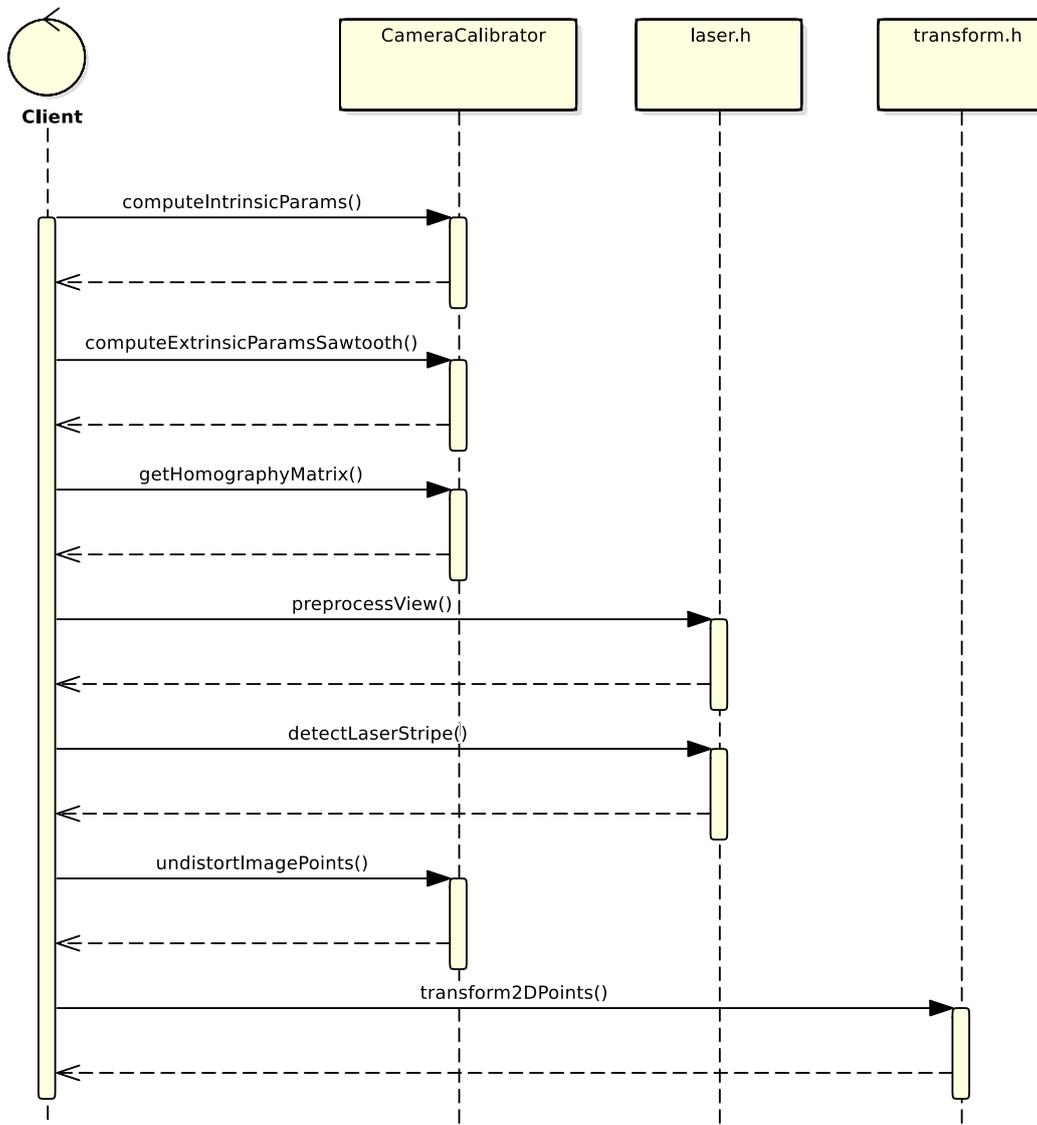


Figura 3.16: Diagramma UML di interazione che schematizza la sequenza di chiamate alla libreria `CvMeasLib` per effettuare la calibrazione della telecamera e la triangolazione di un profilo laser. I parametri delle funzioni sono stati omessi per una migliore leggibilità.

Capitolo 4

Implementazione

4.1 Implementazione del sistema di trascinamento

Nella Figura 4.1 è mostrato lo schema di connessione dei componenti elettronici che costituiscono il sistema di trascinamento. Come si può notare, la scheda Big Easy Driver necessita di un'alimentazione separata (in questo caso è stato utilizzato un alimentatore da 9 V-1 A DC). Il potenziometro CUR ADJ è stato regolato in modo da limitare la massima corrente assorbita dalle due fasi del motore a circa 170 mA; infatti, come detto in precedenza, i consumi di corrente del motore passo-passo a bassi numeri di giri risultano davvero esigui. Si noti inoltre che nella sua configurazione di *default* prevede già la massima risoluzione di *micro-stepping*, ovvero 1/16 di step. La scheda Arduino si interfaccia al *driver* tramite il collegamento dei pin STEP e DIR, che consentono il pilotaggio del motore, ed il pin ENABLE, che può essere utilizzato per disabilitare l'alimentazione alle fasi del motore in caso di inutilizzo.

Per agevolare la codifica della logica Arduino ci si è appoggiati alla libreria AccelStepper, la quale mette a disposizione una classe che implementa metodi per il controllo del motore passo-passo ad alto livello. La classe AccelStepper è stata in questo caso estesa per aggiungere le funzionalità dell'*encoder* emulato, emettendo i due impulsi in quadratura A e B (ed i

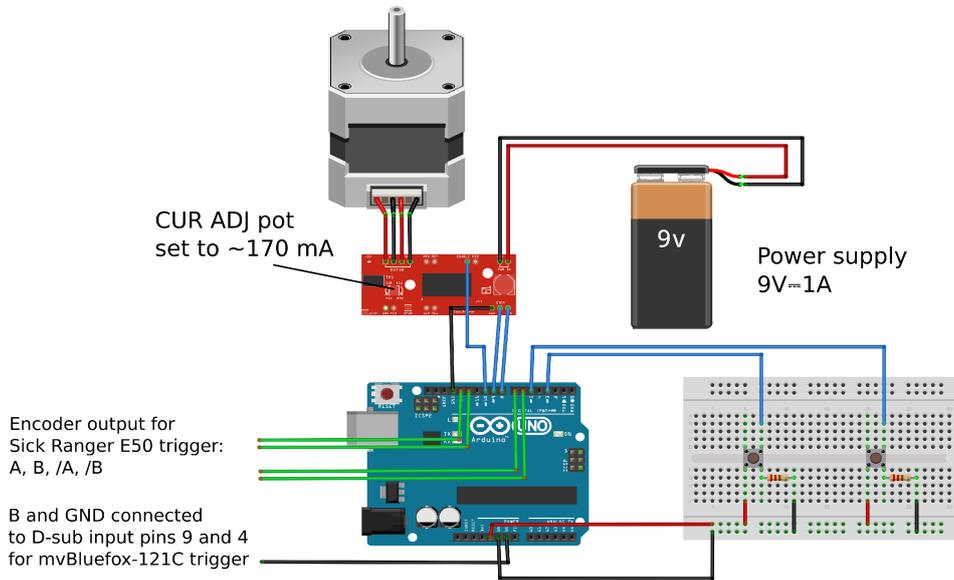


Figura 4.1: Schema di connessione dei componenti elettronici che costituiscono il sistema di trascinamento.

rispettivi negati /A e /B) ogni determinato numero di step percorsi.

La logica di controllo implementa un *parser* di stringhe inviate mediante comunicazione seriale per comandare la rotazione del motore passo-passo e variarne i parametri:

- *SET_SPEED*:<float>: imposta il valore di velocità di rotazione v del motore (espressa in step/s).
- *SET_STEPS*:<int>: imposta il numero di passi s da percorrere.
- *SET_STEPS_TRIG*:<int>: imposta il numero di passi compresi tra l'invio di due impulsi di *trigger*.
- *TRIP_FW*: esegue la rotazione del motore in senso orario (spostamento del piano verso la telecamera) a velocità costante v per un numero di passi s .
- *TRIP_BW*: esegue la rotazione del motore in senso antiorario (spostamento del piano nel verso opposto alla telecamera) a velocità costante v per un numero di passi s .

Un'ulteriore possibilità di controllo è data dalla presenza di due pulsanti, che se premuti attivano la rotazione del motore nei due sensi fino al rilascio.

All'interno del codice, per eseguire uno step viene invocato il metodo `runSpeed()` della classe `AccelStepper`, il quale richiama il metodo privato `step1()`, che setta i livelli di tensione sui pin `STEP` e `DIR` del driver. Tale metodo è stato ridefinito per estenderlo con la possibilità di inviare i quattro segnali digitali dell'encoder ogni numero prefissato di step, come si può osservare nel seguente stralcio di codice.

```
1 //Called by runSpeed(), overridden method from AccelStepper ↵
    class
2 void MyAccelStepper::step1(long step)
3 {
4     ...
5
6     //new code
7     _count++;
8     if(_count == _steps_for_trig)
9     {
10        doEncoderPulse(_direction);
11        _count=0;
12    }
13 }
14
15 //Called by step1()
16 void MyAccelStepper::doEncoderPulse(bool isCW)
17 {
18     //_t4us is 1/4 of the encoder pulse period in ↵
        microseconds. We use 200us as value.
19
20     if(isCW) { // clockwise rotation
21
22         //00
23         writeLevel(LOW,LOW);
24         delayMicroseconds(_t4us);
25
26         //01
```

```
27     writeLevel (LOW, HIGH);
28     delayMicroseconds (_t4us);
29
30     //11
31     writeLevel (HIGH, HIGH);
32     delayMicroseconds (_t4us);
33
34     //10
35     writeLevel (HIGH, LOW);
36     delayMicroseconds (_t4us);
37
38     }else {      // counter-clockwise rotation
39
40         //10
41         writeLevel (HIGH, LOW);
42         delayMicroseconds (_t4us);
43
44         //11
45         writeLevel (HIGH, HIGH);
46         delayMicroseconds (_t4us);
47
48         //01
49         writeLevel (LOW, HIGH);
50         delayMicroseconds (_t4us);
51
52         //00
53         writeLevel (LOW, LOW);
54         delayMicroseconds (_t4us);
55     }
56 }
57
58 //Called by doEncoderPulse()
59 void MyAccelStepper::writeLevel(int levA, int levB)
60 {
61     digitalWrite(_pinA, levA);
62     digitalWrite(_pinAn, HIGH-levA);
63     digitalWrite(_pinB, levB);
64     digitalWrite(_pinBn, HIGH-levB);
65 }
```

4.2 Implementazione della libreria software

Si passa ora alla trattazione dei dettagli implementativi della libreria `Cv-MeasLib`, concentrandosi in particolare sugli algoritmi per la calibrazione della telecamera e il rilevamento del profilo laser.

