

POLITECNICO DI MILANO

Dipartimento di Meccanica

Corso di Dottorato in Ingegneria Meccanica



STUDIO E QUALIFICAZIONE METROLOGICA DI
SOLUZIONI DI MISURA BASATE SULLA STEREOSCOPIA

Tesi di Dottorato di:
Giacomo MAINETTI

Supervisor:

Prof. Remo Sala

Tutor:

Prof. Fabio Fossati

Coordinatore del Programma di Dottorato:

Prof.ssa Bianca Maria Colosimo

2015 - XXVII Ciclo

A Ludovica

Ringraziamenti

Vorrei prima di tutto ringraziare il prof. Remo Sala che mi ha dato la possibilità di intraprendere questa esperienza tanto costruttiva quanto interessante e mi ha aiutato e supportato in questo percorso. Ringrazio inoltre il mio tutor prof. Fabio Fossati per i suoi importantissimi consigli e la Coordinatrice del Programma di Dottorato prof.ssa Bianca Maria Colosimo per il lavoro che ogni giorno compie egregiamente per amministrare le carriere dei dottorandi, insieme a Silvia Barattieri. Ringrazio il prof. Alfredo Cigada per il suo prezioso aiuto che mi ha sempre prontamente offerto e gli altri professori del dipartimento con cui ho avuto il piacere di lavorare in questi anni. Un ringraziamento speciale va allo staff del VBLab ed ai miei colleghi Ambra e Silvio, con i quali ho condiviso buona parte delle avventure di questi anni e hanno reso le giornate in laboratorio sempre piacevoli. Ringrazio Ale Basso, che fino all'ultimo non ha smesso di darmi suggerimenti preziosissimi e i colleghi Mario, Ale P., Andrea e Ale C. oltre che tutto il personale della Sezione di Misure e della Galleria del Vento. Un grazie va anche ai tesisti con cui ho collaborato, con la speranza di avergli dato almeno quanto loro hanno dato a me. Durante la mia attività di ricerca mi sono trovato più volte a lavorare anche con società esterne al Politecnico di Milano e ci tengo a ringraziare tutte le persone che ho incontrato durante questo cammino, interne o esterne all'università, che hanno contribuito alla mia crescita professionale. Infine ringrazio tutte le persone che al di fuori dell'ambito lavorativo mi supportano costantemente. In particolare tutti gli amici con cui ho condiviso i sempre più importanti momenti di svago: una bella sciata, una sgambata o una birra al bar. Ringrazio la mia enorme famiglia che mi è sempre vicina con il suo affetto e che in questi ultimi anni si è ufficialmente allargata in maniera spropositata; in particolare i miei genitori Chicca ed Enea e mio fratello Teo che sono sempre una spalla sicura su cui appoggiarsi. In ultimo per ordine, ma primo per importanza, un grazie speciale va a mia moglie Ludovica, alla quale è dedicato questo lavoro, per aver accettato il pesante incarico di sopportarmi tutti i giorni della nostra vita.

Abstract

Lo sviluppo della tecnologia ha permesso negli ultimi anni la forte diffusione di tecniche di misura senza contatto 3D in molti ambiti applicativi e in particolar modo in campo industriale. Tra le tecniche 3D esistenti si vuole approfondire quella stereoscopica. Scopo di questo lavoro è dunque lo studio e la qualificazione metrologica di sistemi stereoscopici basati sull'utilizzo di più telecamere. La scelta della stereoscopia è dettata dalla necessità di realizzare un sistema di misura utilizzabile in un'ampia varietà di ambiti applicativi, con diverse condizioni ambientali e volumi di lavoro e con la capacità di effettuare misure dinamiche.

In questo lavoro è descritta la qualificazione metrologica di un sistema stereoscopico a tre telecamere e l'applicazione della stereoscopia in tre diversi ambiti. Per l'analisi delle incertezze di misura di un sistema stereoscopico è necessario prima di tutto studiare il processo di taratura e capire come questo influenzi la misura finale. Una parte del lavoro è dunque dedicata allo studio ed al confronto dei principali algoritmi di calibrazione di telecamere esistenti in letteratura. La caratterizzazione metrologica del sistema stereoscopico ha portato alla luce numerosi accorgimenti da rispettare in fase di taratura e ha mostrato come da questa dipenda fortemente l'accuratezza delle misure tridimensionali. Inoltre, mediante il metodo di Monte Carlo, è stata studiata la propagazione dell'incertezza in fase di taratura del sistema stereoscopico ed è stata poi effettuata un'analisi di sensitività della procedura di ricostruzione tridimensionale.

I risultati della qualificazione metrologica del sistema stereoscopico a tre telecamere sono stati poi utilizzati per realizzare tre sistemi di misura ottimizzati per tre diversi ambiti applicativi. In particolare le tre problematiche affrontate sono: misure in galleria del vento, programmazione automatica di robot industriali e guida robot per la presa da cassone di parti alla rinfusa. Per ognuna di queste è stato realizzato un sistema stereoscopico dedicato che è stato metrologicamente qualificato e validato mediante una serie di analisi sperimentali.

Abstract

In many application field, especially in the industrial one, 3D non-contact sensors are always more diffused, thanks to technology development. One of these techniques is the object of this work, which aim is the study and metrological characterization of multi-cameras stereoscopic systems. The need of a measuring system usable in a wide variety of application fields, with different environment conditions and working volumes and able to perform dynamic measurements, has led to the choice of stereoscopy.

In order to perform a metrological characterization of a stereoscopic system, it is very important to analyze the calibration process and investigate how this affects the final 3D measures. The first part of the work is therefore dedicated to the study and comparison of the main camera calibration algorithms existing in literature. The study of the accuracy of a three camera stereoscopic system has underlined how this is strongly influenced by calibration procedures. Different experiment has been performed to investigate this dependency and all the precautions that must be taken to improve measurement accuracy. Furthermore, using Monte Carlo method, the uncertainty propagation in calibration process has been studied and a sensitivity analysis of the three-dimensional reconstruction procedure has been performed.

The results of the metrological qualification of a three-camera system has been used to implement solutions optimized for three different application fields: wind tunnel measurements, automatic robot programming and robot vision for bin-picking of randomly disposed mechanical parts. For each of these applications, a dedicated stereoscopic system has been implemented and it has been metrologically qualified and validated through a series of experimental analysis.

Indice

Indice	5
1. Introduzione	7
2. Stato dell'arte	11
2.1 Stato dell'arte delle tecniche di ricostruzione 3D.....	12
2.1.1 Tecniche passive	13
2.1.2 Tecniche attive	16
2.2 Stato dell'arte delle tecniche di calibrazione di telecamere	20
3. Metodologia di misura	24
3.1 Stereoscopia.....	25
3.2 Il modello lineare della telecamera.....	28
3.3 Il modello non lineare della telecamera	36
3.4 Geometria epipolare.....	40
4. Calibrazione di telecamere	46
4.1 Calibrazione mediante il metodo di Zhang e di Levenberg-Marquardt	47
4.1.1 Acquisizione dei dati di calibrazione	48
4.1.2 Calcolo del modello lineare	51
4.1.3 Ottimizzazione non lineare	55
4.2 La tecnica di Heikkilä	61
4.2.1 Confronto con la tecnica di Zhang-LM	70
5. Attività sperimentale	76
5.1 Setup del sistema	77
5.1.1 Parametri delle telecamere	77
5.1.2 Scelta delle telecamere	81
5.1.3 Altri componenti hardware	83
5.1.4 Setup finale	84
5.1.5 Software	86

5.2 Taratura dello strumento	86
5.3 Studio del legame tra errore 3D e errore 2D	88
5.4 Efficacia della calibrazione al variare del numero di pose utilizzate ...	95
5.5 Studio dell'accuratezza di misura all'esterno dal volume di calibrazione.....	101
5.6 Studio della propagazione dell'incertezza mediante il metodo Monte Carlo.....	103
5.7 Analisi di sensitività.....	124
6. Applicazioni innovative	132
6.1 Misure in Galleria del Vento	133
6.1.1 Setup del sistema	133
6.1.2 Taratura del sistema	135
6.1.3 Misura della forma di vele	139
6.1.4 Misura di strutture civili	149
6.2 Robot pittore.....	153
6.2.1 Setup e taratura del sistema	154
6.2.2 Analisi dell'incertezza di misura	160
6.2.3 Incertezza della ricostruzione 3D	160
6.2.4 Incertezza del sistema lato Robot	163
6.2.5 Misurazioni Camera Vs Robot	165
6.2.6 Incertezza del sistema di acquisizione finale	168
6.2.7 Conclusioni	174
6.3 Bin picking	176
6.3.1 Setup e taratura del sistema	178
6.3.2 Illuminazione	187
6.3.3 Scelta delle posa della telecamera	192
6.3.4 Validazione del sistema stereoscopico	200
6.3.5 Conclusioni e confronto con la tecnica a triangolazione laser	213
7. Conclusioni	217
Bibliografia	224

1. Introduzione

L'utilizzo di tecniche di misura 3D senza contatto sta diventando sempre più diffuso in diversi ambiti applicativi ed in particolar modo in campo industriale. Lo sviluppo di sistemi hardware più performanti a costi contenuti ha contribuito ad una progressiva diffusione di queste tecniche in molti ambiti applicativi. Queste tecniche consentono di effettuare misure di posizione e spostamento del misurando in totale assenza di effetti di carico su di esso, in quanto sfruttano la radiazione elettromagnetica riflessa e/o emessa. Tra le diverse soluzioni esistenti, in questo lavoro, verrà studiata nel dettaglio quella stereoscopica, vale a dire un metodo ottico di ricostruzione tridimensionale basato sull'utilizzo di più immagini. La scelta della stereoscopia è stata dettata dalla necessità di realizzare un sistema di misura che potesse essere utilizzato in un'ampia varietà di ambiti applicativi, con diverse condizioni ambientali e volumi di lavoro e diverse e con la capacità di effettuare misure dinamiche.

Il lavoro descrive la qualificazione metrologica di un sistema stereoscopico generico e l'applicazione della stereoscopia a tre diversi ambiti: misure in galleria del vento, programmazione automatica di robot industriali e guida robot per la presa da cassone di parti alla rinfusa.

La qualificazione metrologica è stata effettuata su un sistema stereoscopico a tre telecamere realizzato in laboratorio. Pur essendo possibile effettuare misure

stereoscopiche utilizzando due telecamere, tuttavia si è preferito utilizzarne tre per irrobustire la procedura di accoppiamento automatico dei punti omologhi delle diverse immagini. Allo scopo di stimare l'incertezza di misura tipica dei sistemi stereoscopici è necessario studiare il processo di calibrazione delle singole telecamere che lo compongono e capire come esso influenza la misura finale. Una parte del lavoro è dunque dedicata allo studio ed al confronto dei principali algoritmi di calibrazione esistenti in letteratura. In particolare si è voluto indagare le potenzialità della tecnica di Heikkilä [1], confrontandola con la più nota e consolidata tecnica di Zhang [2].

Si è poi passati a una qualificazione metrologica del sistema stereoscopico a tre telecamere. In particolare si è voluto analizzare l'influenza dell'intera procedura di taratura sull'accuratezza delle misure tridimensionali. Sono poi stati approfondite alcune problematiche in cui solitamente ci si imbatte durante il processo di taratura di un sistema stereoscopico. Infine si è passati a un'analisi della propagazione dell'incertezza in fase di calibrazione mediante il metodo Monte Carlo. Partendo dai risultati di questo studio è stato poi possibile effettuare un'analisi di sensitività della procedura di ricostruzione tridimensionale.

I risultati della qualificazione metrologica del sistema stereoscopico a tre telecamere sono stati utilizzati per realizzare tre sistemi di misura ottimizzati per tre differenti ambiti applicativi.

Il primo nasce dalla necessità della Galleria del Vento del Politecnico di Milano [34] di effettuare misure di spostamento in tre dimensioni senza introdurre effetti di carico sul misurando, in modo da non disturbare la vena fluida del vento. Il sistema stereoscopico realizzato deve garantire scalabilità e flessibilità in quanto può essere utilizzato per effettuare misure da posizioni molto diverse e su oggetti con svariate possibili dimensioni, a seconda dell'applicazione in cui viene utilizzato il sistema. Per questo ambito applicativo è stato realizzato un sistema di misura del tutto simile a quello utilizzato per la qualificazione metrologica. Esso è stato utilizzato in diverse campagne sperimentali, alcune delle quali sono state brevemente descritte in questo lavoro per quanto concerne la parte di misure senza contatto.

Il secondo ambito applicativo nasce da un progetto di ricerca effettuato con la società ABB Robotica Italia [32] nell'ambito del quale è stato realizzato un dimostratore robotizzato in grado di riprodurre fedelmente le traiettorie manuali eseguite da un operatore che afferra un opportuno tracciatore completo di marcatori catarifrangenti. Le problematiche legate alla ricostruzione 3D della traiettoria del tracciatore sono state risolte mediante l'uso di un sistema stereoscopico che consente di acquisire in tempo reale la posizione 3D e l'orientamento nel tempo dell'utensile. I dati raccolti poi sono trattati in modo da generare traiettorie riproducibili da parte di un robot antropomorfo.

Infine nell'ultima applicazione è stato realizzato un sistema di presa di oggetti posizionati in maniera casuale in un cassone mediante un robot antropomorfo. La problematica era già nota all'interno del VBLab del Politecnico di Milano, in quanto era già stata affrontata utilizzando come sistema di scansione un triangolatore laser. In questo lavoro si vuole invece realizzare una soluzione basata sulla tecnica stereoscopica.

Per tutte le applicazioni descritte sono state effettuate le prove necessarie a qualificare metrologicamente i sistemi realizzati; i risultati sono riportati nel seguito.

La tesi è così strutturata:

- il capitolo 2 è interamente dedicato allo stato dell'arte delle tecniche di misura 3D senza contatto, spiegando le motivazioni che hanno portato a scegliere la stereoscopia tra le altre tecniche di ricostruzione tridimensionale. Viene inoltre approfondito lo stato dell'arte delle tecniche di calibrazione di telecamere esistenti in letteratura;
- nel capitolo 3 viene descritta la metodologia di misura: partendo dal modello di telecamera utilizzato in questo lavoro, viene descritta la stereoscopia e sono riportati alcuni cenni teorici sulla geometria epipolare che sta alla base della ricostruzione tridimensionale;

- il capitolo 4 è dedicato alla calibrazione di telecamere, processo importantissimo in fase di taratura del sistema stereoscopico. Vengono descritte alcune tecniche di calibrazione e confrontate tra loro;
- il capitolo 5 riporta una descrizione del sistema stereoscopico realizzato e utilizzato per l'attività sperimentale di qualificazione metrologica. Vengono riportate una serie di analisi atte a definire l'accuratezza di tale sistema e la dipendenza di essa dalla qualità della taratura. Viene inoltre studiata la propagazione dell'incertezza nel processo di taratura mediante il metodo di Monte Carlo e viene poi riportata un'analisi di sensitività della procedura di ricostruzione 3D;
- il capitolo 6 è dedicato alle tre applicazioni innovative sviluppate in questo lavoro. Esse si basano tutte sull'utilizzo di un sistema stereoscopico per effettuare misure 3D. Ogni soluzione è stata metrologicamente verificata e per ognuna è stata studiata l'accuratezza della procedura di misura. Sono state effettuate una serie di prove per la validazione dei diversi sistemi e i risultati ottenuti sono riportati in questo capitolo;
- Infine nel capitolo 7 sono riassunte le principali conclusioni e i risultati ottenuti in questo lavoro.

2. Stato dell'arte

In letteratura sono presenti numerosissime tecniche differenti di ricostruzione tridimensionale senza contatto. In questo capitolo viene delineato un quadro generale dello stato dell'arte relativo ai sistemi di ricostruzione 3D, indicando inizialmente una possibile tassonomia ed analizzando sinteticamente ogni singola tecnologia al fine di mostrarne sia il principio di funzionamento su cui si basa sia le principali caratteristiche e limitazioni. In questo modo è possibile motivare la scelta della stereoscopia quale oggetto di studio di questo lavoro. Si ricorda che le specifiche che la tecnologia di misura scelta deve soddisfare sono: possibilità di effettuare misure dinamiche, flessibilità e scalabilità, in modo da poter realizzare un sistema di misura che potesse essere utilizzato in un'ampia varietà di ambiti applicativi, con diverse condizioni e volumi di lavoro. Una parte fondamentale nello studio delle accuratezze di un sistema stereoscopico è svolta dalla taratura del sistema ed in particolare dalla calibrazione delle singole telecamere che lo compongono. Per questo viene riportato lo stato dell'arte delle tecniche di calibrazione di telecamere.

2.1 Stato dell'arte delle tecniche di ricostruzione 3D

Le tecniche di misura 3D coprono un'ampia serie di tecnologie data la vastità dei campi applicativi nei quali sono utilizzate. Uno schema riassuntivo di tutte le tecniche di misura tridimensionale [3], [4] è espresso in figura 2.1. Tra questi, i *metodi ottici* sono i più compatibili con le applicazioni in cui verrà utilizzato il sistema di misure in esame. Queste tecniche, basate sull'analisi della scena mediante telecamere, non richiedono contatto con l'oggetto da misurare e consentono un'analisi piuttosto veloce e a bassi costi. Esse poi raggiungono accuratezze elevate e possono essere applicate a campi di misura di differenti vastità. Le limitazioni includono il fatto di poter acquisire solo la parte di superficie visibile e la sensibilità ad alcune proprietà dei materiali quali trasparenza, brillantezza e colore.

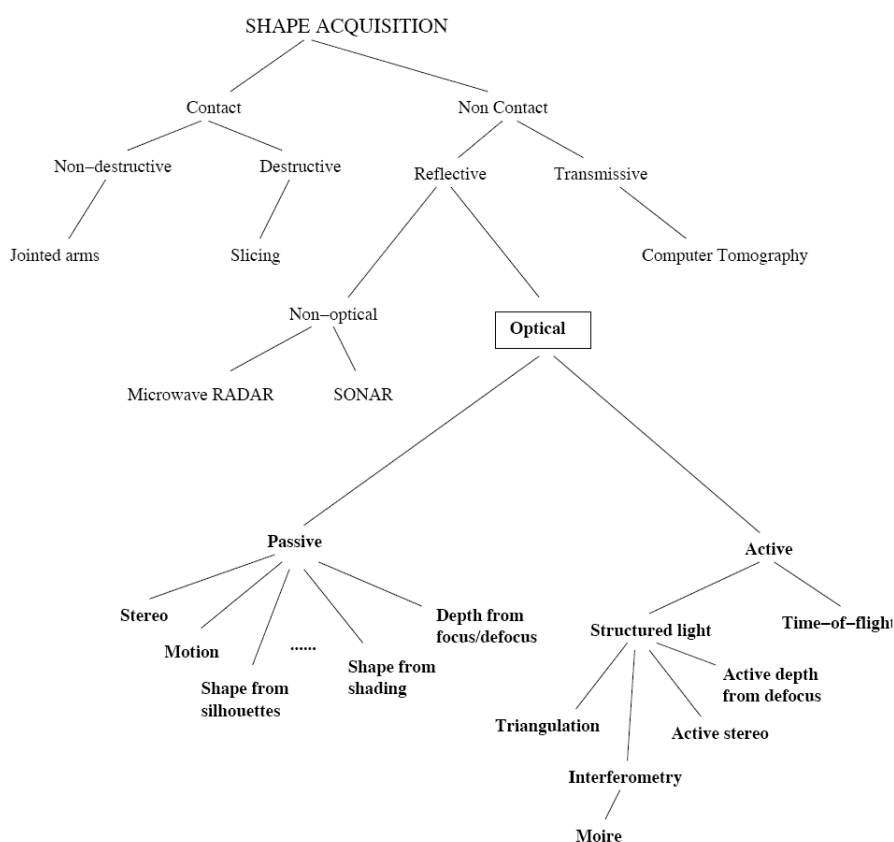


Figura 2-1: Schema riassuntivo delle varie tecniche di ricostruzione tridimensionale.

La distinzione fondamentale tra le tecniche ottiche riguarda l'impiego o meno di sorgenti d'illuminazione strutturate. In particolare, si può fare una distinzione tra i *metodi attivi*, che irradiano la scena con radiazioni elettromagnetiche opportune (pattern luminosi, luce laser, radiazioni infrarosse, ecc...) e i *metodi passivi*, che si basano sull'analisi d'immagini della scena in condizioni variabili di luce naturale. I primi hanno il vantaggio di raggiungere risoluzioni elevate, ma sono più costosi e non sempre applicabili. I secondi sono economici e pongono meno vincoli per l'applicabilità; la procedura di riconoscimento dei punti noti è generalmente meno affidabile e rischia di introdurre incertezze nel problema.

2.1.1 Tecniche passive

Ricostruzione dal chiaroscuro (Shape from shading): in cui la superficie 3D viene generata analizzando la variazione di luminosità di un'immagine 2D dell'oggetto e ricavando la normale associata ad ogni punto. Questo metodo ha notevoli svantaggi allorché si operi in condizioni non ideali (superfici lambertiane, albedo uniforme ed illuminazione parallela).

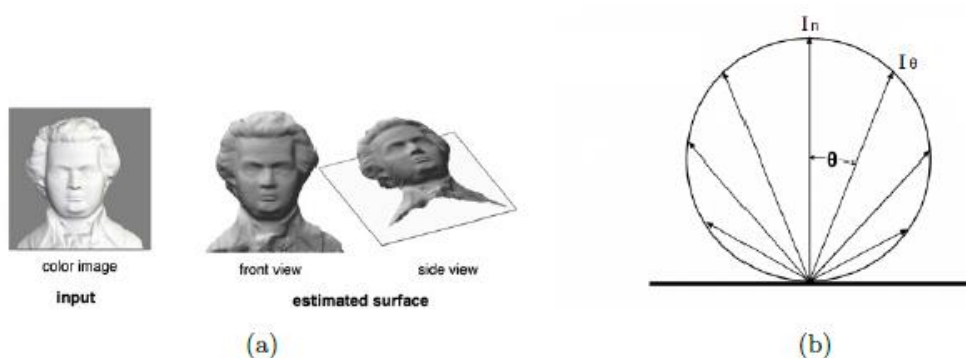


Figura 2-2: Ricostruzione 3D mediante shape from shading (a) e comportamento delle superfici lambertiane (b).

Ricostruzione dallo stereo fotometrico (Shape from photometric stereo): si basa sul medesimo principio concettuale dello shape from shading utilizzando tre differenti immagini acquisite variando, tra uno scatto ed il successivo, le condizioni di illuminazione, al fine di risolvere alcune problematiche relative alla definizione

univoca della normale propria di ogni punto dell'immagine. Il metodo presenta però limitazioni simili a quelle della tecnica precedentemente illustrata e risulta quindi inutilizzabile per la ricostruzione tridimensionale necessaria all'applicazione in esame.

Ricostruzione dalla tessitura (Shape from texture): è in grado di ricostruire la forma acquisita da una superficie su cui sono presenti pattern (texture) ripetuti in modo deterministico (ad es. un muro costituito da mattoni o una tappezzeria) oppure statistico (ad es. disposizione di sassi nella spiaggia o dei fili di erba in un prato). L'operazione è effettuata analizzando l'entità delle distorsioni e la velocità delle variazioni di opportune caratteristiche degli elementi della tessitura (chiamati texel). Poiché nelle applicazioni in esame gli oggetti da ricostruire non presentano necessariamente superfici texturizzate, questo metodo non è applicabile al presente caso di studio.

Visione stereoscopica o stereoscopia (Shape from stereopsis): questo metodo simula il sistema di visione umano ricavando le informazioni di profondità dall'analisi di immagini provenienti da due o più telecamere inquadranti la stessa scena da posizioni diverse. La ricostruzione 3D viene suddivisa in due fasi consecutive: la prima, basata sul principio di somiglianza, è finalizzata al reperimento, nelle due (o più) immagini, di punti omologhi (cioè che si riferiscono al medesimo punto fisico), la seconda, basata sui parametri di calibrazione di ogni singola camera (intrinseci), sulle loro posizioni reciproche (estrinseci) e su opportuni algoritmi, permette di effettuare la ricostruzione vera e propria determinando, per ogni n-upla di punti omologhi, il passaggio da coordinate 2D in pixel a coordinate 3D in millimetri. È chiaro come la ricerca di punti corrispondenti sia ostacolata dal verificarsi di problemi di accoppiamento legati ad occlusioni (punti visibili nella prima immagine ma nascosti nella seconda a causa di oggetti presenti nella scena) o a punti nella stessa immagine che presentano caratteristiche molto simili tra di loro e quindi non sono facili da associare a un punto della seconda immagine. Per semplificare la procedura di accoppiamento dei punti omologhi in alcuni casi è preferibile utilizzare tecniche stereoscopiche attive o semiattive, come descritto in seguito. I vantaggi di tale tecnica sono dati dalla sua flessibilità e

scalabilità, essendo infatti utilizzabile in varie applicazioni con anche campi di misura di dimensioni molto diverse. Altro vantaggio importante è la velocità di acquisizione che consente di effettuare misure 3D in tempo reale.

Ricostruzione dal movimento (Shape from motion): questo metodo si basa sullo studio di immagini catturate con una singola telecamera (di cui sono noti i parametri intrinseci) in moto relativo rispetto all'oggetto da ricostruire. Si ottiene quindi una serie di acquisizioni, riprese da punti di vista diversi, le quali sono paragonabili ad immagini scattate simultaneamente con un sistema di visione stereoscopico; i due metodi infatti sono concettualmente simili. È chiaro come l'ottenimento di buoni risultati in termini di ricostruzione 3D sia subordinato al fatto che non vi debba essere alcuno spostamento relativo tra le componenti dell'oggetto. Esistono due diversi approcci al problema: i metodi discreti ed i metodi differenziali. I primi fondano il loro funzionamento sul reperimento dei punti omologhi (analogamente alla stereoscopia), mentre i secondi analizzano le velocità dei punti dell'immagine. Questa tecnica è applicabile solamente se il soggetto da misurare si presenta come un unico corpo rigido.

Fochettatura (Depth from focus): la posizione di un punto nel sistema di riferimento della camera viene determinata dall'analisi di diverse immagini, scattate in sequenza, in cui l'oggetto in esame è acquisito con un livello di messa a fuoco via via migliore. Questa tecnica è piuttosto costosa in quanto richiede ottiche dotate di lenti di alta fuoco ottimale (condizione non sempre banale da stabilire). Le suddette considerazioni determinano l'incompatibilità del presente metodo con le applicazioni analizzate.

Sfocamento (Depth from defocus): la distanza di ciascun punto viene determinata basando l'analisi di varie immagini, aventi differenti livelli di messa a fuoco, sulla relazione che lega profondità, sfocamento e caratteristiche dell'ottica (focale, ecc...). La necessità di dover acquisire numerose immagini, di calibrare la relazione sfocamento-profondità e di dover misurare in modo accurato lo sfocamento, in aggiunta all'utilizzo di ottiche professionali (al fine di ottenere risultati accurati), caratterizzate da elevate prestazioni, rende questo metodo sia lento che costoso.

Ricostruzione dalle sagome (Shape from silhouettes): la ricostruzione 3D generata da questo metodo si fonda sull'acquisizione di immagini binarizzate (silhouette), ciascuna delle quali è composta da pixel o bianchi o neri corrispondenti a raggi ottici rispettivamente intersecanti o non intersecanti la superficie in esame. A ciascuna immagine può essere associato un volume, a forma di cono generalizzato, all'interno del quale è contenuto l'oggetto dell'analisi. Diverse viste definiscono altrettanti coni dalla cui intersezione deriva la ricostruzione 3D. I principali vantaggi di questo metodo consistono nel non dovere effettuare la complessa operazione di ricerca di punti omologhi, e nella possibilità di ottenere direttamente un modello tridimensionale dell'oggetto analizzato. Tuttavia lo shape from silhouettes è un metodo concettualmente non banale e, nonostante l'esistenza di diversi algoritmi ottimizzanti, computazionalmente oneroso. In aggiunta a ciò si evidenzia come esso non sia adatto per effettuare misure dinamiche.

Ricostruzione da foto coerenza (Shape from photo consistency): è un secondo metodo di ricostruzione volumetrica il quale concettualmente si presenta come un'estensione del precedente. Esso infatti si basa sull'utilizzo di immagini a colori (o in scala di grigi) le quali aggiungono vincoli al problema migliorando il processo di ricostruzione. Si ha quindi una tecnica da un lato più performante della precedente ma dall'altro computazionalmente più onerosa.

2.1.2 Tecniche attive

Le tecniche attive si dividono in *metodi a tempo di volo* e *tecniche a luce strutturata*. I primi sfruttano il medesimo principio di funzionamento del RADAR, ma utilizzano come sorgente di onde elettromagnetiche un laser: uno scanner rileva punto a punto il ritardo fra l'emissione e la ricezione del raggio riflesso. La ricostruzione è possibile conoscendo la geometria del sistema di visione e la velocità di propagazione delle onde luminose. Quasi tutte le tecniche a tempo di volo hanno il difetto di essere lente, in quanto richiedono una scansione punto a punto. Esse non sono dunque di interesse per le applicazioni in esame. L'unica tecnica che permetterebbe di acquisire nuvole di punti 3D in tempo reale è quella

delle telecamere a tempo di volo. Esse non sono altro che dei sensori a tempo di volo matriciali. Questi sistemi però hanno un campo di misura ristretto e spesso non hanno accuratezze sufficienti per le applicazioni in esame. I sistemi a luce strutturata più utilizzati sono: lo *sfocamento attivo*, la *triangolazione*, l'*interferometria* e la *stereoscopia attiva*.

Con lo *sfocamento attivo* (*Depth from defocus*), utilizzando un'illuminazione strutturata, da varie immagini caratterizzate da diversi parametri di fuoco si determina la profondità di tutti i punti della scena, grazie alle relazioni fra profondità, sfocamento e parametri dell'ottica. Le problematiche più evidenti sono la misurazione dello sfocamento e la calibrazione della relazione profondità – sfocamento, oltre alla necessità di acquisire diverse immagini. E' evidente come questo metodo risulti lento e richieda ottiche costose. Di conseguenza lo sfocamento non è utilizzabile nel caso in esame.

La *triangolazione attiva* consente di utilizzare una sola telecamera, proiettando una lama laser sul misurando, resa però solidale al proiettore. In un certo senso, quest'ultimo e la telecamera vedono la scena da posizioni diverse ed è quindi possibile determinare la terza dimensione intersecando il raggio ottico di un punto dell'immagine con il corrispondente piano di luce prodotto dal proiettore. Per un corretto funzionamento della triangolazione è necessario realizzare un modello geometrico del proiettore e la calibrazione di quest'ultimo. Il grande vantaggio di questo metodo è quello di non richiedere la valutazione di alcuna corrispondenza, l'aspetto negativo è la richiesta, di un elevato numero di immagini e la possibilità di eseguire solamente misure statiche. Le applicazioni in cui viene utilizzato il sistema in esame richiedono la possibilità di effettuare misure dinamiche.

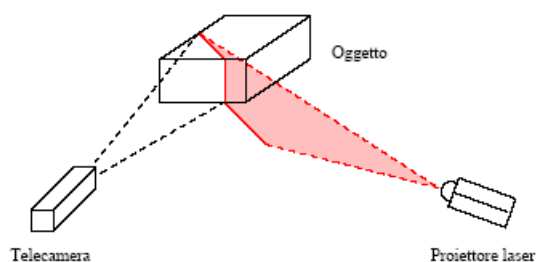


Figura 2-3: Applicazione della triangolazione attiva.

Con il termine *Interferometria* viene essenzialmente considerato il metodo delle *frange di interferenza di Moiré*. La teoria che è alla sua base non è banale e in realtà esistono diverse tecniche di acquisizione che si rifanno ad essa. L'idea di fondo è quella di proiettare una griglia su un oggetto e di acquisirne l'immagine attraverso una seconda griglia, analoga ma leggermente ruotata: le due griglie interferiscono e si creano delle frange di interferenza, dette appunto di moiré (*moiré fringe patterns*). Esse risultano essere fasce alternate di luce ed ombra: un'opportuna analisi di questi pattern permette di risalire alle variazioni di profondità nella scena e quindi alla ricostruzione 3D. Esistono diverse tecniche per generare le griglie, fra le altre esse possono essere proiettate tramite diapositiva oppure essere create tramite software. Il metodo moiré è veloce, accurato e non costoso, tuttavia presenta alcuni inconvenienti: innanzitutto è computazionalmente oneroso, inoltre non è immediato ricostruire la scena laddove esistano nella scena dei salti bruschi di profondità e le superfici troppo angolate talvolta non sono misurabili, poiché presentano una densità delle bande troppo elevata. Dato che la tecnica di misura in esame deve poter misurare diversi oggetti, indipendentemente dalla loro forma, non è consigliabile utilizzare un metodo che abbia queste limitazioni. Per questo e per la complessità computazionale del metodo, esso non è stato considerato in questo lavoro.



Figura 2-4: Applicazione del metodo di Moirè.

La stereoscopia attiva (*Active stereo*) consente, utilizzando almeno due telecamere, di ricostruire la terza dimensione analizzando le corrispondenze tra i punti nelle immagini acquisite dai diversi dispositivi. La luce strutturata facilita la risoluzione del problema dell'identificazione delle corrispondenze. Sono utilizzate diverse tipologie di luce strutturata (infrarossi, laser, etc...). Spesso (come anche in alcune applicazioni oggetto di questo lavoro) per facilitare l'individuazione di punti omologhi viene utilizzata una tecnica a marker catarifrangenti soggetti a illuminatori all'infrarosso. Questi marker vengono posizionati sull'oggetto da misurare e grazie a un filtro sull'ottica della telecamera è possibile acquisire solamente la radiazioni infrarosso riflesse dai marker, riducendo notevolmente il campo del visibile. Così la determinazione dei punti noti, e dunque l'identificazione delle corrispondenze, risulta più facile. Questa tecnica permette di effettuare misure in maniera dinamica, specifica indispensabile per le applicazioni di interesse in questo lavoro. Uno svantaggio di questo metodo è la possibilità di misurare soltanto un numero limitato di punti, ovvero soltanto i punti noti di cui si conosce le corrispondenze. Questa tecnica, che è descritta più in dettaglio nei capitoli successivi, è spesso considerata una tecnica semiattiva. Infatti la scena viene irraggiata da un fascio di luce che viene riflessa dai marker e quindi i marker non sono veramente attivi. Si utilizzano, però, illuminatori infrarossi e quindi non luce naturale. Il confine tra tecnica attiva e semiattiva è molto sottile e quindi esse vengono solitamente entrambe classificate come tecniche attive.

2.2 Stato dell'arte delle tecniche di calibrazione di telecamere

L'accuratezza di un sistema stereoscopico dipende per buona parte dal suo processo di taratura. Il processo di taratura di un sistema stereoscopico è composto dalla calibrazione delle singole telecamere e da una procedura di registrazione delle stesse (vale a dire individuazione della posa relativa). La parte più critica è la calibrazione delle singole telecamere, che verrà dunque studiata nel dettaglio in questo lavoro di tesi. Per questo si riporta qui lo stato dell'arte delle tecniche di calibrazione di telecamere esistenti in letteratura.