4.2.1 Implementazione tecniche di calibrazione della telecamera

Calibrazione parametri intrinseci e coefficienti di distorsione

La calibrazione dei parametri intrinseci (matrice \mathbf{K}) e dei coefficienti di distorsione è implementata nel metodo `computeIntrinsicParams` della classe `CameraCalibrator`. La struttura `PatternData` passata in input contiene informazioni sulle caratteristiche del *pattern* planare:

- `patternType`, il tipo di pattern. L'algoritmo ne supporta tre: scacchiera (`CHESSBOARD`), griglia di cerchi (`CIRCLE_GRID`) e griglia di cerchi sfasati (`ASYMMETRIC_CIRCLE_GRID`), che è quello adottato nella sperimentazione.
- `patternSize`, che specifica la larghezza e l'altezza del pattern in termini di *features*.
- `patternFeatureDistance`, la distanza (espressa in una data unità di misura) tra le features del pattern lungo l'altezza e la larghezza.
- `subPixCornerSearchWinSize`, dimensione della finestra per il raffinamento a *sub-pixel* dei *corner* estratti nel caso di pattern a scacchiera.

È inoltre possibile scegliere quali coefficienti di distorsione stimare impostando il flag `DistortionFlag`. Le scelte possibili sono quattro:

- `ALL_DIST`, vengono stimati tutti e cinque i coefficienti di distorsione.
- `NO_K3`, vengono stimati solo k_1 , k_2 , p_1 e p_2 , fissando k_3 a zero.

- ONLY_K1, viene stimato solo k_1 , fissando tutti gli altri a zero.
- NO_DIST, tutti i coefficienti di distorsione sono fissati a zero.

L'algoritmo procede iterando il vettore dei *paths* delle n immagini fornito in input, rilevando in ognuna di esse le features del pattern planare. Nel caso di pattern a scacchiera viene invocata la funzione OpenCV `findChessboard`, seguita da `cornerSubPix` per il raffinamento a sub-pixel dei corner. Nel caso dei due pattern a griglia di cerchi viene invocata la funzione `findCirclesGrid`, che stima i centroidi dei cerchi. In tutti e tre i casi, prima del rilevamento viene effettuata anche un'espansione del contrasto per migliorare l'accuratezza della procedura di estrazione delle features.

In seguito, per ognuna delle immagini, le features estratte vengono accoppiate con i rispettivi punti sul piano del pattern nella scena, le cui coordinate possono essere calcolate avendo noti i valori di distanza delle features lungo altezza e larghezza.

Il metodo di calibrazione vero e proprio è implementato all'interno della funzione OpenCV `calibrateCamera`, basato sulla tecnica descritta in [18], che stima la matrice \mathbf{K} , i coefficienti di distorsione e le n coppie (\mathbf{R}, \mathbf{t}) che descrivono le diverse pose della telecamera rispetto al pattern. In output viene restituito anche il valore RMSE (*Root Mean Squared Error*) in pixel della riproiezione dei punti effettuata mediante i parametri ottenuti.

Calibrazione parametri estrinseci

La calibrazione dei parametri estrinseci (matrice \mathbf{R} e vettore \mathbf{t}) è implementata nel metodo `computeExtrinsicParamsSawtooth` della classe `CameraCalibrator`. Esso procede iterando il vettore dei *paths* delle n immagini fornito in input; per ciascuna di esse vengono invocate le funzioni `preprocessView` e `detectLaserStripe`, dichiarate nel modulo `laser.h`, per estrarre la linea laser. A partire dai punti del profilo laser così ottenuti, viene corretta la distorsione e vengono stimate, mediante *fitting* ai minimi quadrati, le rette che seguono l'andamento degli spigoli del dente di sega. Dall'intersezione di tali rette vengono in seguito calcolate le coordinate

immagine dei punti di picco e valle del calibratore a dente di sega (ignorando la raggiatura), come mostrato in Figura 4.2.

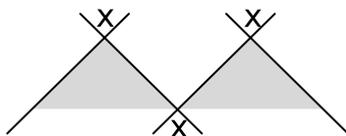


Figura 4.2: L'intersezione delle rette lungo gli spigoli del calibratore a dente di sega consente di calcolare le coordinate immagine dei picchi e delle valli (ignorando la raggiatura).

I punti dell' i -esima immagine analizzata vengono messi in corrispondenza con i rispettivi punti di picco e valle sul piano laser, dei quali è possibile assegnare le coordinate in un sistema di riferimento relativo all'oggetto, in quanto si conosce l'altezza e la larghezza in millimetri del singolo dente, come mostrato nella Figura 4.3. A partire da tali corrispondenze si stima così l'omografia $\tilde{\mathbf{H}}_i$.

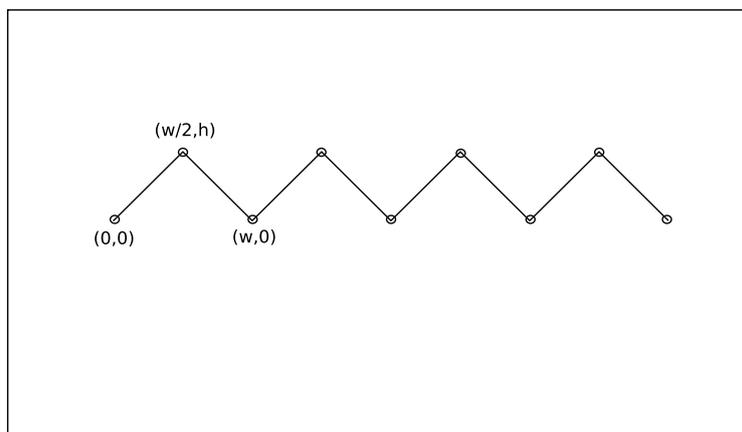


Figura 4.3: Il sistema di riferimento adottato per le coordinate dei punti del calibratore a dente di sega sul piano laser. L'origine è fissata sulla valle più a sinistra e tutte le altre sono assegnate rispetto a quest'ultima.

Dato che le n matrici di omografia stimate (una per ogni vista di input) mappano i punti del piano immagine sul piano laser, ma ciascuna in un differente sistema di riferimento per quest'ultimo, si vuole calcolare la trasformazione euclidea che consente di mappare le coordinate dei punti sul piano

laser da un sistema di riferimento relativo ad uno assoluto. Presa l' i -esima immagine, la proiezione da piano laser con sistema di riferimento assoluto a piano immagine è descritta dall'omografia

$$\mathbf{H} = \tilde{\mathbf{H}}_i \mathbf{T}_i \mathbf{R}_i$$

È possibile stimare la traslazione \mathbf{T}_i e la rotazione \mathbf{R}_i del sistema di riferimento avendo note due corrispondenze:

- il punto immagine \mathbf{u}_0 è la proiezione dell'origine \mathbf{X}_0 del sistema di riferimento assoluto sul piano laser;
- il punto immagine \mathbf{u}_r è la proiezione di un punto \mathbf{X}_r sull'asse X del sistema di riferimento assoluto sul piano laser.

Si fa presente che anche in questo caso i punti sono espressi in coordinate omogenee e le uguaglianze sono definite a meno di un fattore di scala.

Dalla seguente relazione

$$\mathbf{u}_0 = \mathbf{H}\mathbf{X}_0$$

si ottiene

$$\tilde{\mathbf{H}}_i^{-1} \mathbf{u}_0 = \mathbf{T}_i \mathbf{X}_0$$

in quanto la trasformazione di rotazione non ha effetto sull'origine. Dunque la matrice \mathbf{T}_i può essere ricavata come

$$\mathbf{T}_i = \begin{bmatrix} 1 & 0 & | & \\ 0 & 1 & | & \tilde{\mathbf{H}}_i^{-1} \mathbf{u}_0 \\ 0 & 0 & | & \end{bmatrix}$$

dove il vettore $\tilde{\mathbf{H}}_i^{-1} \mathbf{u}_0$ è espresso in coordinate omogenee normalizzate.

Similmente, dalla relazione

$$\mathbf{u}_r = \mathbf{H}\mathbf{X}_r$$

si ottiene

$$\mathbf{T}_i^{-1} \tilde{\mathbf{H}}_i^{-1} \mathbf{u}_0 = \mathbf{R}_i \mathbf{X}_0$$

Per ricavare \mathbf{R}_i è necessario calcolare l'angolo θ tra l'asse X ed il vettore

$$\mathbf{r} = \begin{bmatrix} r_X \\ r_Z \\ 1 \end{bmatrix} = \mathbf{T}_i^{-1} \tilde{\mathbf{H}}_i^{-1} \mathbf{u}_0$$

dove il vettore $\mathbf{T}_i^{-1} \tilde{\mathbf{H}}_i^{-1} \mathbf{u}_0$ è espresso in coordinate omogenee normalizzate. L'angolo θ è calcolabile come

$$\theta = \arctan\left(\frac{r_Z}{r_X}\right)$$

A questo punto \mathbf{R}_i può essere determinata

$$\mathbf{R}_i = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

e di conseguenza anche la matrice \mathbf{H} .

Per ottenere il set finale di corrispondenze tra piano immagine e piano laser mediante le quali stimare i parametri estrinseci, si suddivide l'immagine in una griglia di punti (ad esempio 5×5). Per ciascuno di questi punti, il corrispondente punto sul piano laser è dato dalla media delle n riproiezioni del punto immagine utilizzando le inverse delle n matrici \mathbf{H} stimate precedentemente, ipotizzando che siano affetti da errore gaussiano. I parametri estrinseci vengono infine calcolati richiamando la funzione di OpenCV `solvePNP`.

4.2.2 Implementazione tecniche di rilevamento del profilo laser

La funzione `detectLaserStripe` definita nel modulo `laser.h` implementa le tre tecniche di rilevamento dei punti del profilo laser introdotte in fase di progetto:

- calcolo del *center-of-gravity* del profilo d'intensità (COG);
- *fitting* di una funzione gaussiana sui pixel del profilo di intensità e ricerca del picco (GAUSSIAN_FIT);
- interpolazione dei pixel del profilo di intensità mediante spline cubica naturale e ricerca del picco (CUBIC_SPLINE_INTERP).