Con il termine *calibrazione* di telecamere s'intende il processo di determinazione dei parametri che definiscono il modello della telecamera, descritto in dettaglio nel capitolo 3. La calibrazione è un processo necessario per ricavare informazioni metriche dall'immagine e potere così utilizzare una telecamera come strumento di misura. Ad oggi non esiste una norma che impone una particolare metodologia di calibrazione. In base all'applicazione in cui le telecamere sono utilizzate, si può ricorrere a una determinata tecnica piuttosto che un'altra.

Ad oggi le tecniche di calibrazione presenti si possono suddividere in due principali categorie: i metodi basati sull'uso di un calibratore e le tecniche di auto calibrazione (*self-calibration*). I primi consistono nell'utilizzo di un calibratore, cioè un oggetto di cui si conosce con elevata accuratezza la disposizione nello spazio di alcuni punti noti, come riferimento nello spazio. Attraverso le corrispondenze che vengono individuate tra i punti nella scena e la loro proiezione nell'immagine, vengono determinati i parametri del modello della telecamera. Questi metodi consentono di raggiungere ottime prestazioni in termini di accuratezza [5], [6], [7]. I più famosi metodi tradizionali sono: il "*Direct Line Transformation*" (DLT) method, l'*ottimizzazione non lineare*, il *Tsai two-step method* e il *metodo di Zhang*.

Il DLT fu proposto per la prima volta da Abdel-Aziz e Karara nel 1971 [8]. Esso si basa sul calcolo dei parametri della matrice di proiezione, vale a dire la

matrice che descrive la corrispondenza tra i punti nella scena e nell'immagine. Il grande vantaggio di questo metodo consiste nel fatto che i parametri intrinseci ed estrinseci sono calcolati risolvendo equazioni lineari senza l'utilizzo di algoritmi iterativi. Nonostante la sua velocità computazionale e la semplicità derivante dall'utilizzo di un numero ristretto di parametri, esso non è ad oggi una delle tecniche più utilizzate, poiché non prende in considerazione le distorsioni non lineari e quindi non riesce a raggiungere accuratezze elevate. Esso è stato poi perfezionato per includere i fattori non lineari, aumentando la complessità del problema. Queste tecniche di calibrazione non lineare hanno un'elevata accuratezza, ma sono computazionalmente lente e complesse; inoltre richiedono nella maggior parte dei casi una buona stima dei parametri iniziali.

Nel 1987 Tsai [5] sviluppò un metodo chiamato *two-step*, che consiste in due passaggi principali: prima di tutto viene risolto il problema lineare basandosi sui vincoli che legano i parametri della matrice di proiezione, poi si procede con un'ottimizzazione non lineare, inizializzata con i risultati trovati al primo punto, che tiene conto delle distorsioni. Questo metodo può raggiungere elevate accuratezze pur operando in uno spazio di parametri meno esteso rispetto al primo [6]. Il metodo *two-step* proposto da Tsai considera solamente le distorsioni al secondo ordine, cioè tiene conto solo delle distorsioni radiali senza considerare quelle tangenziali. In generale è possibile raggiungere buoni risultati di calibrazione anche trascurando le distorsioni tangenziali, ma per una maggiore accuratezza è conveniente considerare il modello sviluppato da Brown, che considera entrambe le tipologie di non linearità [9]. Il modello non lineare di Brown è stato considerato durante le prove in questa tesi ed è descritto in seguito nel paragrafo 3.3.

Nel 1998 Zhengyou Zhang propose un nuovo metodo [2], che richiede solamente l'osservazione di un *pattern planare* (figura 2.5) in diverse pose casuali nello spazio (almeno tre). Si può decidere di muovere il calibratore piuttosto che la telecamera, senza alterare i risultati di calibrazione e soprattutto senza il bisogno di misurare la posizione nello spazio del pattern rispetto alla telecamera a priori. Accoppiando i punti del calibratore nelle immagini con un modello di esso, è possibile calcolare la matrice di omografia e da questa ricavare i parametri della

telecamera. L'algoritmo di Zhang per la stima dei parametri lineari può essere raffinato implementando un problema di minimizzazione non lineare, che quindi tiene conto delle distorsioni della lente. Il vantaggio di questa tecnica consiste nel *setup* semplice ed economico avendo comunque risultati molto accurati in fase di calibrazione. Questa tecnica è molto diffusa e ormai validata ed è stata in questo lavoro presa come riferimento per il confronto con altre tecniche più innovative. Tale tecnica verrà quindi approfondita più nel dettaglio nel paragrafo 4.1.

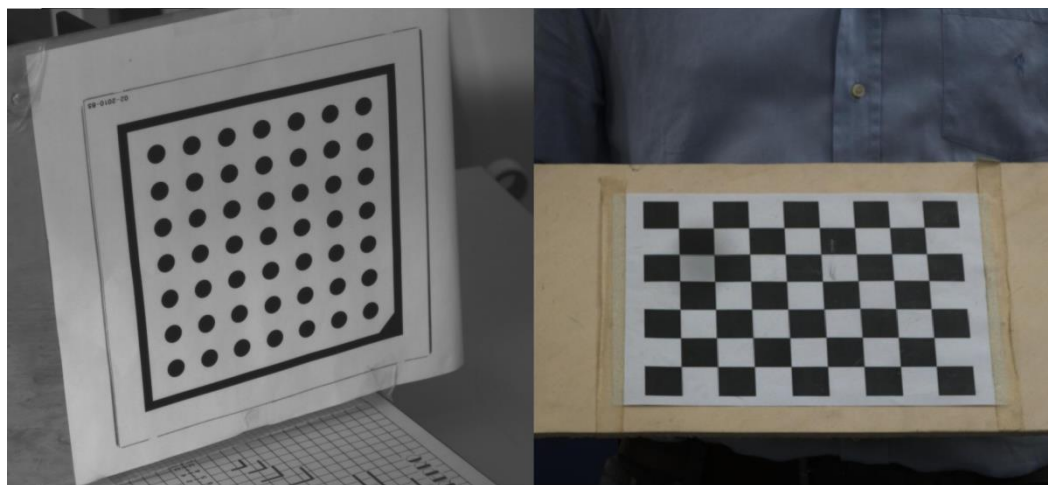


Figura 2-5: Esempi di pattern planari utilizzati durante calibrazioni col metodo di Zhang.

Una tecnica molto simile è stata sviluppata nel 2000 da Janne Heikkilä [1]. Questa tecnica utilizza un calibratore tridimensionale con alcuni cerchi (*blob*) in posizione nota e di raggio noto. L'innovazione di tale tecnica sta in una semplificazione del modello non lineare che permette di invertire il modello delle distorsioni senza ricorrere a metodi ricorsivi. Heikkilä inoltre propone un metodo innovativo per identificare il centro dei blob del calibratore con maggiore accuratezza. Questa tecnica è stata implementata nelle librerie HALCON di MVTec [33] in uso presso il VBLab del Politecnico di Milano ed verrà dunque analizzata più nel dettaglio nel paragrafo 4.2.

In alcune applicazioni non è possibile calibrare la telecamera utilizzando un calibratore e così sono state sviluppate negli ultimi anni numerose tecniche di *self-calibration*. Esse infatti non richiedono l'utilizzo di un calibratore particolare, ma, attraverso la movimentazione della telecamera in una scena statica si riescono a

ricavare i vincoli che legano i parametri di calibrazione. Questo metodo è adatto a quelle applicazioni in cui i parametri della telecamera non possono essere mantenuti fissi ed è necessario dunque ricalibrare spesso e velocemente il sistema. A differenza dei metodi tradizionali questa tecnica è molto flessibile, tuttavia è anche più complessa e non permette di raggiungere accuratezze altrettanto elevate. Tra queste tecniche riscontrano grande interesse quelle basate sulla *visione attiva*, vale a dire le tecniche in cui i parametri di calibrazione vengono determinati attraverso il controllo attivo dei movimenti della telecamera. Esse hanno il vantaggio di basarsi su algoritmi semplici ma lo svantaggio di richiedere un sistema talvolta complesso di controllo.

Il costante sviluppo di alcune metodologie per l'ottimizzazione non lineare, quali reti neurali e algoritmi genetici (GA), ha permesso lo sviluppo di alcuni metodi di calibrazione in cui il problema non lineare viene risolto attraverso l'utilizzo di queste tecniche. Gli algoritmi ad apprendimento sono però ancora troppo lenti e spesso convergono a un minimo locale portando errori considerevoli in fase di calibrazione. Presso il VBLab sono state implementate e validate alcune tecniche basate sugli algoritmi genetici [10], che però non hanno portato a risultati soddisfacenti in termini di robustezza e soprattutto velocità. Un'innovativa tecnica che permette di combinare la capacità di fitting non lineare delle reti neurali con la buona capacità di ricerca globale degli algoritmi genetici di calibrazione è descritta in [11].

3. Metodologia di misura

Nel capitolo precedente sono state descritte le principali tecniche di ricostruzione tridimensionale senza contatto presenti in letteratura e per ognuna sono stati enunciati vantaggi e svantaggi. Al fine di realizzare uno strumento di misura che potesse essere consono a diverse applicazioni tra le tecniche elencate si è deciso di approfondire la stereoscopia. Essa infatti presenta vantaggi in termini di flessibilità, scalabilità, economicità e consente di effettuare misure dinamiche. In questo capitolo viene descritta più in dettaglio la stereoscopia. Prima di tutto si vuole descrivere tale tecnica in generale e approfondire le motivazioni che hanno portato alla scelta di essa come metodologia di misura. Essendo la stereoscopia una tecnica ottica, e in quanto tale basata sull'utilizzo di telecamere, si vuole inoltre descrivere nei particolari il modello di telecamera utilizzato in questo lavoro. Questo risulterà molto utile per comprendere le procedure di calibrazione descritte al capitolo successivo. In seguito viene qui introdotta la geometria che sta alla base della stereoscopia, chiamata geometria epipolare e vengono approfondite alcune problematiche legate alle misure mediante stereoscopia.

3.1 Stereoscopia

La *stereoscopia*, o *stereopsi*, è la capacità percettiva che consente di unire le immagini provenienti dai due punti di vista (ad esempio gli occhi nel caso di esseri umani), che grazie alla loro disposizione strutturale, presentano uno spostamento laterale. Questa disparità è sfruttata dal cervello per trarre informazioni sulla profondità e sulla posizione spaziale dell'oggetto mirato. Di conseguenza la stereopsi permette di generare la visione tridimensionale. Tale metodo naturale è stato studiato e applicato alla visione artificiale dando vita alla stereoscopia (o stereopsi computazionale), ovvero l'insieme di tecniche per ricavare informazioni sulla struttura tridimensionale della scena da almeno una coppia di immagini acquisite in sincronia da posizioni spaziali diverse e note.

Questo problema può essere risolto scomponendolo in due sottoproblemi: l'individuazione nelle immagini delle corrispondenze, cioè di punti immagine riferiti allo stesso punto fisico, e la ricostruzione tridimensionale degli stessi date le posizioni relative tra le due telecamere (processo chiamato triangolazione).

I punti immagine corrispondenti sono chiamati punti omologhi. La loro individuazione può risultare però difficile per la presenza di errori dovuti a occlusioni, distorsioni di luminosità e distorsioni proiettive. Le occlusioni sono i punti di cui non si ha corrispondenza in una delle immagini tipicamente perché essi non sono visibili da entrambe le posizioni per effetto di coperture. Le distorsioni di luminosità si hanno quando lo stesso punto, nelle due immagini, presenta un valore di luminosità differente. Questo accade tipicamente a causa di ombre o riflessi di luce sulla superficie osservata. La distorsione proiettiva invece si manifesta quando un oggetto ripreso da due punti di vista differenti si proietta sulle due immagini in maniera differente.

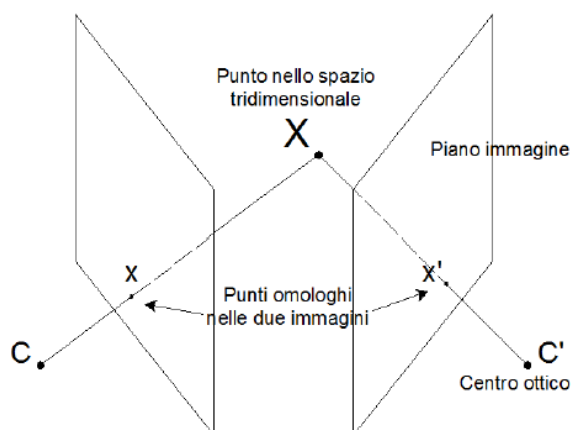


Figura 3-1: I due punti x e x' sono proiezione dello stesso punto della scena e sono dunque detti punti omologhi (o coniugati).

Per agevolare l'individuazione dei punti omologhi si impongono vincoli che semplificano la ricerca. Il più importante è il vincolo epolare che permette di affermare che il corrispondente di un punto in un'immagine può trovarsi solo su una retta (retta epolare) nell'altra immagine. Questo viene spiegato più in dettaglio nel seguito. Una volta noti gli accoppiamenti tra i punti è possibile ottenere la posizione del punto tridimensionale che li ha generati; questo processo viene chiamato triangolazione.

Il grosso svantaggio di questo metodo di ricostruzione tridimensionale è la necessità di avere un sistema di misura calibrato e la possibilità di essere applicato solo a oggetti che presentino caratteristiche facilmente distinguibili in maniera automatica come blob, forme geometriche particolari, fori, spigoli, e altre discontinuità. Queste caratteristiche vengono dette features e possono essere intrinseche dell'oggetto da misurare piuttosto che imposte mediante un sistema d'illuminazione. In questo caso la tecnica utilizzata viene chiamata stereoscopia attiva poiché il misurando viene irraggiato con una fonte luminosa. La luce strutturata facilita la risoluzione del problema dell'identificazione delle corrispondenze. Sono utilizzate diverse tipologie di luce strutturata (infrarossi, laser, etc...). Un'altra tecnica utilizzata in alcune applicazioni di questo lavoro per facilitare l'individuazione di punti omologhi è la stereoscopia semiattiva a marker catarifrangenti soggetti a illuminatori all'infrarosso. Questi marker vengono

posizionati sull'oggetto nelle posizioni in cui vi sono dei punti di interesse per effettuare le misure. La scena viene irradiata con illuminazione all'infrarosso e un filtro interferenziale centrato sulla lunghezza d'onda dell'illuminatore viene posto sull'ottica delle telecamere in modo da tagliare le lunghezze d'onda indesiderate e così da escludere il campo del visibile mettendo in evidenza solamente i marker posti sul misurando. L'immagine acquisita dalle telecamere risulta essere completamente scura con solamente alcuni cerchi (blob) bianchi dove nella scena è presente il marker. La ricerca dei punti omologhi si riduce dunque all'associazione delle corrispondenze tra i centri dei blob e risulta dunque più semplice e robusta. Inoltre la luce naturale dell'ambiente non influisce sulle misure, che dunque possono essere fatte anche in condizioni di illuminazione scarsa o sfavorevole. Come spiegato sopra questa tecnica è spesso considerata una tecnica semiattiva. Infatti la scena viene irraggiata da un fascio di luce che viene riflessa dai marker, che però non emettono loro stessi luce e quindi non sono considerati attivi. Si utilizzano, però, illuminatori infrarossi e quindi non luce naturale. Il confine tra tecnica attiva e semiattiva è molto sottile e quindi esse vengono spesso entrambe classificate come tecniche attive.

I vantaggi della stereoscopia rispetto alle altre tecniche sono i seguenti:

- possibilità di effettuare misure dinamiche (con velocità di acquisizione che dipendono dall'hardware)
- flessibilità delle misure: è possibile effettuare misure di oggetti molto diversi tra di loro, l'unica limitazione è data dalla presenza di feature per l'identificazione di punti omologhi.
- scalabilità: modificando la posizione relativa e l'ottica delle telecamere è possibile avere campi di misura di dimensioni molto diverse fra di loro.

Essendo lo scopo di questo lavoro studiare una tecnica di misura che possa essere utilizzato in diversi ambiti applicativi, i vantaggi sopra descritti sono stati ritenuti fondamentali e per questo si è scelto di studiare in dettaglio la stereoscopia tra tutte le tecniche di misura tridimensionale non a contatto.

3.2 Il modello lineare della telecamera

La stereoscopia è una tecnica ottica e in quanto tale si basa sull'utilizzo di telecamere. La telecamera è uno strumento in grado di acquisire immagini dal mondo circostante, immagazzinando le informazioni ottenute dai fotoni riflessi dagli oggetti che compongono la scena, in un determinato range di frequenze. Questo avviene tramite una conversione degli stessi fotoni in un'unità di misura elaborabile digitalmente, solitamente elettronici. Tale strumento effettua quindi una mappatura tra il mondo tridimensionale e un'immagine bidimensionale. Esistono diversi modelli geometrici [8] che descrivono tale processo e ognuno di questi caratterizza un diverso tipo di telecamera. Questi modelli possono essere distinti in base ai parametri che li compongono; ad esempio è possibile trovare modelli che descrivono telecamere con centro in posizione finita, e telecamere con centro "all'infinito", generalizzazione naturale di una proiezione di linee parallele ("*telecamera affine*"). Quest'ultimo modello viene in genere utilizzato quando il sistema non è riconducibile a un sistema a proiezione centrale, ovvero per ottiche di lunghezza focale molto elevata e non di interesse per questo lavoro. La differenza tra questi due modelli è ben visibile in figura 3.2: in (a) i punti vengono proiettati parallelamente tra loro nell'immagine, in (b) invece esiste un centro di proiezione nel quale i punti sono proiettati.

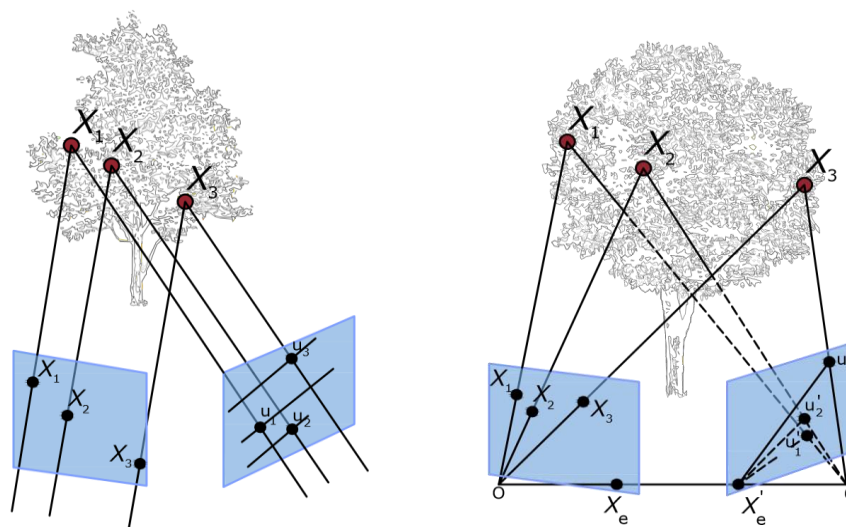


Figura 3-2: Differenza tra modello di telecamera affine (sinistra) e telecamera proiettiva (destra)

In questo lavoro viene utilizzato un modello di telecamera proiettiva, poiché le lunghezze focali qui utilizzate non richiedono l'utilizzo del modello di telecamera affine. In particolare verrà sempre fatto riferimento al modello più comunemente utilizzato, ovvero il *pinhole camera* [8], [12] (figura 3.3). Tale modello descrive la telecamera come un sistema a proiezione centrale, ovvero come una telecamera senza lenti con una singola apertura molto piccola. Si tratta di una scatola chiusa con un piccolo foro attraverso il quale passa la luce proveniente dalla scena e viene quindi proiettata un'immagine capovolta sull'altro lato della scatola. Tale rappresentazione sfrutta il concetto di *foro stenopeico* che descrive in maniera accurata l'effettiva formazione delle immagini attraverso una lente. Il foro stenopeico è usato per rappresentare quello che viene chiamato obiettivo della telecamera. Le telecamere che utilizzano tale struttura producono immagini tanto più nitide quanto più il diametro del foro è piccolo. Minori dimensioni del foro comportano però la diminuzione della luce incidente sul sensore posto dall'altro lato, aumentando i tempi di esposizione necessari. In teoria la dimensione del diaframma dovrebbe essere 1/100 della distanza tra essa e il sensore.

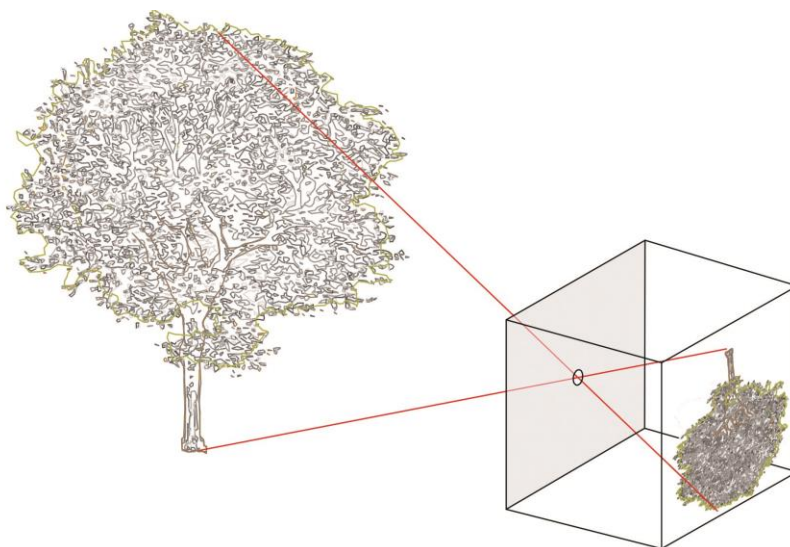


Figura 3-3: Principio base del modello pinhole camera.

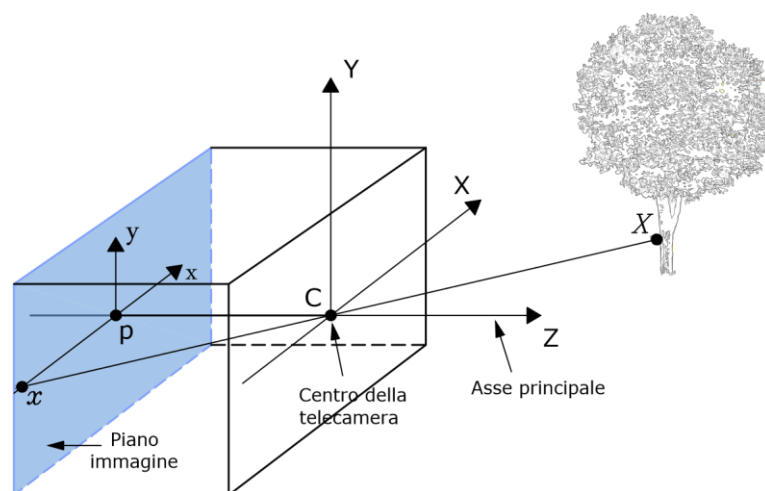


Figura 3-4: Geometria del modello pinhole camera.

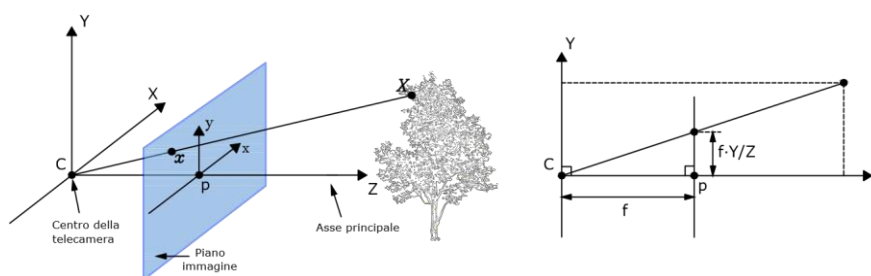


Figura 3-5: Geometri del modello pinhole camera frontale, vale a dire con centro di proiezione posto dietro al sensore.

La geometria che descrive tale modello è mostrata in figura 3.4. Il centro di proiezione del sistema è chiamato *centro della telecamera* o *centro ottico*. Una visione alternativa consiste nel mettere il centro della telecamera dietro il piano immagine, considerando quindi i punti immagine come se fossero già stati invertiti; così facendo si ottiene una formulazione geometrica più facile e comprensibile (figura 3.5). Tale modello viene chiamato *pinhole frontale* ed è quello che sarà preso in considerazione durante tutto il lavoro svolto. La linea dal centro della telecamera perpendicolare al piano immagine viene chiamata *asse principale* o *raggio principale* della telecamera e il punto in cui interseca il piano immagine viene chiamato punto principale. Infine, il piano su cui giace il centro telecamera, parallelo al piano dell'immagine è chiamato *piano principale*. E' intuitivo come tale modello descriva il modo in cui il punto mondo X viene mappato in un punto sul piano immagine x . Esprimendo tali punti nelle tre coordinate spaziali che li compongono è possibile indicarli con dei vettori colonna: $(X, Y, Z)^T$ esprimeranno le coordinate nel mondo, mentre $(f \frac{X}{Z}, f \frac{Y}{Z})^T$ esprimeranno i punti nell'immagine. Ignorando l'ultima coordinata immagine si ottiene una mappatura da uno spazio euclideo in R^3 in uno spazio euclideo in R^2 :

$$(X, Y, Z)^T \rightarrow \left(f \frac{X}{Z}, f \frac{Y}{Z}\right)^T \quad (3.1)$$

Se il mondo e i punti immagine sono espressi mediante vettori omogenei, la proiezione centrale usata dal modello può essere facilmente espressa come mappatura lineare tra le loro coordinate omogenee; in particolare la 3.1 può essere

espressa come:

$$\begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & & 0 \\ & f & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (3.2)$$

La matrice 3×4 che si ottiene è chiamata *matrice di proiezione* o matrice della telecamera (indicata con la lettera P). Usando una notazione più compatta:

$$x = PX \quad (3.3)$$

Per esigenze di chiarezza è opportuno definire i sistemi di riferimento che vengono comunemente utilizzati nel modello della telecamera. Il primo è il SDR globale dell'intero sistema, che è detto SDR mondo ed è nelle nostre applicazioni spesso solidale con la prima telecamera o, in calibrazione, coincidente con la posizione del calibratore. Un successivo SDR è quello che individua il centro geometrico della telecamera, che viene utilizzato per individuare la posizione delle telecamere rispetto al SDR mondo. Come si osserverà in seguito, il centro geometrico del sensore non sempre coincide con il punto principale e la posizione di tale punto non è comunemente utilizzata per indicare le coordinate dei pixel; infatti, per questioni di comodità, essi sono espressi in riferimento all'angolo in alto a sinistra del sensore, in modo da avere sempre coordinate positive. Questo si traduce in un ulteriore SDR di telecamera che permette la definizione dei punti nell'immagine direttamente in coordinate pixel.

Il modello fin qui descritto però è molto semplice e non abbastanza preciso per essere utilizzato in molte applicazioni che richiedono un modello che tenga in considerazione le non idealità della telecamera. È necessario dunque complicare il modello per far sì che esso rispecchi più fedelmente la realtà. In primo luogo si deve, infatti, notare che la posizione dell'origine delle coordinate non sempre coincide con il punto principale; in tale caso l'espressione diventa:

$$\begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (3.4)$$

Dove $(p_x; p_y)$ sono le coordinate del punto principale.

Introducendo:

$$K = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} \quad (3.5)$$

possiamo scrivere in forma concisa:

$$x = K [I \mid 0] X_{cam} \quad (3.6)$$

La matrice K è chiamata *matrice di calibrazione* della telecamera. X_{cam} è scritto in questa forma solo per enfatizzare che la telecamera viene posta nell'origine degli assi cartesiani con l'asse principale diretto lungo l'asse Z . Nel caso in cui il SDR mondo coincida con il SDR di telecamera allora la matrice di calibrazione coincide con la matrice di proiezione. In generale invece i punti nello spazio sono espressi in un diverso sistema di riferimento Euclideo, ovvero il SDR mondo. I due sistemi di coordinate sono collegati tra loro da una rototraslazione. Esprimendo quindi con R la matrice di rotazione tra i due SDR e con \tilde{C} le coordinate del centro della telecamera in coordinate mondo si ottiene:

$$x = KR [I \mid -\tilde{C}] X \quad (3.7)$$

Dove X è ora espresso in coordinate mondo. Questa è la generica mappatura descritta dal modello pinhole camera. La matrice di calibrazione ha dunque 9 GDL: tre dalla matrice K (gli elementi f ; p_x ; p_y), tre dalla matrice R (esprimibile come rotazione su 3 assi), e tre da \tilde{C} (tre coordinate spaziali del centro della telecamera). La matrice K contiene le informazioni dei parametri che dipendono unicamente dalla telecamera e che quindi vengono chiamati *intrinseci*, mentre R e \tilde{C} dipendono dalla posizione e orientamento della telecamera rispetto al mondo esterno, per tale ragione i parametri che li compongono sono chiamati *estrinseci*.

Spesso risulta conveniente non esprimere direttamente il centro della telecamera ma rappresentarlo in termini di trasformazione tra coordinate mondo e immagine. Si può quindi indicare:

$$P = K [R \mid t] \quad \text{con} \quad t = -R\tilde{C} \quad (3.8)$$

In questa tesi sono state utilizzate soltanto *telecamere digitali*; il sensore di tali telecamere è una matrice di semiconduttori chiamati *pixel*. Ognuno di essi compie la conversione da fotoni a elettroni della luce incidente nel suo settore. Un'immagine è formata quindi dalle informazioni coordinate di tutti i pixel che compongono il sensore. Tali componenti sono dotati di diverse sezioni: la prima effettua la conversione fotoelettrica, una seconda, detta di immagazzinamento e separata da quella di rilevamento, immagazzina temporaneamente le cariche acquisite da una precedente conversione, ed infine la terza sezione solitamente formata da un circuito di controllo integrato che gestisce le immagini, permettendone un immagazzinamento ulteriore in una memoria flash e gestisce l'invio in forma normale o compressa delle stesse ad altri dispositivi. La capacità di avere le immagini già in formato digitale è un forte contributo alla possibilità di eseguire successivamente operazioni automatiche su di esse.

I pixel delle telecamere non sono sempre perfettamente quadrati. Siccome normalmente le immagini vengono misurate in pixel, ciò rende diverso il fattore di scala orizzontale e verticale del sensore. Tale costruzione si vede in maniera molto accentuata in figura 3.6: dimensioni non quadrate dei pixel si traducono in un numero diverso di pixel per unità di lunghezza che comporta una scalatura diversa nella proiezione del punto lungo le due direzioni.

Per tener conto di questo fenomeno la forma generale della matrice di calibrazione delle telecamere digitali diventa quindi :

$$K = \begin{bmatrix} f_x & x_0 \\ & f_y & y_0 \\ & & & 1 \end{bmatrix} \quad (3.9)$$

con $f_x = f \cdot m_x$ e $f_y = f \cdot m_y$, dove m_x e m_y rappresentano il numero di pixel per unità di distanza rispettivamente lungo x e y.

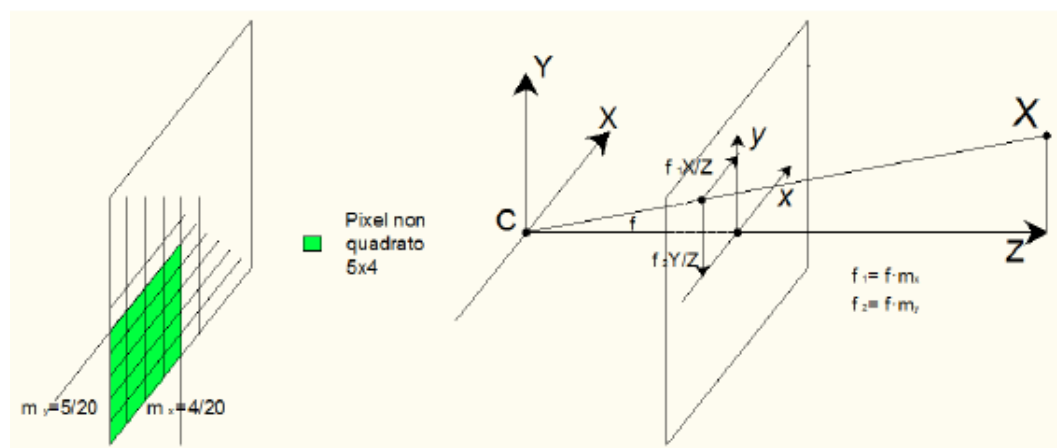


Figura 3-6: Matrice dei pixel di una telecamera digitale con distanze focali diverse lungo x e y e effetto di questo fattore sulle coordinate immagine.

È necessario introdurre un ulteriore parametro per descrivere interamente il modello di una telecamera digitale, ovvero lo *skew*. Esso descrive la possibilità che i pixel abbiano lati non perpendicolari tra loro ed è espresso da:

$$s = \cos(\alpha) * f \quad (3.10)$$

dove α è l'angolo formato dai lati del pixel. Tale parametro è generalmente uguale a 0 per la maggior parte delle telecamere.

In definitiva la matrice di calibrazione viene espressa come:

$$K = \begin{bmatrix} f_x & s & x_0 \\ & f_y & y_0 \\ & & 1 \end{bmatrix} \quad (3.11)$$

Tale modello di telecamera ha ora 11 GDL, due in più rispetto al precedente per l'aggiunta del parametro di skew e per la possibile diversità tra le due lunghezze focali. Il modello pinhole camera descrive solamente relazioni lineari tra i parametri di telecamera e risulta quindi essere molto semplice e in molti casi offre una buona approssimazione della realtà.

3.3 Il modello non lineare della telecamera

In alcuni casi un modello lineare come quello descritto nel paragrafo precedente non risulta sufficientemente adeguato a descrivere la realtà. Infatti la semplicità del modello pinhole lineare non è sufficiente nei casi in cui sono presenti delle distorsioni ottiche considerevoli. Tali distorsioni possono presentarsi in diversi modi, ma generalmente sono individuabili come delle non-linearità del sistema di visione. Il modello viene quindi modificato ulteriormente per avvicinarsi maggiormente alla geometria reale di una telecamera, nella quale i punti del mondo, nel piano immagine e il centro ottico non sono più collineari (ovvero collegati da una linea retta) come nel caso ideale. Tali effetti sono molto visibili in figura 3.7: le linee dell'immagine, che dovrebbero essere rette e parallele tra loro sono invece curvate in maniera tanto più accentuata quanto più si trovano ai bordi dell'immagine.