In ognuna di esse viene eseguita un'analisi *column-by-column* del profilo d'intensità del laser lungo le righe; quindi, iterando dalla prima all'ultima riga, si procede alla ricerca di sequenze di pixel adiacenti diversi da zero, e per ciascuna di esse si esegue il calcolo dell'indice di riga del picco stimato, mantenendo solo quello corrispondente al valore di intensità maggiore. Si noti che l'analisi effettuata all'interno di una colonna è indipendente da tutte le altre, quindi per una maggiore efficienza è stata implementata la possibilità di parallelizzare la computazione in *tasks* indipendenti mediante direttive OpenMP 2.0¹. Il modo in cui viene calcolato l'indice di riga del picco varia a seconda della tecnica scelta:

- Nel caso dell'algoritmo COG, tale indice di posizione viene calcolato come media degli indici di riga pesata sui valori d'intensità dei pixel. Questa tecnica risulta più robusta dell'algoritmo Hi3D COG implementato nella telecamera Sick Ranger E50, che effettua invece il calcolo del *center-of-gravity* sull'intera colonna, tenendo anche conto quindi di eventuali pixel di rumore nello sfondo sopravvissuti alla sogliatura.
- Nel caso dell'algoritmo GAUSSIAN_FIT, per il fitting di una funzione gaussiana sul profilo d'intensità è stato implementato il metodo descritto in [10]. La risoluzione della valutazione della spline interpolata è pari a 1/64 di pixel.
- Nel caso dell'algoritmo CUBIC_SPLINE_INTERP, per l'interpolazione di una spline cubica naturale è stato implementato il metodo descritto in [13]. La risoluzione della valutazione della funzione fittata è pari a 1/64 di pixel.

¹<http://www.openmp.org/>

La funzione `preprocessView` definita nel modulo `laser.h` implementa il *pre-processing* dell'immagine, composto, come detto nel Paragrafo 3.3.2, da una fase di sogliatura e da una di *smoothing*. Per quanto riguarda la sogliatura sono state inserite tre tecniche, per ognuna delle quali si effettuerà una prima analisi qualitativa dei risultati in modo da scegliere la migliore da utilizzare nella successiva fase di sperimentazione:

- sogliatura globale con valore di soglia scelto automaticamente mediante metodo di Otsu (`GLOBAL_OTSU`);
- sogliatura globale con valore di soglia passato come parametro (`GLOBAL_MANUAL`);
- sogliatura adattiva in cui il valore di soglia è pari a $mean - C$, dove $mean$ è la media delle intensità calcolata in un intorno quadrato del pixel. Dunque tale metodo richiede due parametri: la dimensione della finestra in cui calcolare la media locale ed una costante C (`ADAPTIVE`).

Lo *smoothing* è stato invece implementato mediante convoluzione dell'immagine con un filtro gaussiano, del quale è possibile settare la dimensione ed il valore di σ (nel caso di $\sigma = 0$, verrà calcolato automaticamente in base alla dimensione del filtro).

Nella Figura 4.4 sono mostrati i punti del profilo rilevato utilizzando la tecnica COG sull'immagine del calibratore a dente di sega (pezzo non diffusivo) e dell'osso compatto (pezzo diffusivo) acquisite con la telecamera mvBluefox-121C, senza aver eseguito alcun pre-processing. Si può facilmente notare come in entrambi i casi il profilo presenti un andamento alquanto irregolare e poco accurato. Nella Figura 4.5a e 4.5c si osservano degli ottimi risultati per la sogliatura globale automatica e manuale del calibratore, mentre per l'osso si ottiene una sogliatura poco soddisfacente, come si può notare nella Figura 4.5b e 4.5d. Infatti, il metodo di Otsu calcola un valore di soglia che non riesce a cancellare per intero l'alone di *scattering*, mentre impostandone manualmente uno in grado di eliminarlo si taglia anche parte del segnale utile. Per tali motivi, in caso di una marcata presenza di scattering intorno al profilo laser, una sogliatura di tipo globale si dimostra poco flessibile, e per di

più, con quella manuale, il valore di soglia ottimale deve essere valutato empiricamente, fattore che costituisce un ulteriore punto a sfavore. Nella Figura 4.5e e 4.5f si può osservare come con la sogliatura adattiva si riescano ad ottenere ottimi risultati per entrambe le immagini, seppur anche in questo caso vi sia il difetto di dover tarare a priori i valori ottimali da adottare per i due parametri.

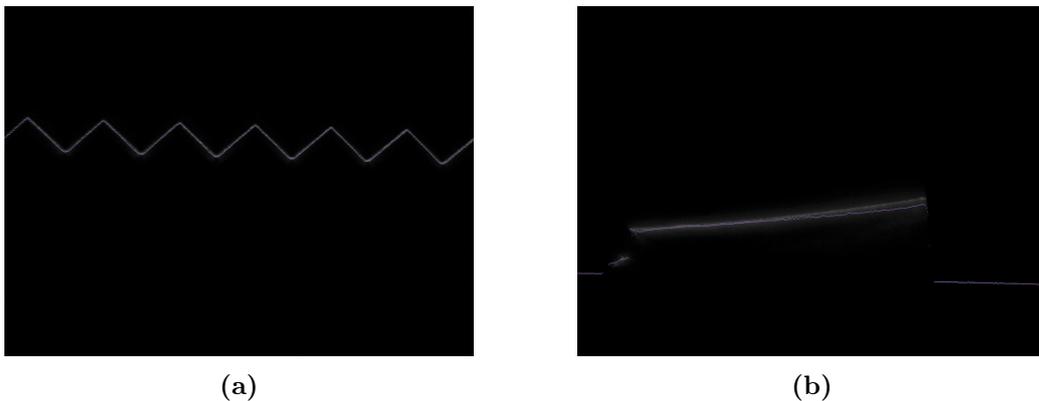


Figura 4.4: Profilo laser rilevato con COG su immagine del calibratore a dente di sega (a) e osso compatto (b) senza pre-processing. In questa e nelle prossime figure, i punti del profilo sovrainposti alle immagini sono disegnati a risoluzione di pixel interi.

La Figura 4.6 mostra il risultato della procedura di smoothing gaussiano effettuata a valle della sogliatura adattiva, nel caso del calibratore a dente di sega e osso compatto. Il valore ottimale per la dimensione del filtro va scelto in base allo spessore della striscia laser; dimensioni opportune nel caso delle immagini presentate vanno da 7×7 a 11×11 .

Infine, nella Figura 4.7 viene mostrato il rilevamento con COG a seguito delle due fasi di pre-processing.

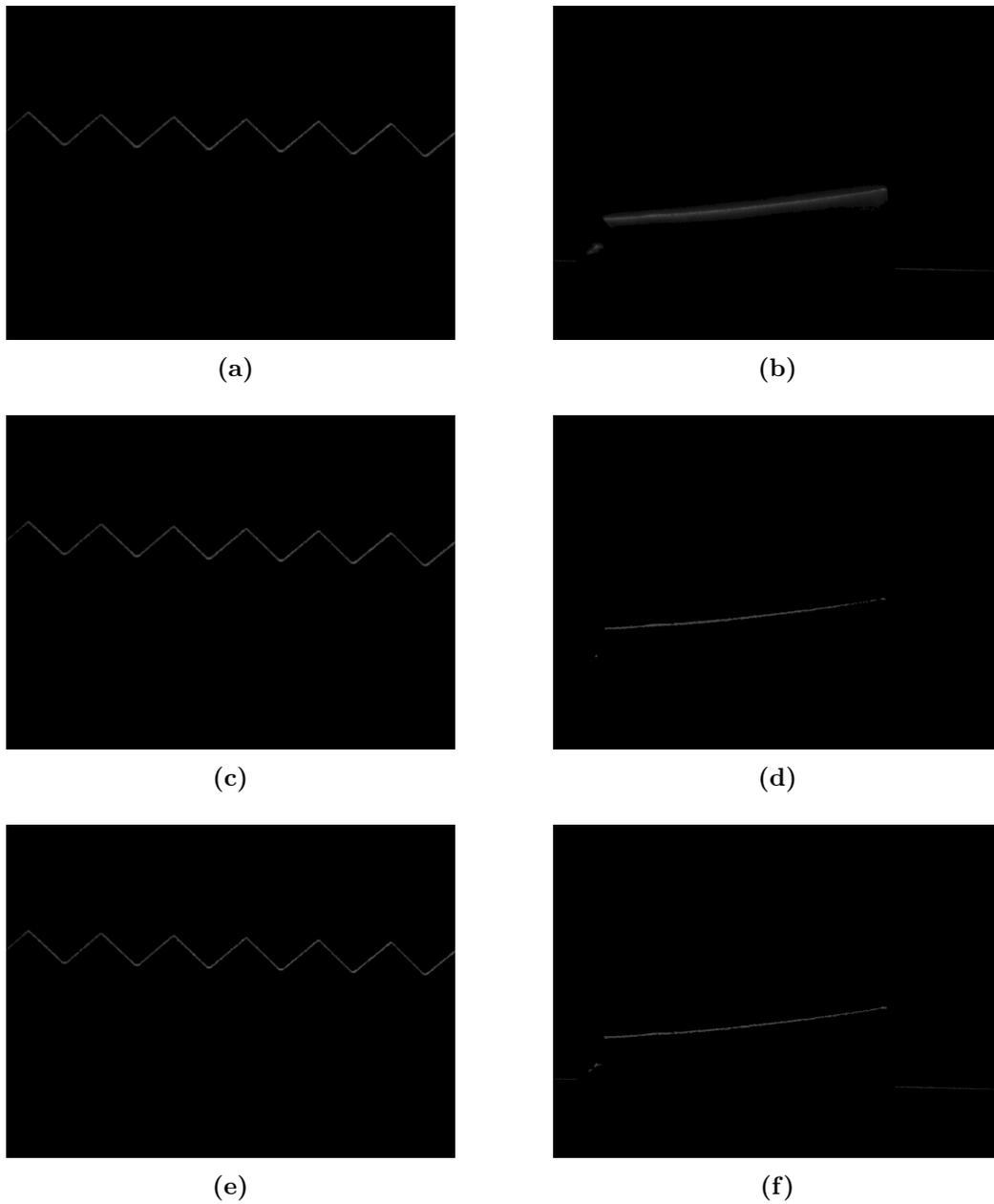


Figura 4.5: Sogliatura globale automatica su immagine del calibratore a dente di sega (a) e osso compatto (b). Sogliatura globale manuale su immagine del calibratore a dente di sega ($thresh = 26$) (c) e osso compatto ($thresh = 51$) (d). Sogliatura adattiva su immagine del calibratore a dente di sega ($size = 17 \times 17$, $C = -7$) (e) e osso compatto ($size = 17 \times 17$, $C = -9$) (f).

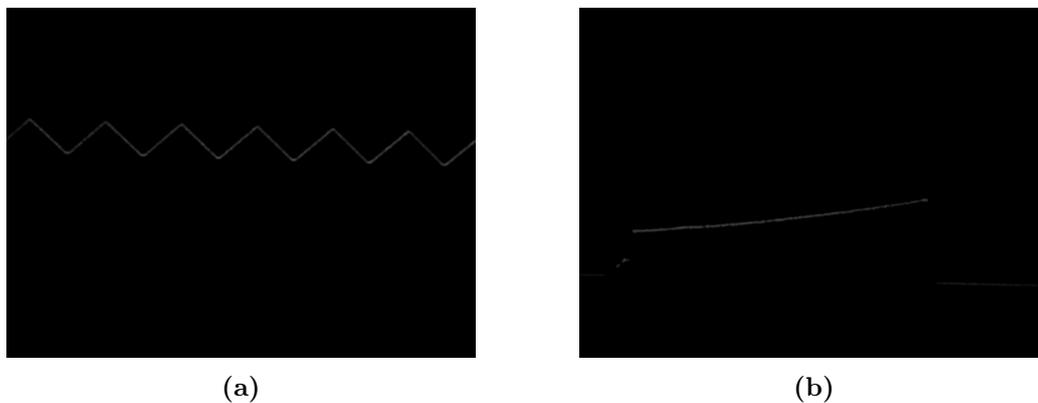


Figura 4.6: Risultato dello smoothing gaussiano a seguito della sogliatura adattiva su immagine del calibratore a dente di sega ($size = 7 \times 7$) (a) e osso compatto ($size = 7 \times 7$) (b).

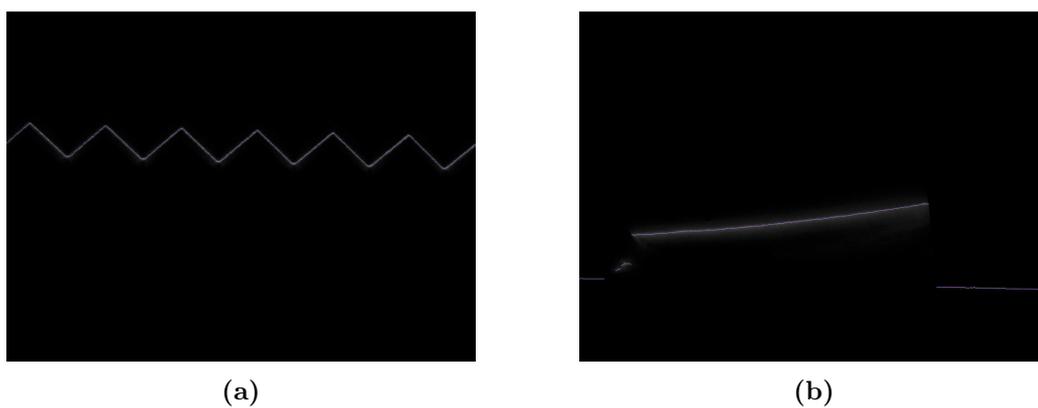


Figura 4.7: Profilo laser rilevato con COG su immagine del calibratore a dente di sega (a) e osso compatto (b) a seguito del pre-processing.

Capitolo 5

Risultati sperimentali

5.1 Risultati della calibrazione delle telecamere

In questo paragrafo vengono presentati i risultati della calibrazione della telecamera mvBluefox-121C eseguita con gli algoritmi di libreria implementati in questa tesi, e quelli della calibrazione della telecamera Sick Ranger E50 eseguita tramite il software in dotazione 3D Coordinator.

5.1.1 Risultati calibrazione mvBluefox-121C

Per la calibrazione dei parametri intrinseci e dei coefficienti di distorsione della telecamera mvBluefox-121C sono state utilizzate sei viste del *pattern* 3×9 di tipo *asymmetric circles*, mostrate in Figura 5.1. Al fine di eseguire una buona calibrazione è importante fornire viste non ridondanti tra loro, cercando di variarne il più possibile l'orientazione e la posizione all'interno del campo visivo. Nel caso della telecamera utilizzata si è notato che la profondità di campo risultava assai ristretta, dunque non è stato possibile inclinare fin troppo il pattern senza evitare che esso sfocasse. Si è deciso di fissare a zero il coefficiente k_3 , in quanto è consigliabile stimarlo solo nel caso di lenti che presentano un'elevata distorsione radiale (per esempio obiettivi *fish-eye*) [4]. Avendo a che fare con obiettivi professionali ed un FOV di pochi centimetri, si potrebbe pensare di semplificare il calcolo dei coefficienti di distorsione

stimando solamente k_1 . In ogni caso, la calibrazione eseguita in questa sede ed adottata nelle successive sperimentazioni stima k_1 , k_2 , p_1 e p_2 .

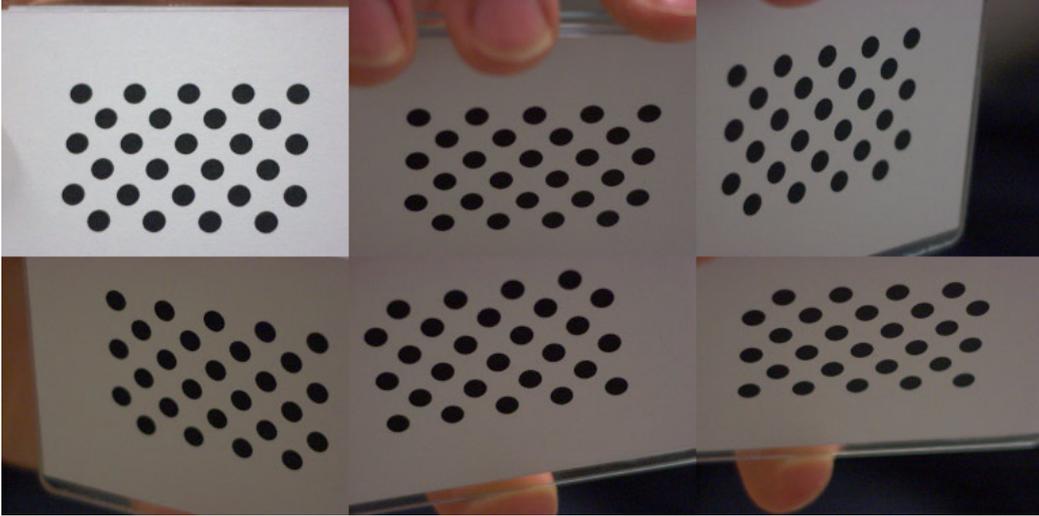


Figura 5.1: Le sei viste del pattern asymmetric circles utilizzate per la calibrazione dei parametri intrinseci e dei coefficienti di distorsione della telecamera mvBluefox-121C.

Effettuando esperimenti con svariati set di viste, questo è quello con cui si è ottenuto il valore di RMSE di riproiezione più basso, ovvero pari a 0.24 pixel ($RMSE_u = 0.11$ pixel, $RMSE_v = 0.21$ pixel). La Figura 5.2 mostra le *features* del pattern rilevate nelle sei viste. La Tabella 5.1 mostra le stime dei valori di α_u , α_v , u_0 e v_0 contenuti nella matrice \mathbf{K} , e dei quattro coefficienti di distorsione.

α_u	α_v	u_0	v_0	k_1	k_2	p_1	p_2
5717.130	5849.547	546.931	297.162	0.344	-38.411	-0.002	0.003

Tabella 5.1: Parametri intrinseci e coefficienti di distorsione della telecamera mvBluefox-121C stimati dall'algoritmo.

Per la calibrazione dei parametri estrinseci è stato utilizzato un set di undici viste del calibratore a dente di sega colpito dalla lama laser, variandone la posizione all'interno del campo visivo della telecamera. Il *pre-processing* è stato eseguito con sogliatura adattiva ($size = 17 \times 17$, $C = -7$) seguita

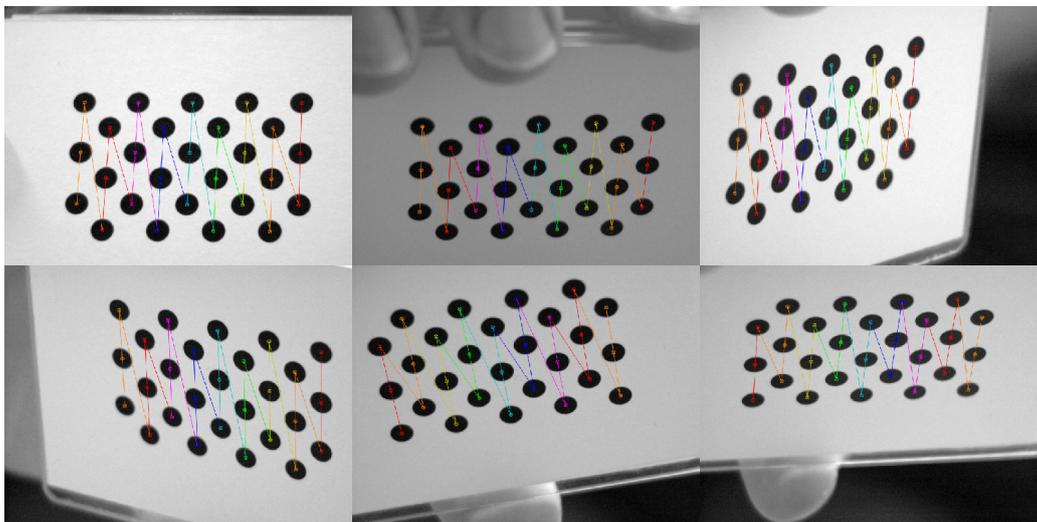


Figura 5.2: Le features del pattern asymmetric circles rilevate nelle sei viste.

da *smoothing* gaussiano con finestra 7×7 , mentre il rilevamento del profilo laser è stato eseguito con il metodo COG. In Figura 5.3 è mostrata una delle viste del calibratore con i punti di picco e valle rilevati dall'algoritmo. Il valore di RMSE di riproiezione delle viste del calibratore è pari a 0.19 pixel ($RMSE_u = 0.16$ pixel, $RMSE_v = 0.11$ pixel).

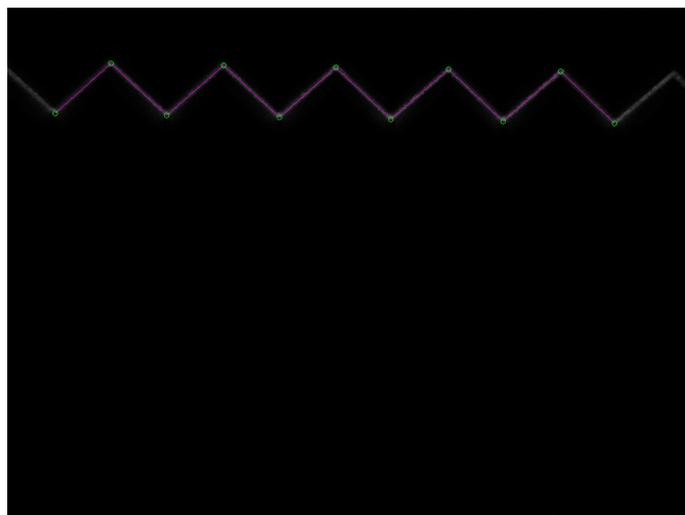


Figura 5.3: Una vista del calibratore a dente di sega con i punti di picco e valle rilevati dall'algoritmo.