Figura 3-7: *Differenza tra un'immagine soggetta a distorsioni radiali (a sinistra) e una non soggetta (a destra).*

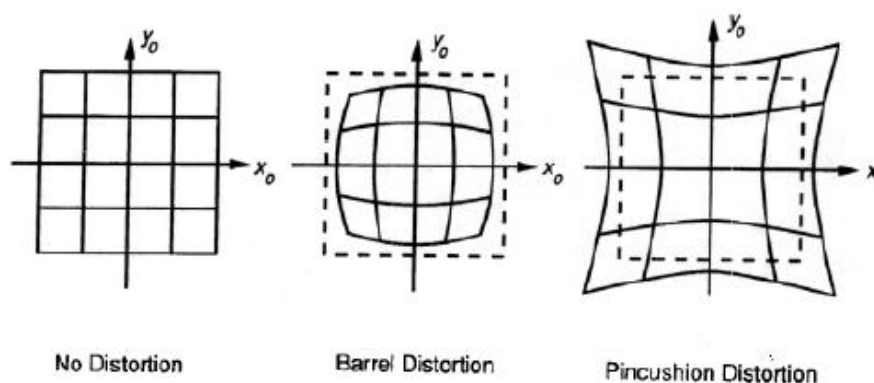


Figura 3-8: Rappresentazione delle tipologie di distorsioni radiali; da sinistra verso destra: indeformato, a botte, a puntaspilli.

Tale distorsione viene chiamata *radiale* e può verificarsi in due modalità, a seconda del segno della distorsione; a botte oppure a puntaspilli (figura 3.8). Questa distorsione solitamente diventa più accentuata con il diminuire della distanza focale.

Una volta corrette le distorsioni, il fenomeno di formazione delle immagini può essere nuovamente ricondotto al caso lineare. Il processo per effettuare questa correzione viene mostrato in figura 3.10; in coordinate pixel risulta:

$$\hat{x} = x_c + L(r)(x - x_c) \quad \hat{y} = y_c + L(r)(y - y_c) \quad (3.12)$$

$$\text{con } r^2 = (x - x_c)^2 + (y - y_c)^2$$

dove $(x; y)$ sono le coordinate immagine misurate, $(\hat{x}; \hat{y})$ sono le coordinate corrette, mentre $(x_c; y_c)$ è il centro delle distorsioni radiali. La funzione $L(r)$ è definita solo per valori positivi di r e viene solitamente risolta come uno sviluppo di un'equazione di Taylor $L(r) = 1 + k_1 r + k_2 r^2 + k_3 r^3 + \dots$

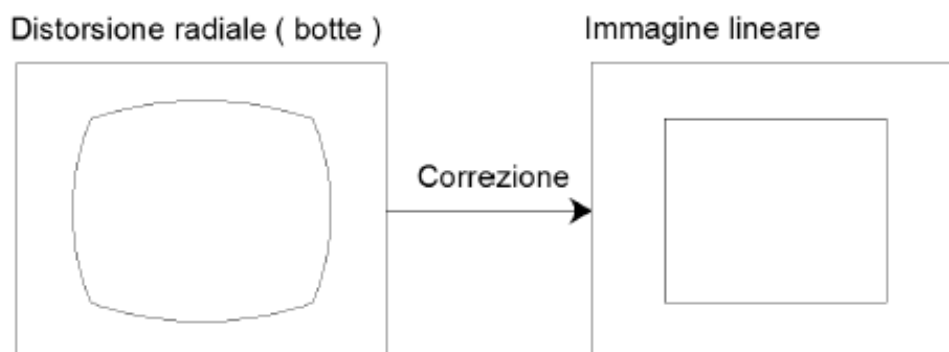


Figura 3-9: Correzione di un'immagine soggetta a distorsione radiale.



Figura 3-10: Esempio di rettificazione di un'immagine soggetta a distorsioni radiali

I coefficienti di distorsione k_1, k_2, k_3, \dots sono considerati parte della calibrazione interna della telecamera.

Un modello migliore per trattare tali aberrazioni ottiche venne introdotto da Brown nel 1966 [9], chiamato modello “*Plumb Bob*” (piombino). Lo spostamento del centro delle distorsioni viene esplicitato come una distorsione tangenziale, dovuta solitamente a un problema di montaggio della lente o a un difetto costruttivo. Indicando con x_n le coordinate normalizzate di telecamera (x, y) la correzione delle coordinate diventa:

$$x_d = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) x_n + dx \quad (3.13)$$

Con dx termine che descrive lo spostamento introdotto dalle distorsioni tangenziali:

$$dx = \begin{bmatrix} 2p_1 xy + p_2 (r^2 + 2x^2) \\ p_1 (r^2 + 2y^2) + 2p_2 xy \end{bmatrix} \quad (3.14)$$

Si è infine giunti al modello completo per descrivere la telecamera. Questo è caratterizzato da 16 GDL, separabili in intrinseci ed estrinseci. La calibrazione è l'insieme delle operazioni finalizzate a determinare tali parametri, che a scopo riassuntivo vengono qui indicati nella loro notazione completa e suddivisi in parametri intrinseci, estrinseci e di distorsione:

Modello pinhole camera completo:

- 5 parametri intrinseci:
 - f_1, f_2 ovvero la distanza focale espressa nelle due direzioni
 - c_1, c_2 le due coordinate del centro ottico
 - s ovvero lo skew
- 6 parametri estrinseci:
 - r_1, r_2, r_3 tre parametri di rotazione esprimibili in angoli di Eulero, Rodrigues etc.
 - t_1, t_2, t_3 tre parametri di traslazione
- 5 parametri di distorsione:
 - k_1, k_2, k_3 tre distorsioni radiali
 - p_1, p_2 due distorsioni tangenziali

Non sempre però è necessario ricorrere all'utilizzo del modello completo; in alcuni casi, infatti, la complessità introdotta da questo modello è eccessiva per le applicazioni in cui esso è utilizzato. Spesso le distorsioni sono trascurabili; in particolare è abbastanza frequente che non si presentino le prime due distorsioni radiali, mentre è ancora più frequente che non si presentino quelle tangenziali e l'ultima radiale [12], [13]. Per motivi legati alla convergenza dei metodi di minimizzazione, considerare nel modello anche parametri che non sono presenti in realtà, conduce spesso a risultati di calibrazione sbagliati.

3.4 Geometria epipolare

Prima di eseguire la triangolazione stereoscopica è necessario individuare le corrispondenze tra i punti nelle due immagini, detti *omologhi*, che sono la proiezione del medesimo punto nello spazio (figura 3.11). Questa procedura non è banale e per fare ciò bisogna ricorrere alla *geometria epipolare*. Essa descrive le relazioni geometriche tra due viste prospettiche della medesima scena 3D.

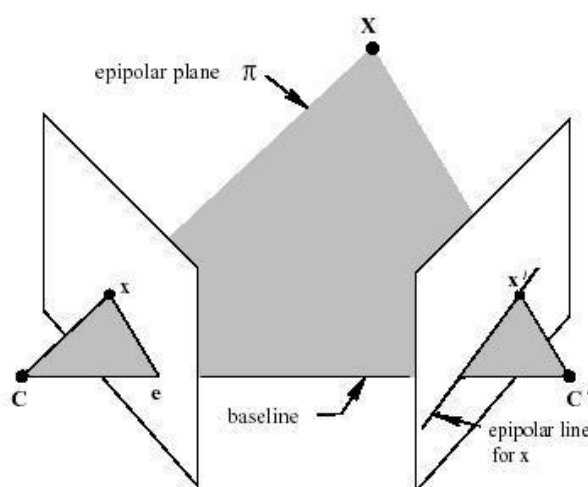


Figura 3-11: Geometria dei punti omologhi. Il piano univocamente definito dalle coordinate dei centri delle telecamere e il punto X è detto piano epipolare. La proiezione della linea congiungente il centro della telecamera di sinistra al punto X (che incontra il piano della telecamera nel punto x) sul piano immagine destro è detta retta epipolare; il punto x' omologo di x giacerà idealmente su tale retta.

Date le coordinate dei due centri telecamera e le coordinate di un punto X nello spazio è possibile definire univocamente un piano chiamato piano epipolare. Collegando i centri delle telecamere tra loro si ottiene una linea chiamata *baseline* (o linea base), che intersecata con i piani dei sensori genera due punti chiamati *epipoli* ($e;e'$). Questi punti possono essere visti come la proiezione nella seconda immagine del centro ottico della prima e viceversa. Se si proietta il raggio di proiezione (ovvero la retta che unisce il centro della telecamera C al punto X) sul piano della seconda immagine, si trova la cosiddetta *retta epipolare*, che passa per il secondo epipolo e' e la proiezione del punto X, cioè x' . È dunque facile intuire

come dato un punto in un'immagine, il suo corrispondente nella seconda immagine possa trovarsi solamente lungo una linea passante per l'epipolo (la retta epipolare). Questa proprietà riduce notevolmente la complessità computazionale della ricerca dei punti omologhi, anche se in realtà, a causa di errori di misura, il punto coniugato non viene sempre a trovarsi esattamente sulla retta epipolare bensì molto vicino ad essa.

Per ricavare le rette epipolari è necessario conoscere la posa relativa tra le telecamere che riprendono la scena, informazione ricavabile grazie al processo di calibrazione ed espressa algebricamente tramite la *matrice fondamentale di telecamera*. Tale matrice, omogenea 3x3, di rango 2 e con 7 gradi di libertà, descrive la corrispondenza tra i punti nelle due immagini, e per definizione risulta che:

$$x'F x = 0 \quad (3.15)$$

Dove x e x' sono le coppie di punti omologhi.

Esistono due tipi di algoritmi che permettono di calcolarla: uno basato direttamente sulle corrispondenze tra i punti delle due immagini (almeno 8), il secondo basato sulla conoscenza della matrici di calibrazione. Il primo metodo è caratterizzato da un errore maggiore di ricostruzione sui punti rispetto il secondo.

Se le matrici di calibrazione delle due telecamere sono ottenute da immagini comuni espresse nel medesimo sistema di riferimento, è possibile ricavare direttamente la matrice fondamentale. Come visto nel paragrafo 3.2 le matrici di proiezione di telecamera sono così definite:

$$P = K[R t] \quad P' = K' [R' t'] \quad (3.16)$$

Dove K e K' sono le matrici dei parametri intrinseci delle due telecamere e $[R t]$, $[R' t']$ la posizione relativa delle due telecamere rispetto ad un SDR comune. Dalla conoscenza di queste matrici di proiezione è possibile ricavare la posizione dei centri delle due telecamere (nello stesso sistema di riferimento) tramite scomposizione ai valori singolari della matrice P . Il centro della telecamera corrisponde, infatti, al vettore associato più piccolo valore singolare della matrice di proiezione:

$$[U; S; V] = SVD(P) \longrightarrow C = v_{end} \quad (3.17)$$

Dove v_{end} indica l'ultima colonna della matrice V. Una volta nota la posizione dei centri è possibile ricavare la posizione degli epipoli come:

$$e' = P'C \quad (3.18)$$

Come in precedenza gli apici indicano la seconda telecamera, mentre i valori senza apice sono riferiti alla prima telecamera. Dagli epipoli è poi ricavabile la matrice fondamentale come:

$$F = e'P'P^+ \quad (3.19)$$

Dove P^+ è la pseudo-inversa della matrice di telecamera.

Una volta ricavata la matrice fondamentale e sapendo che essa rappresenta una correlazione, ovvero una mappatura proiettiva tra i punti nelle immagini e la linea epipolare che essi generano sull'altra immagine, si può scrivere che:

$$l' = Fx \quad l = F^T x' \quad (3.20)$$

Le coordinate della linea che si ottengono sono espresse in un vettore 3x1, che descrive l'equazione $l_1x + l_2y + l_3 = 0$. Può essere riportato all'equazione esplicita $y = mx + q$ con $m = -l_1/l_2$ e $q = -l_3/l_2$. Tale calcolo può essere eseguito per tutti i punti dell'immagine e per ognuna delle due immagini in relazione all'altra. Si possono quindi creare su ogni immagine una griglia di rette corrispondenti alle epipolari ottenute rispetto ai punti nell'altra immagine e utilizzare tali rette non solo per individuare le corrispondenze tra punti ma anche per stimare le incertezze di ricostruzione sui punti stessi (figura 3.12).

Anche se in posizioni totalmente diverse all'interno delle immagini diventa facile tramite questo metodo ricavare la posizione dei punti omologhi ricercandoli solamente nell'intorno delle rette epipolari generate dalla prima immagine. Una volta eseguita l'associazione è possibile eseguire la triangolazione dei punti.

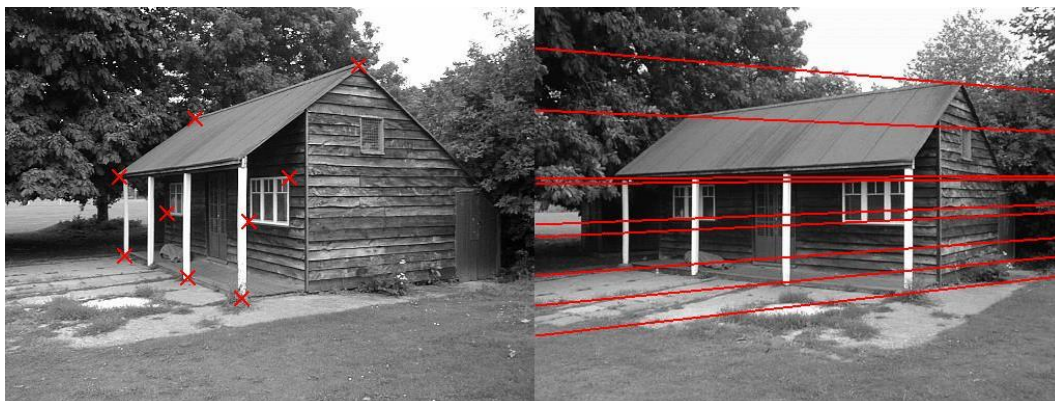


Figura 3-12: Esempio di geometria epipolare. Per ogni punto selezionato nella prima immagine sono state tracciate le rette epipolari nella seconda immagine.

Un limite importante di questa tecnologia proviene dal fatto che, nel caso in cui più marker sono disposti lungo linee parallele alla baseline, s'incorre in casi di ambiguità (figura 3.13). Per risolvere questo problema si può ricorrere all'aggiunta di una terza telecamera, aumentando la complessità del sistema di misura ma aggiungendo stabilità al metodo d'individuazione dei punti omologhi basato sulle rette epipolari. Tale costruzione potrebbe essere affrontata introducendo il tensore trifocale, ovvero la matrice duale della fondamentale ma valida per una tripletta di telecamere. Esso non viene però calcolato in questo lavoro di tesi poiché aumenta la complessità del modello del sistema di acquisizione senza portare benefici considerevoli in termini di accuratezza; è ben più semplice e utile vedere il sistema come costituito da tre coppie di telecamere, esprimendo per ognuna di esse la matrice fondamentale rispetto alle altre due singolarmente. Se vengono tracciate le rette epipolari corrispondenti a due punti omologhi sulla terza immagine esse si incontreranno in un punto, che corrisponde al punto che rispetta i vincoli epipolari per entrambe le immagini. Essendo l'obiettivo corrente quello di individuare i punti omologhi senza ambiguità, si può limitare la ricerca dei punti omologhi solamente nei dintorni di tale punto ideale e non lungo un'intera linea epipolare (figura 3.14).

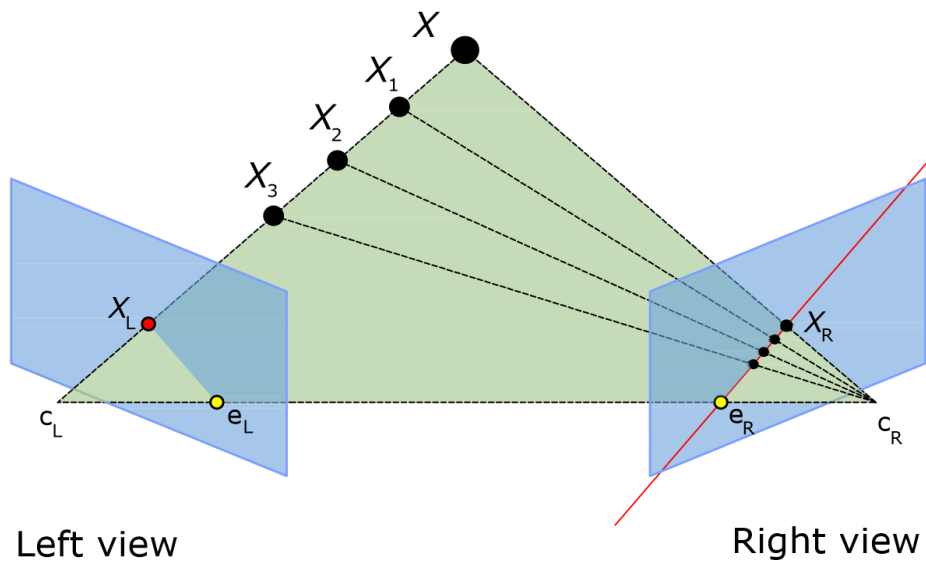


Figura 3-13: Esempio di caso in cui è possibile avere ambiguità nel riconoscimento dei punti omologhi.