5.1.2 Risultati calibrazione Sick Ranger E50

Nella Figura 5.4 sono riportati i risultati della calibrazione della telecamera Sick Ranger E50 eseguita dal software 3D Coordinator. Il riquadro superiore mostra l'errore medio in pixel a seguito della calibrazione delle distorsioni, quello inferiore mostra l'errore medio di misura in millimetri.

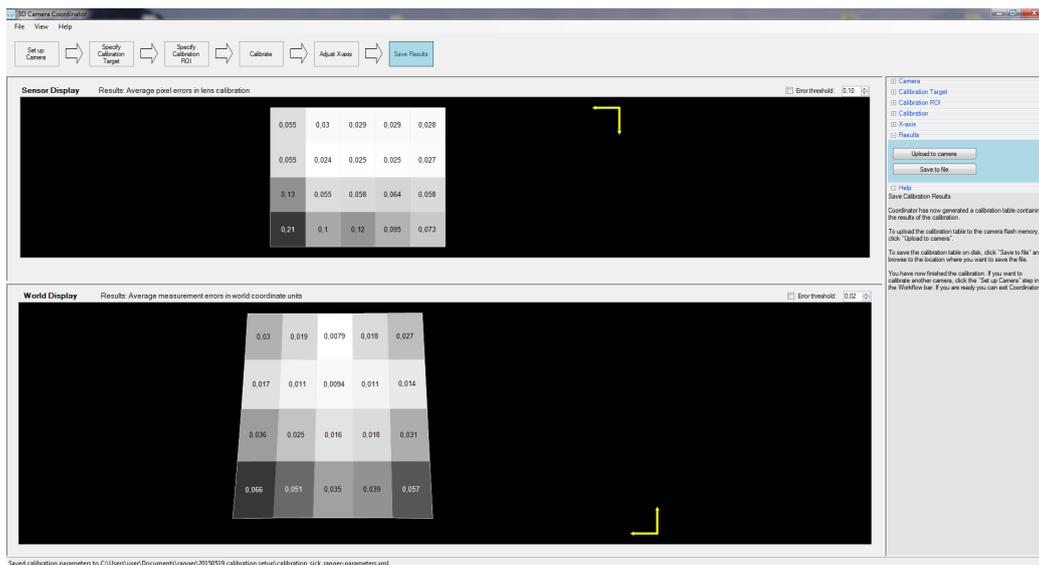


Figura 5.4: Errore di calibrazione della telecamera Sick Ranger E50 riportato dal software 3D Coordinator.

5.2 Risultati della calibrazione del sistema di trascinamento

In questa fase ci si è occupato di caratterizzare in maniera assoluta la velocità (costante) di traslazione del piano mobile (mm/s), per ricavare poi lo spostamento in millimetri compiuto da un singolo step e quantificare quindi il passo di scansione. Il valore di velocità stimato con il motore passo-passo impostato alla rotazione di 400 step/s è pari a 1.523 mm/s. Dato che uno step viene eseguito ogni $1/400 = 0.0025$ s, lo spostamento compiuto da un singolo step è pari a $1.523 \cdot 0.0025 = 0.003808$ mm.

5.3 Risultati della triangolazione

5.3.1 Test con pezzi non diffusivi

Per testare la bontà della calibrazione e dei metodi di rilevamento implementati per la telecamera mvBluefox-121C, si è deciso di effettuare, inizialmente, misurazioni sul caso più semplice, ovvero quello di pezzi non diffusivi. La Figura 5.5 mostra un blocchetto in alluminio del quale è stata misurata la quota in Z della faccia superiore rispetto al piano d'appoggio. Il valore *ground-truth* da utilizzare come riferimento per il confronto delle misure è pari a 25.37 mm, ed è stato stimato in seguito a misurazioni ripetute eseguite con un calibro digitale.

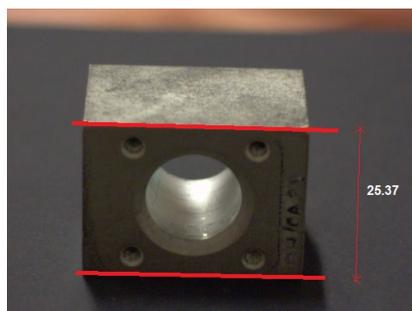


Figura 5.5: Il pezzo in alluminio utilizzato per il test.

Nella Figura 5.6 è riportato il grafico delle quote in Z dei punti del profilo laser rilevati utilizzando le tre tecniche proposte. In tutti e tre i casi si può notare che le quote misurate tendono a distribuirsi intorno al valore di *ground-truth* (si osservi per esempio la distribuzione nel caso COG mostrata in Figura 5.7). L'andamento leggermente oscillante dei dati (visibile soprattutto nel caso del rilevamento tramite CUBIC_SPLINE_INTERP, che è la tecnica più sensibile al rumore), è dovuto all'*aliasing* dei pixel della linea laser; lo *smoothing* gaussiano eseguito in *pre-processing* serve anche a ridurre questo effetto. La Tabella 5.2 mostra, per ognuno dei tre metodi di rilevamento utilizzati, alcune statistiche sui dati acquisiti, come la quota minima e massima, la media e la deviazione standard (per valutare accuratezza e precisione dell'acquisizione), ed il valore RMSE.

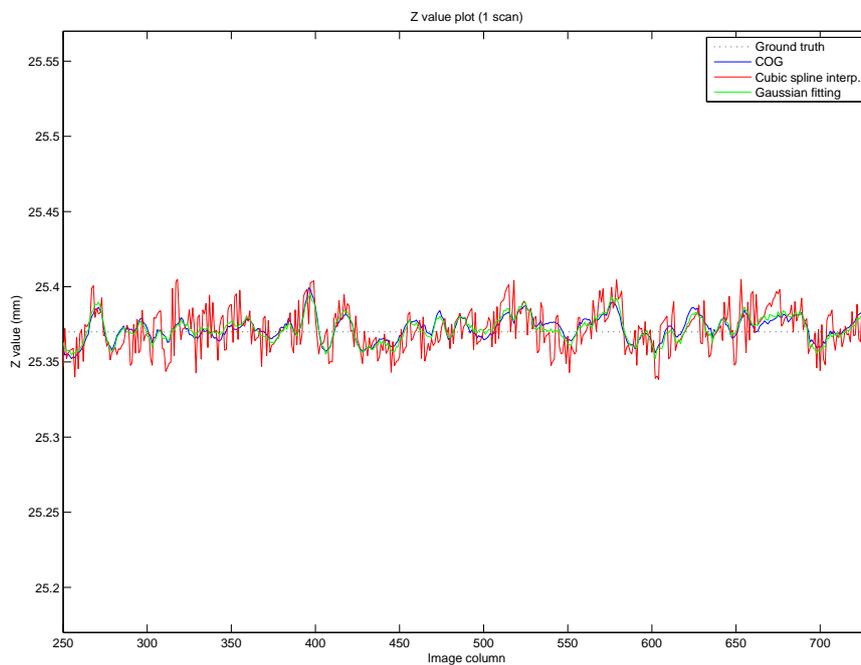


Figura 5.6: Grafico delle quote lungo Z dei punti del profilo laser del pezzo di alluminio. Il pre-processing è stato effettuato con sogliatura adattiva ($size = 17 \times 17$, $C = -3$) seguita da smoothing gaussiano con finestra 11×11 .

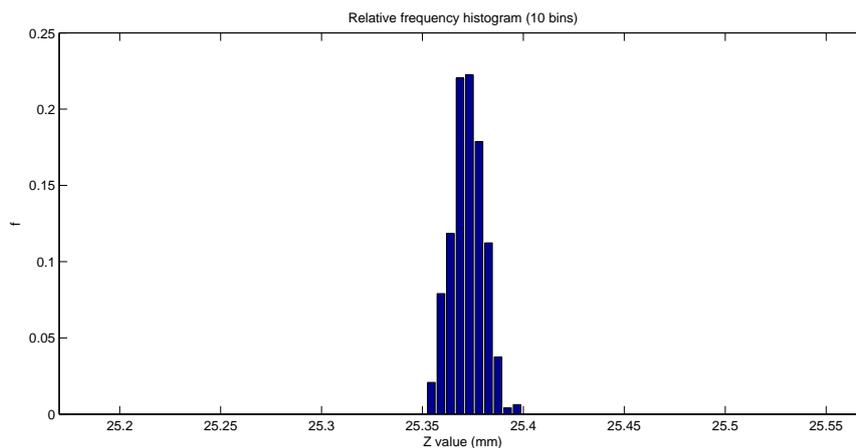


Figura 5.7: Istogramma delle frequenze delle quote lungo Z dei punti del profilo laser del pezzo di alluminio (caso rilevamento tramite metodo COG).

	$minZ$	$maxZ$	\bar{Z}	\bar{s}	$RMSE$
COG	25.352	25.399	25.372	0.008	0.008
GAUSSIAN_FIT	25.352	25.395	25.372	0.008	0.009
CUBIC_SPLINE_INTERP	25.338	25.405	25.372	0.014	0.014

Tabella 5.2

Nella Tabella 5.3 sono riportati i tempi medi in millisecondi di triangolazione del profilo (calcolati su 100 acquisizioni ripetute) per i tre metodi utilizzati, sia nel caso di implementazione sequenziale che di implementazione parallelizzata. Il *time-profiling* è stato eseguito su una macchina che monta una CPU AMD FX-8320 (8 cores) a 3.5 GHz con 8 MB di cache e 8 GB di RAM.

	COG	GAUSSIAN_FIT	CUBIC_SPLINE_INTERP
sequenziale	10.136 ms	27.920 ms	46.276 ms
parallelizzata	9.690 ms	16.727 ms	19.098 ms

Tabella 5.3

Lo stesso profilo è stato acquisito anche con la telecamera Sick Ranger E50 adottando l'algoritmo Hi3D COG. In questo caso, non potendo ottenere le quote rispetto al piano d'appoggio, è stato stimato sui punti il *fitting* ai minimi quadrati di una retta e calcolato l'errore di ciascun punto rispetto ad essa (Figura 5.8).

La Figura 5.9 mostra le *point cloud* ottenute a seguito della scansione del calibratore a dente di sega con le due telecamere, impostando la velocità di rotazione del motore passo-passo a 100 step/s ed il passo di scansione a 0.011423 mm. Le *point cloud* sono state allineate con il modello CAD del calibratore e si è valutata la distanza di ciascun punto rispetto ad esso (Figura 5.10).