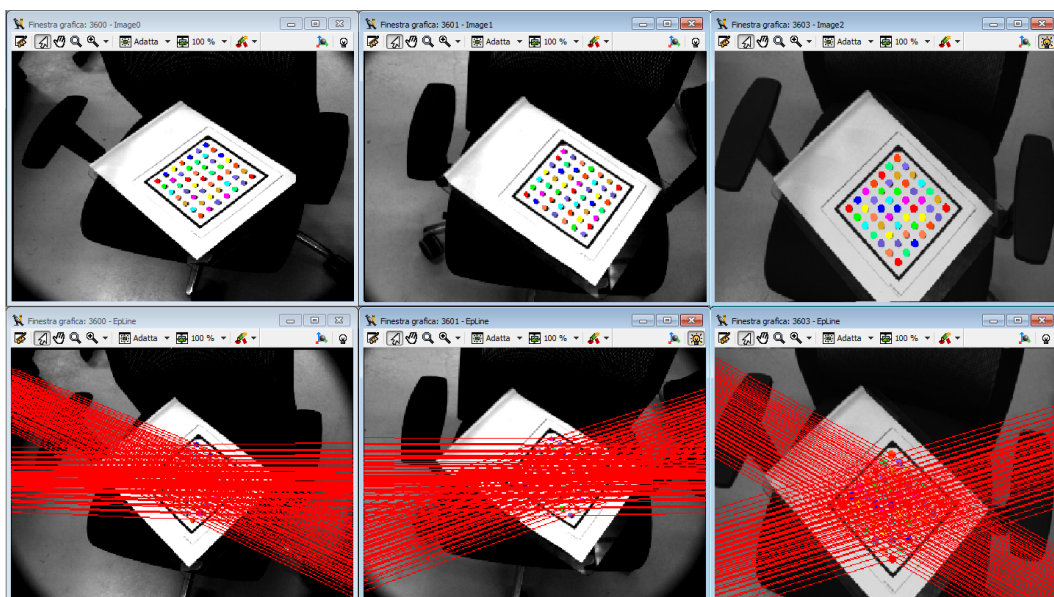


Figura 3-14: Esempio di individuazione delle rette epipolari per un calibratore planare mediante l'uso di una tripletta di telecamere. I punti del calibratore sono stati trovati utilizzando l'algoritmo di blob detection.

La procedura completa per eseguire questa individuazione diventa:

1. Per ogni punto nella prima immagine si traccia la retta epipolare nella seconda.
2. Per tutti i punti con distanza da tale retta entro una determinata soglia (*tolleranza*) si traccia la retta epipolare nella terza immagine.
3. Per tutti i punti con distanza da tale retta epipolare entro una determinata soglia si traccia la retta epipolare nella prima immagine.
4. Se la retta epipolare tracciata incontra il punto dal quale è cominciata l'analisi, allora si seleziona la terna di punti.
5. Per ogni punto della terna si calcola la distanza dal punto d'incontro delle due rette epipolari nell'immagine.
6. Viene calcolata la media delle distanze.
7. Per ogni terna si seleziona quella che ha media delle distanze minore e si passa al punto successivo.

Questa procedura che consente di eseguire una ricerca di corrispondenze nei dintorni di un punto, risulta maggiormente mirata, fornisce cioè un numero minore di false associazioni rispetto alla tecnica con due telecamere.

4. Calibrazione di telecamere

Come ogni strumento di misura, anche un sistema ottico stereoscopico necessita di essere tarato prima di eseguire qualsiasi misura. La taratura di tale strumento consiste nella determinazione dei parametri costituenti il modello matematico per ogni telecamera e nella determinazione delle loro pose rispetto a un sistema di riferimento comune. La calibrazione delle telecamere consiste nella stima dei parametri che riguardano direttamente il dispositivo (intrinseci) e di quelli che ne descrivono la posa (estrinseci). È quindi possibile calibrare le singole telecamere che compongono la tripletta e poi esprimerne le pose in un sistema di riferimento comune a tutti i dispositivi. La calibrazione delle singole telecamere risulta dunque essere la procedura più critica dell'intero processo di taratura del sistema, perciò si vuole dedicare un capitolo solamente ad essa. In letteratura non esiste un metodo standard da utilizzare per calibrare. La tecnica più utilizzata è l'algoritmo di Zhang. Questa è già stata implementata, validata e ampiamente utilizzata per alcune attività di ricerca del VBLab ed è dunque stata presa come riferimento per il confronto con altre tecniche più nuove e meno consolidate. Tra queste in particolare si è voluto valutare le prestazioni della tecnica di Heikkilä una tecnica implementata nelle librerie commerciali HALCON di MVTec che sono in uso presso il VBLab da pochi anni. In questo capitolo si approfondirà la procedura di calibrazione secondo il metodo di Zhang, ottimizzato mediante l'algoritmo di Levenberg-Marquardt e la tecnica di Heikkilä. Verranno poi riportati alcuni risultati sul confronto tra queste due tecniche principali.

4.1 Calibrazione mediante il metodo di Zhang e di Levenberg-Marquardt

La calibrazione è un processo essenziale per poter effettuare misure tramite telecamere. Nel paragrafo 2.2 sono state descritte le principali tecniche presenti in letteratura per affrontare questo problema. Le applicazioni nella quale si utilizzeranno i dispositivi studiati in questo lavoro di tesi richiedono che i risultati di calibrazione siano molto accurati. Altra specifica importante è che il setup del processo di calibrazione possa essere semplice ed economico. L'intera procedura deve essere svolta nell'arco di qualche minuto (al massimo poche ore), quindi si può dire che il tempo non è una specifica critica; la calibrazione può essere infatti effettuata prima di eseguire le misure, una volta impostato l'apertura del diaframma e la focale. In caso venga variato uno di questi parametri l'intera procedura deve necessariamente essere rieseguita. Se invece viene modificata solamente la posa sarà necessario registrare la telecamera, cioè calcolarne la posa rispetto a un sistema di riferimento fissato; andranno quindi ricalcolati solamente i parametri estrinseci, senza ripetere l'intera calibrazione.

Il metodo di Zhang rispetta tutte le specifiche sopra riportate ed è una tecnica molto diffusa e utilizzata in letteratura per risolvere il modello lineare; i risultati di questo poi possono essere utilizzati per inizializzare un problema di ottimizzazione del modello non lineare completo. Questo è stato risolto utilizzando l'algoritmo di Levenberg-Marquardt (LM) come vedremo in seguito. La tecnica di Zhang-LM è già stata validata nel corso degli anni da molti studiosi [6] ed è stata già implementata dallo staff del VBLab. Inoltre in [13] si sono eseguite numerose prove con lo scopo di studiare più a fondo questo approccio alla calibrazione. Partendo allora dalle conoscenze già acquisite su tale tecnica dal VBLab, si vuole quindi confrontarla con altre metodologie. Si passa ora a descrivere più in dettaglio la procedura di calibrazione utilizzando questi algoritmi.

4.1.1 Acquisizione dei dati di calibrazione

L'acquisizione delle immagini del calibratore in pose differenti si può considerare la parte più critica dell'intera procedura di calibrazione. Infatti se non svolta in maniera corretta, si rischia di introdurre numerosi errori che limitano le prestazioni degli algoritmi di calibrazione. Prima di tutto bisogna scegliere un calibratore adeguato per l'applicazione, scelta essenziale per poter ottenere una buona accuratezza di calibrazione. Di norma esso deve occupare da $1/8$ a $1/4$ dell'immagine [13]. Una volta scelta la focale e l'apertura del diaframma, sarà possibile acquisire immagini di oggetti solamente a una data *profondità di campo*, cioè l'insieme di distanze in cui l'oggetto appare a fuoco. È, infatti, necessario che tutti i punti noti del calibratore siano a fuoco. Questo fattore limita gli spostamenti dell'oggetto sull'asse principale della telecamera; le dimensioni del calibratore allora dipendono dalla profondità di campo in cui si vuole acquisire misurazioni.

Una volta individuate le dimensioni del calibratore da utilizzare, bisogna determinarne le caratteristiche; esso, infatti, deve necessariamente possedere alcune proprietà fondamentali:

- Punti noti facilmente distinguibili (almeno 4).
- Distanze tra i punti noti conosciute con accuratezza.
- Caratteristiche morfologiche accurate (tolleranze di planarità e parallelità molto elevate).
- Osservabile con la telecamera utilizzata (ad esempio costruito con marker rifrangenti se vengono utilizzati filtri ad infrarossi)

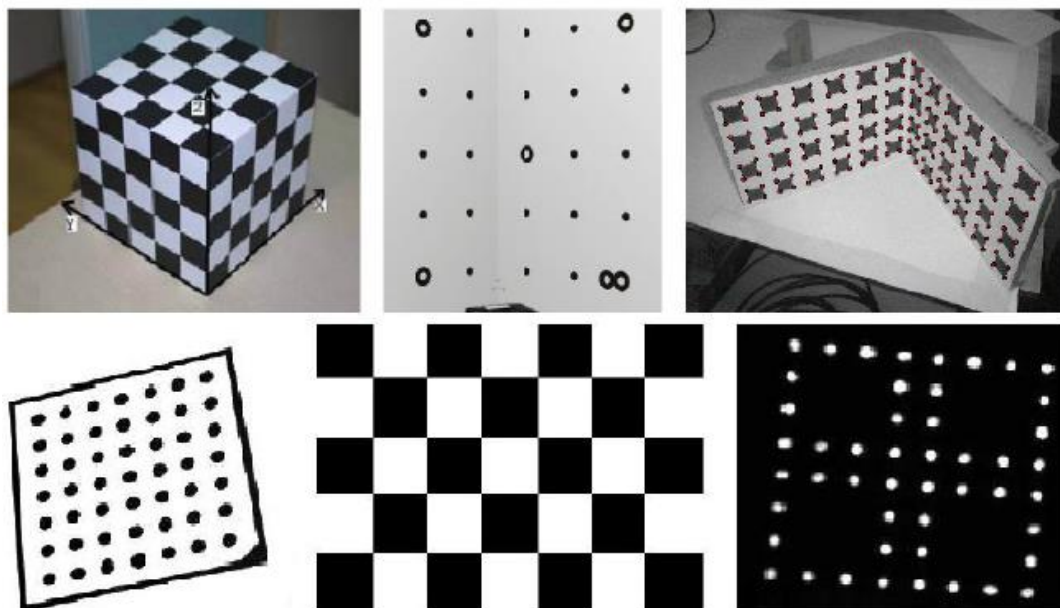


Figura 4-1: Esempi di calibratori 3D e planari.

In virtù delle caratteristiche sopra elencate, generalmente i punti notevoli sono rappresentati da *blob* (regioni nell'immagine che sono più chiare o scure rispetto alle circostanti) o *corner*. Gli spigoli sono invece dei punti che corrispondono a brusche variazioni d'intensità di scala di grigi. Nella figura 4.1 sono rappresentati diversi esempi di calibratore. Le figure sotto rappresentano calibratori planari, mentre sopra sono raffigurati esempi di calibratori tridimensionali, che non sono altro che un insieme di calibratori planari messi in pose differenti nello spazio. Si nota poi come alcuni di essi presentano alcuni cerchi (*blob*) facilmente distinguibili dal resto della scena, mentre altri sono caratterizzati da zone quadrate che presentano *corners* facilmente determinabili.

Dopo aver scelto il calibratore più adatto, si procede all'acquisizione delle immagini. Esse devono essere acquisite in modo da caratterizzare al meglio tutto il sensore e con una qualità elevata, così da poter determinare con accuratezza la posizione dei punti d'interesse. Più si varia la posizione del calibratore, in modo da coprire tutto lo spazio in cui avverranno le misure, e più la calibrazione sarà accurata. Anche le pose devono essere più varie possibili; come si vedrà nel seguito, è indispensabile per risolvere il problema di calibrazione che nelle immagini

acquisite il calibratore giaccia su almeno tre piani differenti. Chiaramente la procedura di acquisizione delle immagini si può eseguire sia muovendo il calibratore che spostando le telecamere in diverse posizioni.

Anche l'illuminazione della scena rappresenta un importante parametro per la riuscita della calibrazione. I punti noti del calibratore, infatti, devono ricevere un'illuminazione omogenea e tale da creare un forte contrasto tra essi e il resto della scena (senza però saturare l'immagine), in modo da renderli ben visibili e riconoscibili.

Le tecniche più utilizzate per l'estrazione delle coordinate nelle immagini dei punti noti sono l'analisi dei *blob* [14] e l'algoritmo di *corner detection* di Harris [15]. L'analisi dei *blob* individua i punti d'interesse tramite la binarizzazione dell'immagine e dalla successiva individuazione delle regioni adiacenti caratterizzate dalla medesima polarità. Per fare questo occorre solitamente definire una soglia di taglio (*threshold*), che viene scelta per discriminare in chiaro o scuro i 255 livelli dell'immagine, separando in tal modo quello che viene chiamato sfondo dalle regioni di interesse. L'immagine viene così binarizzata e può successivamente essere analizzata con gli strumenti tipici della *blob analysis*, ovvero tutto quell'insieme di algoritmi che di tali regioni calcolano area e forme.

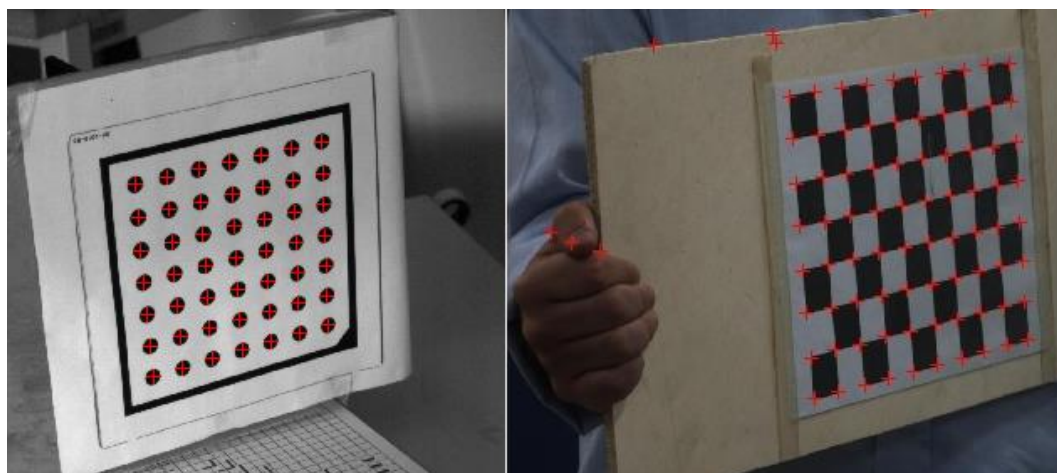


Figura 4-2: Esempio di applicazione dell'algoritmo di *blob detection* (a sinistra) e di *Harris corner detection* (a destra).

Il metodo di Harris permette di determinare le coordinate degli spigoli (o *corners*) presenti nell'immagine. Questo, proposto da C. Harris e Stephens [15],

tratta il problema dell'individuazione dei corners attraverso funzioni locali di autocorrelazione dell'intensità d'immagine. Vengono create tante piccole finestre di dimensioni fisse all'interno della quale si esegue la ricerca dei corners. Per ognuna di queste si ricava una funzione che esprime la variazione di intensità nelle due direzioni. Tale intensità viene utilizzata per individuare se la finestra corrente contiene un angolo. La finestratura viene eseguita per tutta l'immagine, individuando così una matrice di punti con la posizione di tutti gli angoli individuati.

Sono disponibili software che permettono a un utente dotato di competenze non specializzate di impostare direttamente le soglie corrette per effettuare le analisi sopra descritte. Presso il laboratorio VBLab, ad esempio, è stato sviluppato un software chiamato "FeatureLinker". Esso consente di trovare in maniera manuale le corrispondenze tra i punti del modello che descrive il calibratore e i punti nell'immagine. Si è deciso di utilizzare un approccio manuale, più lento e laborioso di un approccio completamente automatico, perché con un riscontro visivo immediato, si riduce la possibilità di avere errori in fase di correlazione dei punti. Le librerie HALCON invece hanno al loro interno un toolbox per l'acquisizione delle immagini di calibrazione chiamato "Calibration Assistant". Questo consente di trovare automaticamente un particolare calibratore nella scena e verificare la bontà dell'immagine acquisita. Si aggiunge infine che per ottenere informazioni sull'orientamento di calibratori quadrati o rettangolari è bene inserire un'asimmetria nel calibratore.

4.1.2 Calcolo del modello lineare

Se l'applicazione per cui viene eseguita la calibrazione non richiede accuratze elevate o se le distorsioni delle lenti sono trascurabili, il calcolo del modello lineare rappresenta l'intero processo di calibrazione. Se invece è importante utilizzare un più completo modello non lineare, allora i risultati provenienti dal problema lineare possono essere utilizzati per inizializzare un algoritmo di ottimizzazione non lineare. Alcune tecniche di calibrazioni non

utilizzano l'algoritmo di Zhang, ma si limitano a inizializzare i parametri con valori sensati (dimensioni del sensore, dimensioni del pixel, ottica montata), lasciando alla sola ottimizzazione finale il compito di convergere ai valori veri. Chiaramente non sempre è possibile conoscere questi dati con precisione e dunque risulta più utile partire dalla tecnica di Zhang.

L'algoritmo di Zhang prende in ingresso le coordinate che descrivono il calibratore planare e i corrispondenti punti nelle varie immagini e consente di calcolare i parametri intrinseci lineari ed estrinseci. Dato che il sistema di riferimento è considerato solidale al calibratore, i parametri estrinseci della telecamera saranno diversi per ogni immagine acquisita. Inoltre è possibile considerare nulla la coordinata Z dei punti nello spazio, cioè la coordinata ortogonale al piano del calibratore. Questo ci consente di limitare il problema al semplice calcolo di più omografie. L'*omografia* è la relazione tra punti in due spazi tali per cui ogni punto di uno spazio corrisponde ad uno e un solo punto del secondo spazio.

Per eseguire il calcolo dell'omografia si procede come segue:

1. Si normalizzano i punti X ovvero i punti mondo bidimensionali (poiché $Z=0$).
2. Si normalizzano i punti x ovvero i punti ricavati dalle immagini.
3. Si applica l'omografia H fra i punti del modello del calibratore e i punti trovati nell'immagine.
4. Si denormalizza la matrice H .

La normalizzazione è necessaria per poter avere un migliore condizionamento del problema. Così facendo si riferiscono i punti del calibratore ad una posizione fissa nello spazio. Solitamente tale posizione è scelta in modo da posizionare il centro di massa dei punti nell'origine $(0; 0)^T$, scalati in modo che la loro distanza media dall'origine sia $\sqrt{2}$.