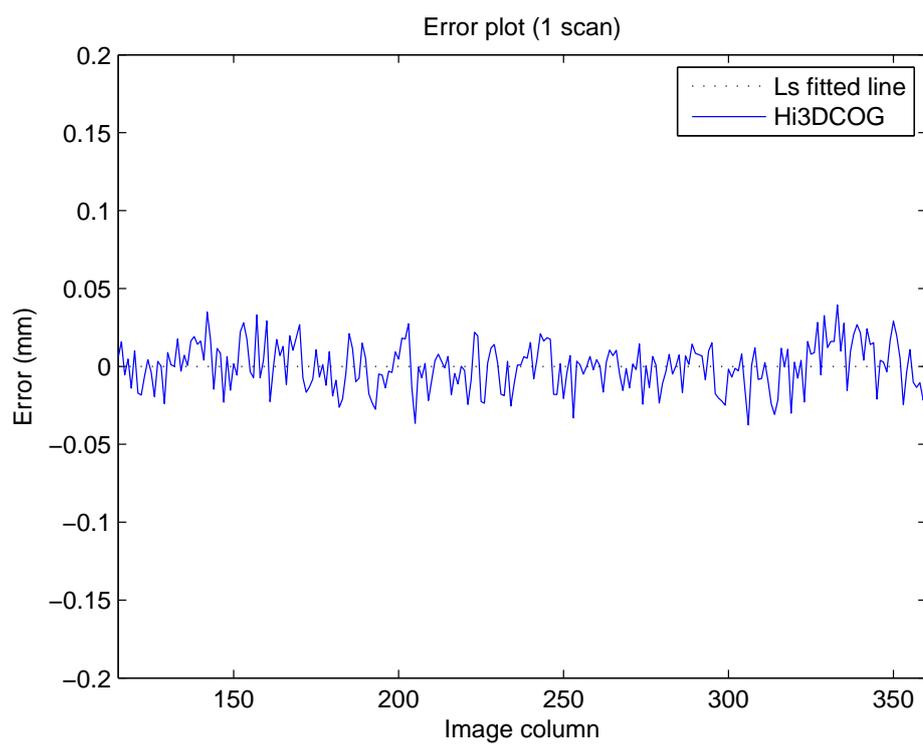
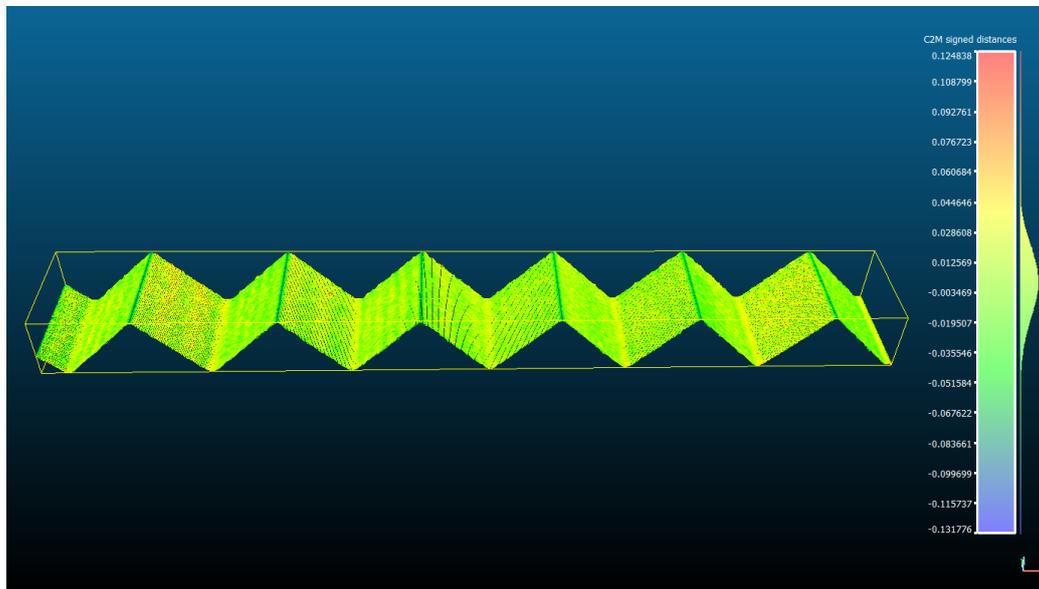
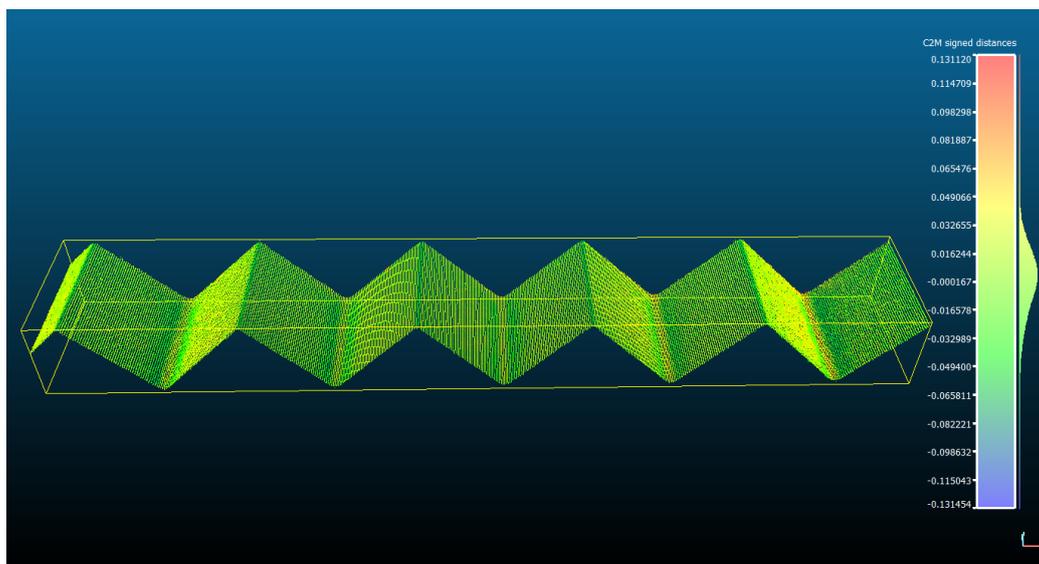


Figura 5.8: Grafico dell'errore dei punti del profilo laser del pezzo di alluminio rispetto alla retta fittata ai minimi quadrati. Acquisizione con telecamera Sick Ranger E50, rilevamento con metodo *embedded* Hi3D COG ($thresh = 14$). $minErr = -0.038$, $maxErr = 0.039$, $\bar{Err} = 0.000$, $\bar{s} = 0.015$.

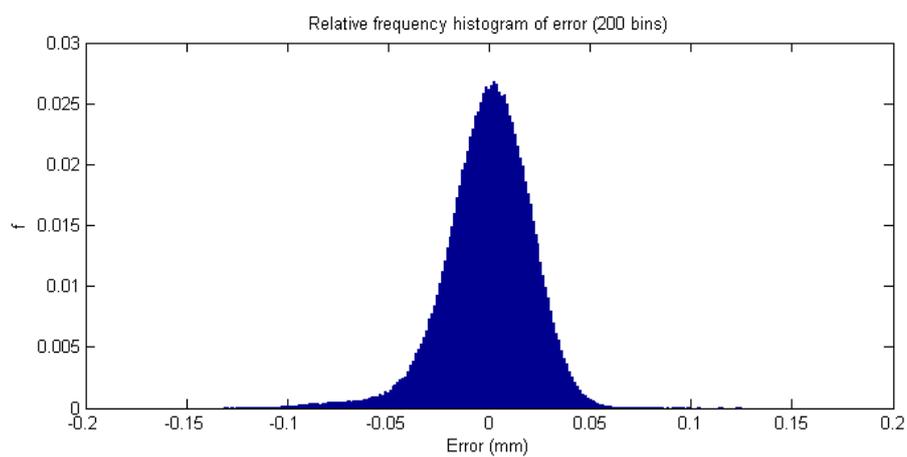


(a)

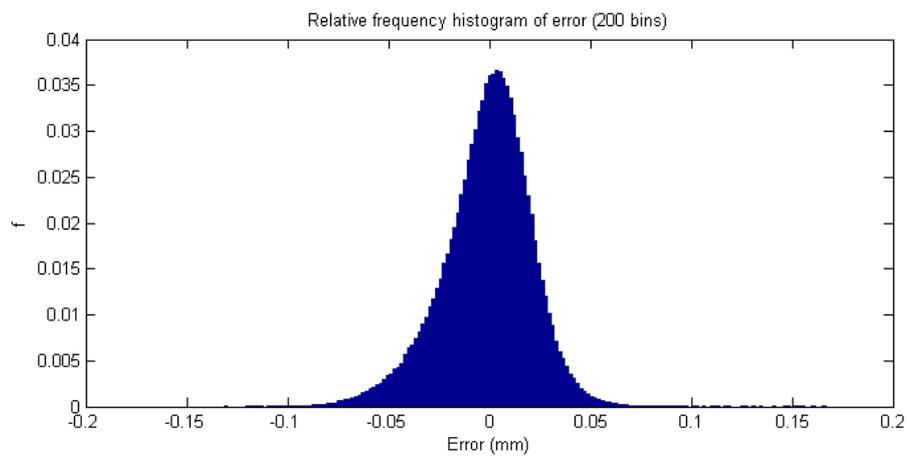


(b)

Figura 5.9: Le point cloud ottenute a seguito della scansione del calibratore a dente di sega e allineate con il modello CAD. (a) mvBluefox-121C, rilevamento con COG, pre-processing con sogliatura adattiva ($size = 17 \times 17$, $C = -7$) seguita da smoothing gaussiano con finestra 7×7 . (b) Sick Ranger E50 (Hi3D COG, $thresh = 14$).



(a)



(b)

Figura 5.10: Istogramma delle frequenze dell'errore della point cloud rispetto al modello del calibratore nel caso di (a) mvBluefox-121C e (b) Sick Ranger E50.

5.3.2 Test con pezzi diffusivi

La Figura 5.12 mostra le *point cloud* ottenute a seguito della scansione di un pezzo in poliossimetilene (materiale ad elevato potere diffusivo) con le due telecamere, impostando la velocità di rotazione del motore passo-passo a 100 step/s ed il passo di scansione a 0.030460 mm. Le point cloud sono state allineate con il modello CAD del pezzo (Figura 5.11).