La normalizzazione si riconduce nell'individuare una matrice di trasformazione che la effettua, la cui equazione può essere ricavata da libri di testo (ad esempio [8]) e risulta essere:

$$T = \begin{bmatrix} \frac{\sqrt{2}}{std(x)} & 0 & -\left(\frac{\sqrt{2}}{std(x)}\right)m_x \\ 0 & \frac{\sqrt{2}}{std(y)} & -\left(\frac{\sqrt{2}}{std(y)}\right)m_y \\ 0 & 0 & 1 \end{bmatrix} \quad (4.1)$$

In cui m_x e m_y sono le medie dei punti nelle due direzioni e $std(x)$ e $std(y)$ sono le deviazioni standard. I punti normalizzati mediante la matrice di trasformazione T saranno $\tilde{X} = TX$. Per i punti immagine si applica la medesima procedura ottenendo anche qui i loro corrispondenti normalizzati $\tilde{x} = Tx$. Eseguita la normalizzazione si può applicare l'algoritmo per il calcolo delle omografie che, partendo dalle corrispondenze $\tilde{X} \leftrightarrow \tilde{x}_i$, determina la matrice H per cui $\tilde{x}_i = H\tilde{X}_i$. La procedura viene descritta in dettaglio in [8]. La matrice H così individuata è una 3x3 con 8 gradi di libertà. Definendo con h_i le colonne di H e con r_i le colonne della matrice di rotazione è possibile esprimere, per un calibratore planare, la relazione che lega i parametri intrinseci all'omografia:

$$H = [h_1 \ h_2 \ h_3] = K [r_1 \ r_2 \ t] \quad (4.2)$$

Questa espressione è possibile se il calibratore è planare ovvero $Z=0$; infatti la terza colonna della matrice di rotazione, essendo moltiplicata per zero, non viene indicata.

La matrice H è dipendente dunque sia dai parametri intrinseci che da quelli estrinseci. Ricordando che per ogni immagine di calibrazione viene calcolata una omografia, degli 8 GDL di H ne vengono utilizzati 6 per definire la posa della telecamera, ovvero la matrice di rotazione e il vettore di traslazione, quindi ne rimangono solo 2 per definire i parametri intrinseci. Questo vincolo a volte non è sufficiente poiché, a seconda del modello utilizzato, per rappresentare la telecamera, si deve fare ricorso ad un numero di parametri intrinseci maggiore di 2. Per far fronte a questo problema bisogna disporre di corrispondenze tra punti in più

immagini in cui il calibratore è posto in pose differenti tra loro. Per definire 5 parametri intrinseci si dovranno acquisire almeno 3 immagini. A questo punto, per determinare i parametri intrinseci, è necessario imporre un vincolo costante in tutte le immagini che viene utilizzato per correlarle tra loro; tale vincolo è quello di ortonormalità tra r_1 e r_2 , che equivale ad imporre che gli assi x e y siano ortogonali tra loro in tutte le immagini:

$$\begin{cases} h_1^T K^{-T} K^{-1} h_2 = 0 \\ h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \end{cases} \quad (4.3)$$

La soluzione viene trovata, come mostrato da Zhang in [2], indicando con $B = K^{-T} K^{-1}$ e con $b = [B_{11} \ B_{12} \ B_{22} \ B_{13} \ B_{23} \ B_{33}]^T$. L'equazione 4.3 diventa:

$$h_i^T B h_i = v_{ij} b \quad (4.4)$$

Con v_{ij} costruito come:

$$v_{ij} = [h_{i1} h_{j1}; h_{i1} h_{j2} + h_{i2} h_{j1}; h_{i2} h_{j2}; h_{i3} h_{j1} + h_{i1} h_{j3}; h_{i3} h_{j2} + h_{i2} h_{j3}; h_{i3} h_{j3}]$$

È possibile allora risolvere l'equazione formulando il problema come:

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \quad (4.5)$$

ovvero nella forma $Vb = 0$ risolvibile tramite la tecnica della decomposizione ai valori singolari SVD. I parametri intrinseci si estraggono direttamente dalla matrice B tramite le equazioni:

$$\begin{aligned} x_0 &= (B_{12} B_{13} - B_{11} B_{23}) / (B_{11} B_{22} - B_{21}^2) \\ \lambda &= B_{33} - [B_{13}^2 + x_0 (B_{12} B_{13} - B_{11} B_{23})] / B_{11} \\ f_1 &= \sqrt{\lambda / B_{11}} \end{aligned} \quad (4.6)$$

$$f_2 = \sqrt{\lambda B_{11} / (B_{11} B_{22} - B_{21}^2)}$$

$$y_0 = s x_0 / f_2 - B_{13} f_1^2 / \lambda$$

dove λ indica il coefficiente di scala.

Per quanto riguarda i parametri estrinseci essi possono essere ricavati, una volta determinata la matrice K , come:

$$\begin{aligned} r_1 &= \lambda K^{-1} h_1 \\ r_2 &= \lambda K^{-1} h_2 \\ r_3 &= r_1 \times r_2 \\ t &= \lambda K^{-1} h_3 \end{aligned} \quad (4.7)$$

In generale, la matrice di rotazione ottenuta potrebbe non rispettare il vincolo di ortonormalità ($R^T R = I$) a causa di errori sui punti; si riporta la matrice di rotazione ad una che rispetta il vincolo risolvendo il problema di minimizzazione:

$$\min_R \|R - Q\|_F^2 \quad \text{soggetto a } R^T R = I \quad (4.8)$$

Questo problema può essere risolto più semplicemente mediante la scomposizione ai valori singolari, esprimendo $Q = USV^T$ dove S è una matrice diagonale contenente gli autovalori di Q , mentre U e V sono matrici unitarie. Si ricava R come:

$$R = UV^T \quad (4.9)$$

La matrice qui trovata rappresenta la soluzione del problema ai minimi quadrati [8].

4.1.3 Ottimizzazione non lineare

É ora possibile utilizzare i risultati ottenuti mediante la procedura lineare di Zhang fin qui descritta come soluzione di primo tentativo per la risoluzione del problema non lineare di calibrazione. Per fare ciò si utilizza uno dei metodi iterativi per la risoluzione del problema non lineare di minimizzazione; tali tecniche presentano solitamente alcuni svantaggi:

- la convergenza può essere lenta;
- richiedono valori iniziali prossimi alla soluzione;
- esiste la possibilità di non-convergenza, o convergenza ad un minimo locale anziché globale;
- la scelta dei criteri di terminazione dell'iterazione non è sempre banale;

- il problema numerico fa perdere ai parametri il loro rigido significato fisico, generando soluzioni inaccettabili fisicamente e difficili da verificare.

Di conseguenza, tali tecniche richiedono generalmente una più attenta implementazione, valutazione dei risultati ottenuti. Le tecniche di minimizzazione iterativa consistono generalmente in cinque passaggi:

1. Funzione di costo: una funzione di costo è scelta come nucleo della minimizzazione. Per il problema corrente è possibile impostare diverse funzioni.
2. Parametrizzazione: le trasformazioni devono essere espresse in funzione di un finito numero di parametri. Non sempre è necessario considerare il set minimo di parametri, ma in alcuni casi una sovrparametrizzazione può portare dei vantaggi.
3. Scelta della funzione di costo.
4. Inizializzazione: il calcolo dei parametri iniziali deve essere eseguito con accuratezza; nel caso corrente si può utilizzare il metodo di Zhang descritto sopra.
5. Iterazioni: iniziando dalla soluzione di primo tentativo, i parametri sono iterativamente migliorati minimizzando la funzione di costo.

È possibile utilizzare differenti funzioni di costo; in letteratura [8], [16] sono utilizzati ad esempio:

- Distanza algebrica
- Distanza geometrica in una o due immagini
- Errore di retroproiezione
- Errore di Sampson
- Funzioni di costo statistiche (un'immagine, due immagini, distanza di Mahalanobis)

In questa tesi, come nella maggior parte della letteratura [6], [7], [8], è stato scelto di utilizzare, come funzione obiettivo per l'ottimizzazione non lineare, l'*errore di retroproiezione*, cioè la distanza tra i punti individuati nelle immagini (i dati di partenza) e i punti del calibratore trasformati sulla base dei parametri di telecamera, ovvero retroproiettati. Questa scelta è stata fatta poiché l'errore di

retroproiezione mantiene fortemente un significato fisico ed è quindi immediato da verificare con riscontri visivi, quali grafici o immagini. Inoltre questa funzione di costo è relativamente semplice da implementare ed equivale a cercare il migliore set di parametri che minimizzi la funzione:

$$\sum_i d(x'_i, H_i x_i)^2 \quad (4.10)$$

Dove $d(x, y)$ rappresenta la distanza euclidea tra i punti individuati nell'immagine x'_i e i punti del calibratore retroproiettati $H_i x_i$.

Per ogni funzione di costo ci sono differenti scelte di parametrizzazione. In generale nessun effetto negativo è stato riscontrato se il problema di minimizzazione viene sovra-parametrizzato. In particolare, algoritmi come quello di Gill-Murray [17] (che è una modifica di quello di Gauss-Newton) o quello di Levenberg-Marquardt [8], [16], [18], [19] gestiscono parametri ridondanti facilmente.

Per semplificare il processo di parametrizzazione si osserva che $H_i = K[R_i t_i]$, vale a dire che la matrice H può essere scomposta in due parti separate: un insieme di parametri dipendente solamente dai parametri intrinseci (e dunque comuni a tutte le immagini) e un insieme di parametri estrinseci dipendente dalla particolare immagine. L'equazione complessiva potrà essere ottenuta utilizzando un insieme di parametri creato in modo da tenere conto solamente una volta della matrice K e separatamente, per ogni immagine, della sua rototraslazione relativa. La parametrizzazione del problema diventa:

$$f : (K, R_1, t_1, \dots, R_n, t_n) \quad (4.11)$$

L'equazione da minimizzare risulta quindi:

$$\sum_{i=1}^n \sum_{j=1}^m ||x_{ij} - \hat{x}_i(K, R_i, t_i, X_j)|| \quad (4.12)$$

Dove X_j sono i punti del calibratore e $\hat{x}_i(K, R_i, t_i, X_j)$ la loro retroproiezione sull'immagine i secondo l'equazione $\hat{x}_i = H X_j$. Si indica con n il numero di punti in un'immagine, e con m il numero di immagini considerate. Inoltre, per ridurre il numero dei parametri, solitamente la matrice $R_{3 \times 3}$ viene parametrizzata con un

vettore di 3 elementi corrispondenti a angoli di rotazioni rappresentati in una forma standard quale ad esempio gli angoli di Rodrigues o di Eulero.

Per completare il modello della telecamera bisogna considerare anche i parametri non lineari che si riferiscono alle distorsioni delle lenti. Come descritto in precedenza, tali distorsioni dipendono dalla conformazione della lente e non variano da un'immagine all'altra. Tali parametri possono essere considerati intrinseci. Il problema finale diventa quindi la minimizzazione della funzione:

$$\sum_{i=1}^n \sum_{j=1}^m \|x_{ij} - \hat{x}_i(K, k_1, k_2, k_3, p_1, p_2, R_i, t_i, X_j)\| \quad (4.13)$$

Non sempre è utile inserire tutti i parametri che rappresentano le distorsioni nel modello; alle volte, anzi, è meglio trascurarne alcune, come ad esempio l'ultima radiale, che di solito è meno influente delle altre.

Un volta scelto il set di parametri e la funzione di costo, si passa alla selezione del metodo di minimizzazione. Esistono vari metodi iterativi [8], [16] ed i più utilizzati, descritti in seguito, sono:

- Metodo di Gauss-Newton;
- Metodo a discesa di gradiente;
- Metodo di Levenberg-Marquardt;

Supponendo di avere una relazione funzionale $X = f(P)$, dove X è il vettore delle misure e P è il vettore dei parametri rispettivamente in \mathbb{R}^n e in \mathbb{R}^m , si desidera trovare il vettore \hat{P} che meglio soddisfa tale relazione. Più precisamente si cerca il vettore \hat{P} che soddisfa $X = f(P) - \epsilon$, nel quale viene minimizzato $\|\epsilon\|$. Nel caso f sia una funzione lineare, tale sistema è risolto semplicemente con l'equazione DLT. Se invece è una funzione non-lineare, si deve partire da un valore iniziale dei parametri P_0 per poi procedere al raffinamento della stima sotto l'assunzione che la funzione f sia localmente lineare. Sia $\epsilon_0 = f(P_0) - X$ e si assuma che la funzione f sia approssimata in P_0 da $f(P_0 + \Delta) = f(P_0) + J\Delta$, dove J è una mappatura lineare rappresentata dalla matrice Jacobiana $J = \frac{df}{dP}$.

La ricerca del punto successivo d'iterazione $f(P_1)$, avviene con la funzione $P_1 = P_0 + \Delta$, che minimizza $f(P_1) - X = f(P_0) + J\Delta = \epsilon_0 + J\Delta$. Ovvero si riduce al calcolo di un problema di minimizzazione lineare descritto da $|\epsilon_0 + J\omega\Delta|$ su Δ . Il vettore Δ è ottenuto risolvendo l'equazione normale $\Delta = -J^+\epsilon_0$. Dove J^+ è la pseudo inversa di J . Quindi il vettore \hat{P} è ottenuto iniziando da una P_0 stimato e calcolato ricorsivamente in accordo alla formula:

$$P_{i+1} = P_i + \Delta_i \quad (4.14)$$

dove Δ_i è la soluzione di un problema lineare ai minimi quadrati:

$$J\Delta_i = -\epsilon_i \quad (4.15)$$

La matrice J è lo Jacobiano di df/dP valutato in P_i e $\epsilon_i = f(P_i) - X$. Si ipotizza che tale algoritmo converga alla soluzione dei minimi quadrati richiesta; in alcuni casi è però possibile che le iterazioni convergano ad un minimo locale, o non convergano affatto. Il comportamento dell'algoritmo dipende fortemente dalla stima iniziale P_0 . Si voglia ora cercare il minimo di una funzione di molte variabili. Si considera quindi una generica funzione $g(P)$; l'obbiettivo diventa ora la minimizzazione di tale funzione rispetto tutte le variabili che compongono P . Si fanno però due assunzioni: che $g(P)$ abbia un valore minimo ben definito e che sia conosciuto un punto P_0 ragionevolmente vicino al minimo. Espandendo in serie di Taylor la funzione rispetto P_0 si ottiene:

$$g(P_0 + \Delta) = g + g'\Delta + \Delta^T g'' \frac{\Delta}{2} + \dots \quad (4.16)$$

Si indicano con g' le derivate prime della funzione e con g'' le derivate seconde. Si desidera minimizzare questa quantità in rispetto a Δ . A questo fine, si può derivare rispetto a Δ e imporre la derivata a 0, arrivando quindi all'equazione $g + g'\Delta = 0$ oppure:

$$g''\Delta = g' \quad (4.17)$$

In questa equazione g'' è la matrice delle derivate seconde, chiamata Hessiano di g , essa è composta da termini del tipo $d^2g/dp_i dp_j$, con i e j corrispondenti ai i -esimo e j -esimo componente del vettore dei parametri. Il vettore g' è il gradiente di g . L'equazione 4.17 è quella utilizzata dal *metodo di Newton* come parametro d'iterazione. Il problema principale di tale procedura sta nella difficoltà del calcolo dell'Hessiano; una soluzione a tale problema deriva dalla semplificazione introdotta dal *metodo di Gauss-Newton* [8], [16], nel quale l'Hessiano viene calcolato in maniera approssimata come:

$$g'' = J^T J \quad (4.18)$$

ovvero riscrivendo la (4.17):

$$J^T J \Delta = -J^T \epsilon \quad (4.19)$$

Una variante a tale metodo deriva dalla possibilità di utilizzare un passo di incremento variabile. Se si considera, infatti, che $-g' = -J^T \epsilon$ definisce la direzione di più rapida discesa della funzione costo, una strategia di minimizzazione consiste nel muoversi iterativamente nella direzione del gradiente. La lunghezza del passo può essere calcolata ricercando lungo la linea definita dalla direzione negativa del gradiente. In questo caso il parametro di incremento Δ viene calcolato dall'equazione $\lambda \Delta = -g'$, dove λ controlla la lunghezza dello step. Tale metodologia è chiamata *a discesa di gradiente* [8], [16], e può essere interpretata come un metodo di Newton nel quale l'Hessiano è stato approssimato da una matrice scalare λI . La discesa di gradiente solitamente non è una buona strategia di minimizzazione, poiché tipicamente caratterizzata da una bassa convergenza dovuta a spostamenti “a zigzag”.

Il *metodo di Levenberg-Marquardt (LM)* [8], [16], [18], [19] può essere considerato come un'interpolazione tra il metodo di Gauss-Newton e la discesa di gradiente. L'equazione 4.19 viene sostituita da un'equazione normale aumentata :

$$(J^T J + \lambda I) \Delta = -J^T \epsilon \quad (4.20)$$

Il valore di λ varia tra un'iterazione e l'altra e assume tipicamente valori di circa un millesimo della media della diagonale di $J^T J$. Il valore d'incremento Δ ottenuto risolvendo l'equazione aumentata viene ridotto in caso di aumento dell'errore, altrimenti esso viene accettato e λ viene diviso per un fattore (solitamente 10) prima della successiva iterazione. In altre parole, se Δ porta a un incremento dell'errore, allora λ è incrementato dello stesso fattore e l'equazione aumentata viene risolta nuovamente; tale processo continua finché viene individuato un Δ in grado di diminuire l'errore. Per meglio comprendere tale metodo si può osservare come per valori molto piccoli di λ questo si riduce essenzialmente al metodo di Gauss-Newton, in grado di convergere velocemente alla soluzione. Viceversa per valori grandi di λ l'equazione può essere approssimata da $\lambda \Delta = -J^T \epsilon$, corrispondente al metodo a discesa di gradiente, che garantisce una migliore convergenza in situazioni difficili. Tra tutti questi metodi, il più indicato per la risoluzione del problema di ottimizzazione non lineare dei parametri di calibrazione risulta quindi essere il metodo di Levenberg-Marquardt. Esso, pur essendo poco più lento rispetto agli altri due, è sicuramente più robusto; infatti, rispetto agli altri, anche partendo da parametri iniziali leggermente più lontani dal minimo globale, esso ha più possibilità di convergere alla soluzione corretta [8].