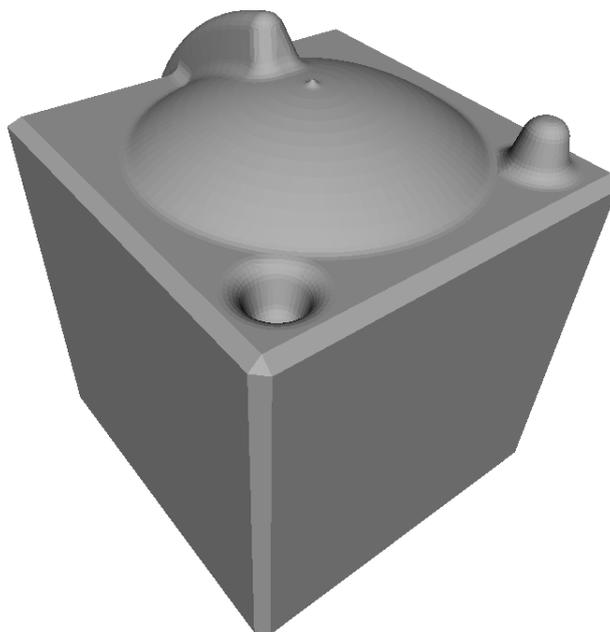
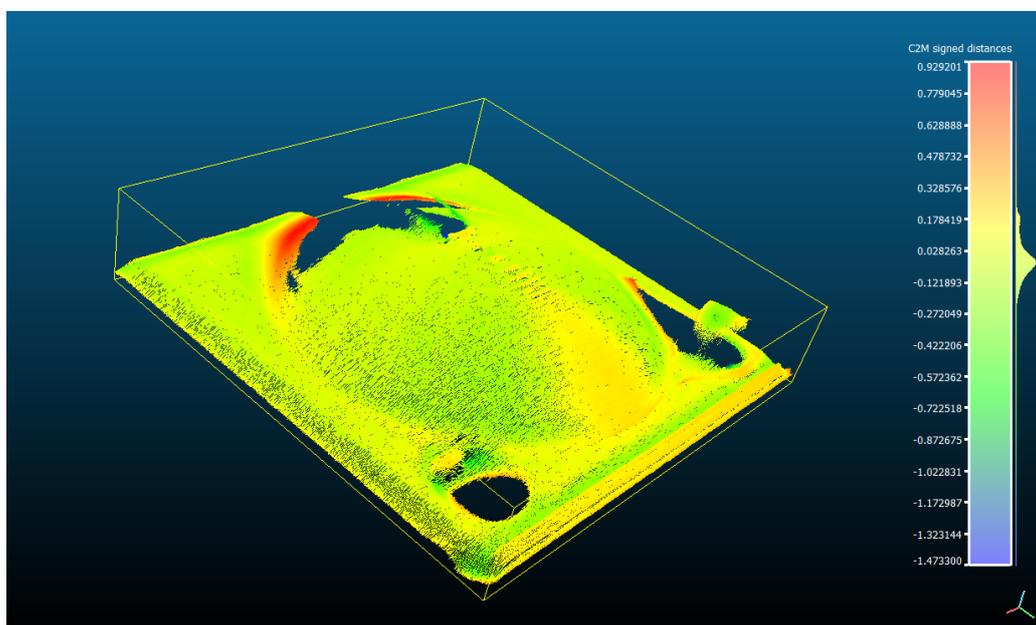


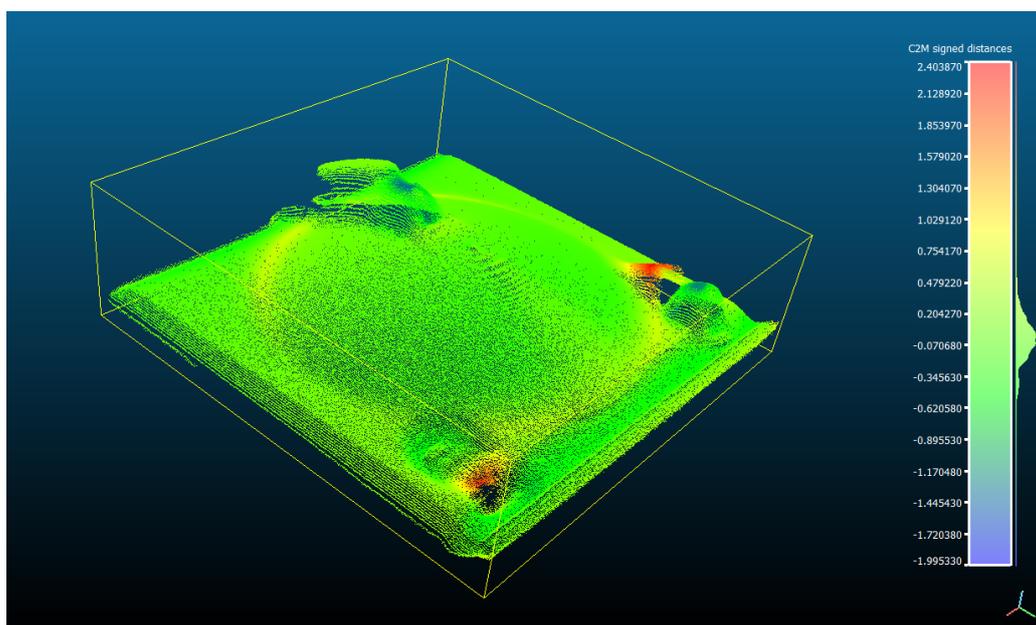
Figura 5.11: Il modello CAD del pezzo in poliossimetilene.

Gli algoritmi sono stati testati anche per ricostruire le due facce di un pezzo d'osso (Figura 5.13). In questo caso ci si è limitati unicamente ad un'analisi qualitativa del profilo laser rilevato, rimandando ad altra sede la caratterizzazione approfondita dell'errore con questo tipo di superficie.

Nella Figura 5.14a e 5.14b sono mostrati due profili sulle due facce dell'osso rilevati con la telecamera mvBluefox-121C, mentre nella Figura 5.14c e 5.14d sono mostrati quelli rilevati con la telecamera Sick Ranger E50. Si noti come in questo caso l'algoritmo Hi3D COG della Sick Ranger E50 non sia in grado di individuare in modo accurato il profilo laser dell'osso nell'immagine a causa dello *scattering* (non del tutto eliminabile dalla sogliatura eseguita dal metodo) che devia sensibilmente la stima del *center-of-gravity*.



(a)



(b)

Figura 5.12: Le point cloud ottenute a seguito della scansione del pezzo in poliossimetilene e allineate con il modello CAD. (a) mvBluefox-121C, rilevamento con COG, pre-processing con sogliatura adattiva ($size = 49 \times 49$, $C = -7$) seguita da smoothing gaussiano con finestra 17×17 . (b) Sick Ranger E50 (Hi3D COG, $thresh = 14$).

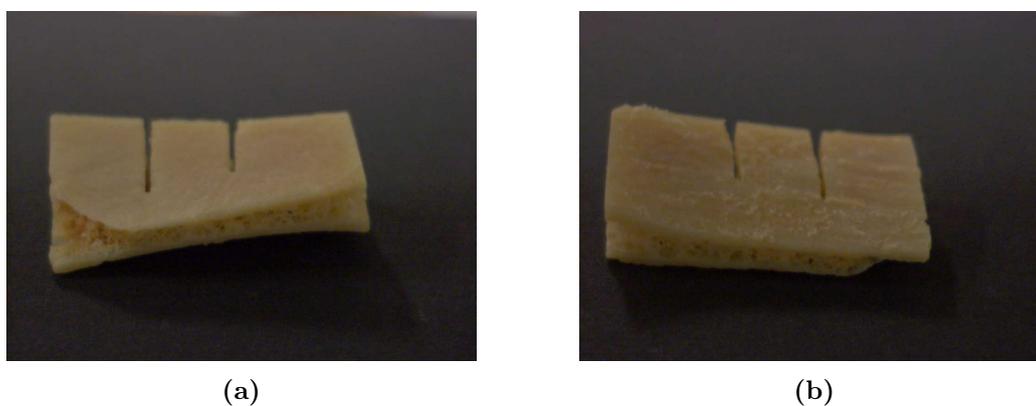


Figura 5.13: (a) Faccia A osso. (b) Faccia B osso.

Infine, nella Figura 5.15 sono mostrate le *mesh* di triangoli delle due facce dell'osso ricostruite a partire dalla scansione con telecamera mvBluefox-121C (scansione a 100 step/sec, passo di 0.030460 mm).

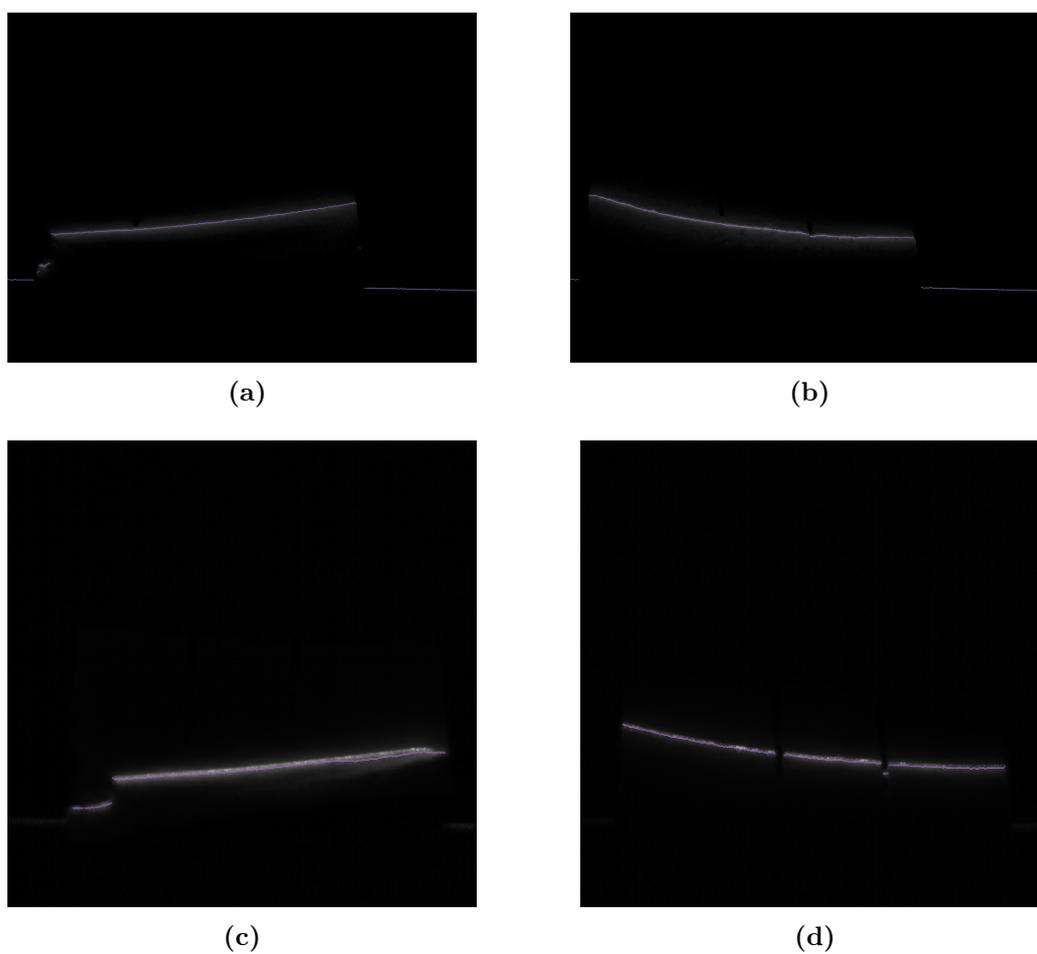
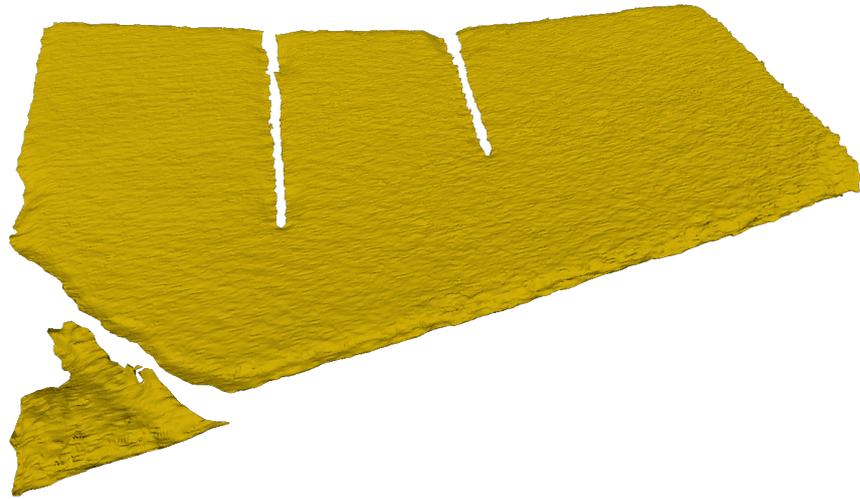
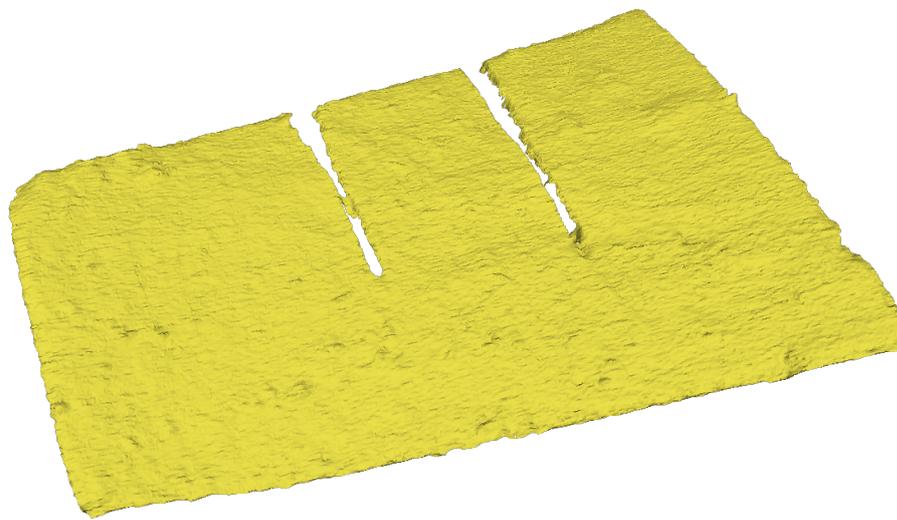


Figura 5.14: Profilo laser dell'osso rilevato con telecamera mvBluefox-121C per la faccia A (a) e B (b). Rilevamento con COG, pre-processing con sogliatura adattiva ($size = 17 \times 17$, $C = -9$) seguita da smoothing gaussiano con finestra 7×7 . Profilo laser dell'osso rilevato con telecamera Sick Ranger E50 (Hi3D COG, $thresh = 30$) per la faccia A (c) e B (d).



(a)



(b)

Figura 5.15: Le mesh di triangoli della faccia A (a) e B (b) dell'osso ricostruite a partire dalle point cloud acquisite con la telecamera mvBluefox-121C. Le mesh sono state calcolate utilizzando l'algoritmo *Delaunay Triangulation* implementato in CloudCompare.

Conclusioni e sviluppi futuri

Nel corso di questo lavoro di tesi è stato trattato il progetto di un sistema di triangolazione a lama laser, sia nella sua parte hardware che software. Si è partiti dalla scelta della telecamera e del proiettore laser da adottare, per poi passare alla loro configurazione geometrica e allo sviluppo di un semplice prototipo per automatizzare la scansione traslando il pezzo attraverso la lama laser. Per quanto riguarda il lato software, è stato effettuato lo studio dei passi procedurali che compongono la risoluzione al problema della triangolazione, concretizzati poi nel successivo sviluppo di una libreria C++ `CvMeasLib` per la calibrazione della telecamera, il rilevamento del profilo nell'immagine (che comprende una fase di *pre-processing* ed il rilevamento vero e proprio, per il quale sono state implementate tre differenti tecniche) e la riproiezione sul piano della lama laser. I componenti adottati sono tutti a basso costo, dunque è stato di nostro interesse valutare la qualità dei risultati rispetto a quelli ottenuti utilizzando un sistema commerciale di triangolazione (Sick Ranger E50).

Nel semplice caso di acquisizioni di pezzi non diffusivi e con una buona riflettività, si è riscontrato per entrambe le telecamere un grado di accuratezza che rientra nei $10\ \mu m$, con errori che tendono a non superare i $50\ \mu m$. Per quanto riguarda il disturbo dovuto al *subsurface scattering*, si è verificato che gli algoritmi di pre-processing sviluppati sono in grado di eliminarlo in modo opportuno, ma solo nel caso in cui vi sia un adeguato gradiente d'intensità tra la componente riflessa del laser (segnale utile) e quella di disturbo, come nel caso dell'osso. Dalle prove con il pezzo in poliossimetilene, infatti, si è notato che tale materiale tende a riflettere la luce laser in un ristretto cono di direzioni; dunque, nel caso in cui la superficie non sia orientata verso il sensore,

l'intensità della componente riflessa si confonde con quella dello scattering, rendendo impossibile un'adeguata segmentazione e portando a sovrastimare e sottostimare svariate aree della superficie, con picchi di errore di qualche decimo di millimetro.

La telecamera Sick Ranger E50, sebbene sia un sistema *high-end*, non risulta adatta a gestire acquisizioni di superfici diffuse, in quanto il semplice pre-processing mediante sogliatura manuale eseguito dall'algoritmo dedicato (Hi3D COG) spesso non è sufficiente ad eliminare per intero lo scattering senza tagliare anche parte del segnale utile. Si tratta comunque di un sistema che fa della velocità di scansione il suo punto di forza, dunque implementa algoritmi semplici e adatti soprattutto per l'ispezione industriale di prodotti.

In definitiva, il lavoro svolto in questa tesi costituisce il punto di partenza per il futuro sviluppo del sottosistema di acquisizione e ricostruzione 3D che farà parte di un apparato per la realizzazione di protesi d'osso su misura in camera bianca. In questo senso, sviluppi futuri del progetto comprendono: il raffinamento delle tecniche di rilevamento della linea laser, in particolare per ciò che concerne la fase di pre-processing, i cui risultati desiderati dipendono da un'adeguata taratura preliminare dei parametri; lo sviluppo di tecniche per la registrazione di nuvole di punti e la successiva integrazione in un unico modello tridimensionale CAD. A tale scopo ci si servirà anche di un sistema di attuazione più complesso del prototipo costruito in questa tesi, in grado di muovere l'oggetto con più gradi di libertà.

Bibliografia

- [1] D. Akça, A. Grün, B. Breuckmann, and C. Lahanier. High Definition 3D-Scanning of Arts Objects and Paintings. In *Optical 3-D Measurement Techniques VIII, Volume II*, pages 50–58, 2007.
- [2] K. Barat. *Laser Safety Management*. CRC Press, 2006.
- [3] F. Bernardini and H. Rushmeier. The 3D Model Acquisition Pipeline. *Computer Graphics Forum*, 21(2):149–172, 2002.
- [4] G. Bradsky and A. Kaehler. *Learning OpenCV*. O’Reilly, 2008.
- [5] F. Chiabrandò, R. Chiabrandò, D. Piatti, and F. Rinaudo. Sensors for 3D Imaging: Metric Evaluation and Calibration of a CCD/CMOS Time-of-Flight Camera. *Sensors*, 9(12):10080–10096, 2009.
- [6] B. Curless and S. Seitz. 3D Photography. ACM Siggraph ’00 Course Notes. 2000.
- [7] R. B. Fisher and K. Konolige. Range Sensors. In B. Siciliano and O. Khattib, editors, *Springer Handbook of Robotics*, pages 521–542. Springer Berlin Heidelberg, 2008.
- [8] J. Geng. Structured-light 3D surface imaging: a tutorial. *Advances in Optics and Photonics*, 3(2):128–160, 2011.
- [9] G. Godin, M. Rioux, J. A. Beraldin, M. Levoy, L. Cournoyer, and F. Blais. An Assessment of Laser Range Measurement of Marble Surfaces. In *Proceedings of the 5th Conference on Optical 3-D Measurement*, pages 49–56, 2001.

-
- [10] H. Guo. A Simple Algorithm for Fitting a Gaussian Function. *IEEE Signal Processing Magazine*, 28(5):134–137, 2011.
- [11] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.
- [12] D. K. Naidu and R. B. Fisher. A Comparative Analysis of Algorithms for Determining the Peak Position of a Stripe to Sub-pixel Accuracy. In *Proceedings of the British Machine Vision Conference*, pages 217–225, 1991.
- [13] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, third edition, 2007.
- [14] U. Schulz and K. Böhnke. Calibration Methods for a 3D Triangulation Based Camera. In A. Gottscheber, D. Obdržálek, and C. Schmidt, editors, *Research and Education in Robotics - EUROBOT 2009*, volume 82 of *Communications in Computer and Information Science*, pages 131–143. Springer Berlin Heidelberg, 2010.
- [15] SickAG. 3D Coordinator Reference Manual. <http://www.sick.com/>, 2014.
- [16] SickAG. Sick Ranger E/D Reference Manual. <http://www.sick.com/>, 2014.
- [17] C. Smochinǎ and V. Manta. Industrial inspection system using triangulation. 2008.
- [18] Z. Zhang. A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.