4.2 La tecnica di Heikkilä

Tra le diverse tecniche esistenti in letteratura si vuole in questo paragrafo approfondire la tecnica presentata da Janne Heikkilä [1]. Questa tecnica utilizza un calibratore planare con alcuni cerchi (*blob*) in posizione nota e di raggio noto. L'innovazione di tale tecnica sta in una semplificazione del modello non lineare che permette di invertire il modello delle distorsioni senza ricorrere a metodi ricorsivi. Heikkilä inoltre propone un metodo innovativo per identificare il centro dei blob del calibratore con maggiore accuratezza, eliminando gli errori sistematici.

La tecnica di Heikkilä utilizza un modello pinhole come quello descritto nel paragrafo 3.2 che considera i parametri non lineare di distorsione, modellizzati

come descritto al paragrafo 3.3. Non esiste dunque una soluzione analitica che consenta di invertire il modello completo (equazioni 3.13). Heikkilä propone dunque alcune semplificazioni da apportare a tale modello che consentano di renderlo invertibile. In particolare riscrivendo la 3.13 come segue:

$$a_c = a_d + \mathcal{F}_D(a_c, \delta) \quad (4.21)$$

dove:

$$a_c = x_d$$

$$a_d = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) x_n$$

$$\mathcal{F}_D(a_c, \delta) = dx \text{ (descritto dall'equazione 3.14)}$$

$\delta = [k_1, k_2, \dots, p_1, p_2, \dots]$ coefficienti per le distorsioni radiali e tangenziali

Sviluppando l'equazione 4.21 in serie di Taylor, si ottiene:

$$a_c \approx a_d + \mathcal{F}_D(a_c, \delta) + D(a_c)(a_d - a_c) \quad (4.22)$$

con:

$$D(a_c) = \left[\frac{\partial}{\partial u} \mathcal{F}_D(a, \delta) \frac{\partial}{\partial v} \mathcal{F}_D(a, \delta) \right] \Big|_{a=a_c} \quad (4.23)$$

Risolvendo a_d dall'equazione 4.22 si trova:

$$a_d \approx a_c - (I + D(a_c))^{-1} \mathcal{F}_D(a_c, \delta) \quad (4.24)$$

Gli elementi di $D(a_c)$ sono molto piccoli ($\ll 1$); quindi è possibile approssimare come segue:

$$a_d \approx a_c - \frac{1}{d_{11}(a_c) + d_{11}(a_c) + 1} \mathcal{F}_D(a_c, \delta) = a_c - \mathcal{F}_D^*(a_c, \delta) \quad (4.25)$$

dove $d_{11}(a_c)$ e $d_{11}(a_c)$ sono rispettivamente l'elemento in alto a sinistra e in basso a destra di $D(a_c)$. Ad esempio se si volesse utilizzare solamente due coefficienti per le distorsioni radiali e due per quelle tangenziali il modello delle distorsioni invertito sarebbe:

$$\mathcal{F}_D^*(a_c, \delta) = \frac{1}{4k_1 r_c^2 + 6k_2 r_c^4 + 8p_1 \bar{v}_c + 8p_2 \bar{u}_c + 1} \mathcal{F}_D(a_c, \delta) \quad (4.26)$$

dove $r_c = \sqrt{\bar{v}_c^2 + \bar{u}_c^2}$.

Sostituendo $u = \bar{u}_c$ e $v = \bar{v}_c$ nell'equazione che descrive il modello della telecamera, ad esempio, riconducendoci all'equazione 3.2:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \propto \begin{bmatrix} \lambda u \\ \lambda v \\ 1 \end{bmatrix} = F \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = PM \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.27)$$

con $M \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$, si trova il modello diretto della telecamera che consente di convertire le coordinate 3D in coordinate immagine distorte. Utilizzando il modello inverso della telecamera invece è possibile trasformare le coordinate distorte in raggi di vista nello spazio 3D o trovare l'intersezione di questi raggi con un piano 2D conosciuto. Ad esempio si prenda un piano 2D Π' nel sistema di riferimento H, le cui origini sono in $h_0 = [X_0, Y_0, Z_0]^T$ e definito dai vettori 3D $h_1 = [X_1, Y_1, Z_1]^T$ e $h_2 = [X_2, Y_2, Z_2]^T$. La trasformazione dalle coordinate immagine corrette $a_c = [u_c, v_c]^T$ prodotta dall'equazione 4.21 sul piano Π' può essere espressa da:

$$\lambda \begin{bmatrix} x_H \\ 1 \end{bmatrix} = (FH)^{-1} \begin{bmatrix} a_c \\ 1 \end{bmatrix} \quad (4.28)$$

dove λ è un fattore di scala, $x_H = [X_H, Y_H]^T$ sono le coordinate retroproiettate in H e $H = \begin{bmatrix} h_1 & h_2 & h_3 \\ 0 & 0 & 1 \end{bmatrix}$.

A causa delle approssimazioni fatte, \mathcal{F}_D^* non rappresenta esattamente la funzione inversa di \mathcal{F}_D . Bisognerebbe utilizzare dei parametri leggermente diversi δ' nell'equazione 4.26. Una tecnica per stimare questi parametri è presentata in [1].

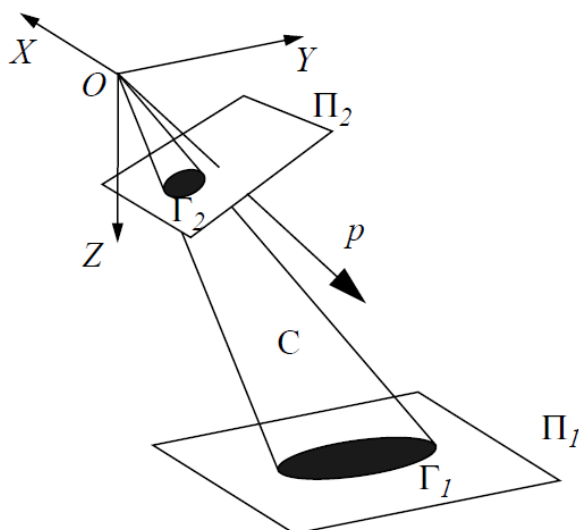


Figura 4-3: Deformazione prospettica di un marcatore circolare.

Oltre a delineare un modello invertibile che considera le distorsioni non lineari, Heikkilä propone un metodo per stimare il centro dei marcatori circolari di un calibratore con ottima accuratezza, evitando gli errori sistematici che si commettono nelle procedure tradizionali. Osservando infatti un cerchio con una telecamera la prospettiva fa sì che a seconda dell'inclinazione del piano sul quale è disegnato il cerchio rispetto al piano immagine, esso venga visto nell'immagine come un'ellisse. Uno schema esemplificativo di questo fenomeno è rappresentato in figura 4.3. In questi casi, utilizzando i normali algoritmi di blob detection per trovare il centro del marcatore circolare, si avrebbero degli errori sistematici che potrebbero peggiorare l'accuratezza della calibrazione. Inoltre la proiezione del cerchio sull'immagine potrebbe essere affetta da distorsioni radiali e tangenziali della quale bisogna tener conto se si vuole individuare con buona accuratezza il centro del marcatore. Si consideri di avere un marcatore circolare R con raggio r posto nel piano Π' in modo che il suo centro è nel centro del sistema di riferimento H . Un cerchio può essere rappresentato dalla generica curva quadratica della seguente equazione:

$$AX_H^2 + 2BX_HY_H + CY_H^2 + 2DX_H + 2EY_H + F = 0 \quad (4.29)$$

Dove A, B, C, D, E e F sono coefficienti che definiscono la forma e la posizione della curva quadratica. In coordinate omogenee la curva può essere espressa da:

$$\begin{bmatrix} x_H \\ 1 \end{bmatrix}^T Q \begin{bmatrix} x_H \\ 1 \end{bmatrix} = 0 \quad (4.30)$$

dove: $Q = \begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix}$.

Per quanto riguarda un cerchio: $Q = \begin{bmatrix} -\frac{1}{r^2} & 0 & 0 \\ 0 & -\frac{1}{r^2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Utilizzando l'equazione 4.28, è possibile proiettare R sul piano immagine, ottenendo la seguente curva:

$$\begin{bmatrix} a_c \\ 1 \end{bmatrix}^T S \begin{bmatrix} a_c \\ 1 \end{bmatrix} = 0 \quad (4.31)$$

dove:

$$S = ((FH)^{-1})^T Q (FH)^{-1} \quad (4.32)$$

Dall'equazione si deduce che il risultato è una curva quadratica che può essere interpretata geometricamente come un cerchio, un'iperbole, una parabola o un'ellisse. In pratica però il cerchio proiettato sul piano immagine può essere un'ellisse o, in casi particolari, un cerchio. Il centro dell'ellisse $e_c = [u_e, v_e]^T$ può essere trovato come segue:

$$\lambda \begin{bmatrix} e_c \\ 1 \end{bmatrix} = S^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (4.33)$$

Combinando poi l'equazione 4.32 con la 4.33 si ottiene:

$$\lambda \begin{bmatrix} e_c \\ 1 \end{bmatrix} = FGf_3 \quad (4.34)$$

dove $G = HQ^{-1}H^T$ e f_3 è il vettore che rappresenta l'ultima riga di F. La matrice G identifica la geometria, la posizione e l'orientamento del cerchio. La trasformazione prospettica F identifica la mappatura dalle coordinate mondo alle coordinate immagine non distorte, che possono essere convertite in coordinate immagini distorte utilizzando l'equazione 4.25:

$$e_d = e_c - \mathcal{F}_D(e_c, \delta) \quad (4.35)$$

Dunque utilizzando l'equazione 4.34 invece che la 4.27, si ottiene una relazione tra i centri osservati delle ellissi e il modello della telecamera che compensa gli errori sistematici.

La tecnica di Heikkilä è utilizzata nelle librerie commerciali per la visione artificiale HALCON, sviluppate da MVTec [33] e in uso presso il VBLab da pochi anni. Questo applicativo consente di eseguire calcoli e manipolazioni su immagini, utilizzando un proprio linguaggio nell'ambiente di sviluppo HDevelop. In esso vi sono moltissime funzioni, che consentono di effettuare l'intera procedura di calibrazione. In alternativa, è possibile utilizzare un tool molto pratico e immediato, chiamato *Calibration Assistant*. In esso si ha la possibilità di effettuare una calibrazione completa, vale a dire con tutti i parametri da stimare, oppure fissarne alcuni a un valore prestabilito e stimare gli altri. Per ogni parametro che viene stimato durante la calibrazione si lascia la possibilità di inicializzarlo o meno a un valore definito dall'utente. In caso di calibrazione completa è comunque necessario impostare un valore iniziale di lunghezza focale e dimensioni della cella (pixel). Mentre la focale e la componente orizzontale del pixel sono solamente valori da cui partire per effettuare la stima, il valore che esprime la componente verticale del pixel rimane fisso durante la calibrazione e non viene stimato. Questo è uno dei principali svantaggi della tecnica di HALCON, ovvero dover conoscere con esattezza le dimensioni del pixel. Questa informazione è facilmente ricavabile da datasheet nel caso in cui si sta utilizzando telecamere industriali, mentre è più difficile da ricavare nel caso si utilizzi ad esempio una webcam. Se non si conosce con esattezza le dimensioni dei pixel che compongono il sensore la tecnica implementata nelle librerie HALCON non può essere utilizzata.

Selezionando una calibrazione completa, è possibile scegliere due modelli di telecamera: uno chiamato *divisionale* e uno *polinomiale*. Il primo ha come parametri l'ampiezza dei pixel nelle due direzioni (impostato dall'utente), la lunghezza focale, la posizione del centro ottico (inicializzata nel centro del sensore) e un parametro che tiene conto delle distorsioni radiali *Kappa*. Il modello polinomiale è invece più complesso e utilizza cinque parametri che rappresentano le distorsioni (invece di uno solo, *Kappa*), tre per le distorsioni radiali (K_1 , K_2 , K_3)

e due per le tangenziale (P1 e P2). Da numerose prove ed esperimenti si è notato che l'utilizzo del secondo modello porta spesso a errori di retroproiezioni minore e molto frequentemente a calibrazioni più accurate.

Una volta selezionato il modello e inizializzati correttamente i parametri, si passa all'acquisizione delle immagini. HALCON richiede che le immagini siano prese in maniera adeguata seguendo una serie di regole:

1. Il calibratore deve essere abbastanza grande da coprire almeno un quarto dell'immagine.
2. Il diametro minimo dei blob deve essere almeno 10 pixel.
3. Usare un'illuminazione che consenta di lasciare lo sfondo molto più scuro rispetto al calibratore (il contrasto tra uno e l'altro deve essere almeno 100 valori di grigi).
4. Il valore di grigio della parte bianca del calibratore deve essere almeno 100.
5. L'illuminazione deve assicurare che il calibratore sia rappresentato da valori di grigi omogenei.
6. L'immagine non deve essere sovraesposta, vale a dire non devono essere presenti valori di grigi maggiori di 255.
7. Il set di immagini deve coprire l'intero campo di visione.
8. Bisogna utilizzare varie pose del calibratore.
9. Per una buona calibrazione è necessario acquisire un numero di immagini compreso tra 15 e 20; per poter calibrare sono richieste comunque almeno 10 immagini.

L'acquisizione delle immagini può essere fatta in due modi. Il primo consiste nell'acquisire una serie di immagini (di solito almeno venti o trenta) a monte della calibrazione, utilizzando HALCON o qualsiasi programma che consente di eseguire un'istantanea dal dispositivo. Le immagini vengono poi caricate nel Calibration Assistant, che effettua un controllo qualità per ognuna di esse e in maniera automatica trova le correlazioni tra i punti nelle immagini e il modello del calibratore. Grazie a questo è possibile controllare che le immagini siano state acquisite in modo adeguato, così da poter eliminare quelle meno adatte o aggiungerne alcune acquisite in modo più accurato. Un esempio di controllo qualità

delle immagini è mostrato in figura 4.4 e 4.5. In particolare nella figura 4.5 è rappresentata l'interfaccia del Calibration Assistant che effettua il controllo qualità sulle immagini caricate. In questo caso è selezionata quella mostrata in figura 4.4 e il controllo è eseguito su di essa. Grazie alla verifica della qualità possiamo vedere che l'immagine presenta un contrasto basso e dunque non è un'immagine buona per la calibrazione. Il secondo metodo di acquisizione consiste nel visualizzare il calibratore in modalità live tramite il dispositivo; in real-time HALCON effettua il controllo della qualità dell'immagine e trova le correlazioni tra i punti e il modello del calibratore. In questo modo è possibile acquisire un'istantanea ogni volta che il controllo qualità è passato ed è stata trovata la corrispondenza tra il modello del calibratore e i punti nell'immagine. Entrambi i metodi dunque permettono anche a un utente poco esperto in calibrazione di non commettere errori in fase di acquisizione delle immagini, che possono compromettere la calibrazione e dunque l'intera misurazione.

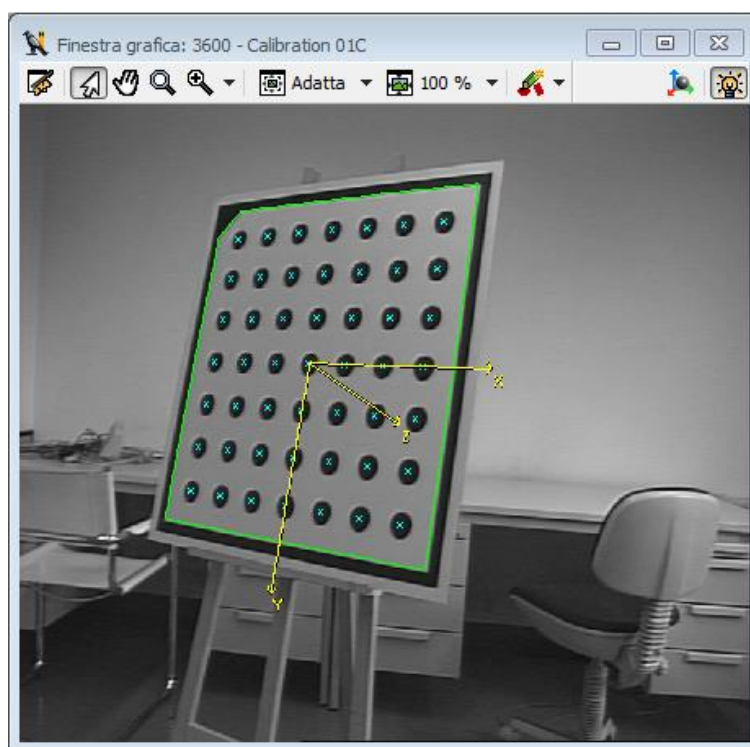


Figura 4-4: Esempio di immagine di calibrazione nel quale è stata trovata la corrispondenza tra il modello di calibratore e i punti nell'immagine. Come si vede dal controllo qualità (figura 4.5) l'immagine ha però il contrasto basso e dunque non è un'immagine buona per calibrare.

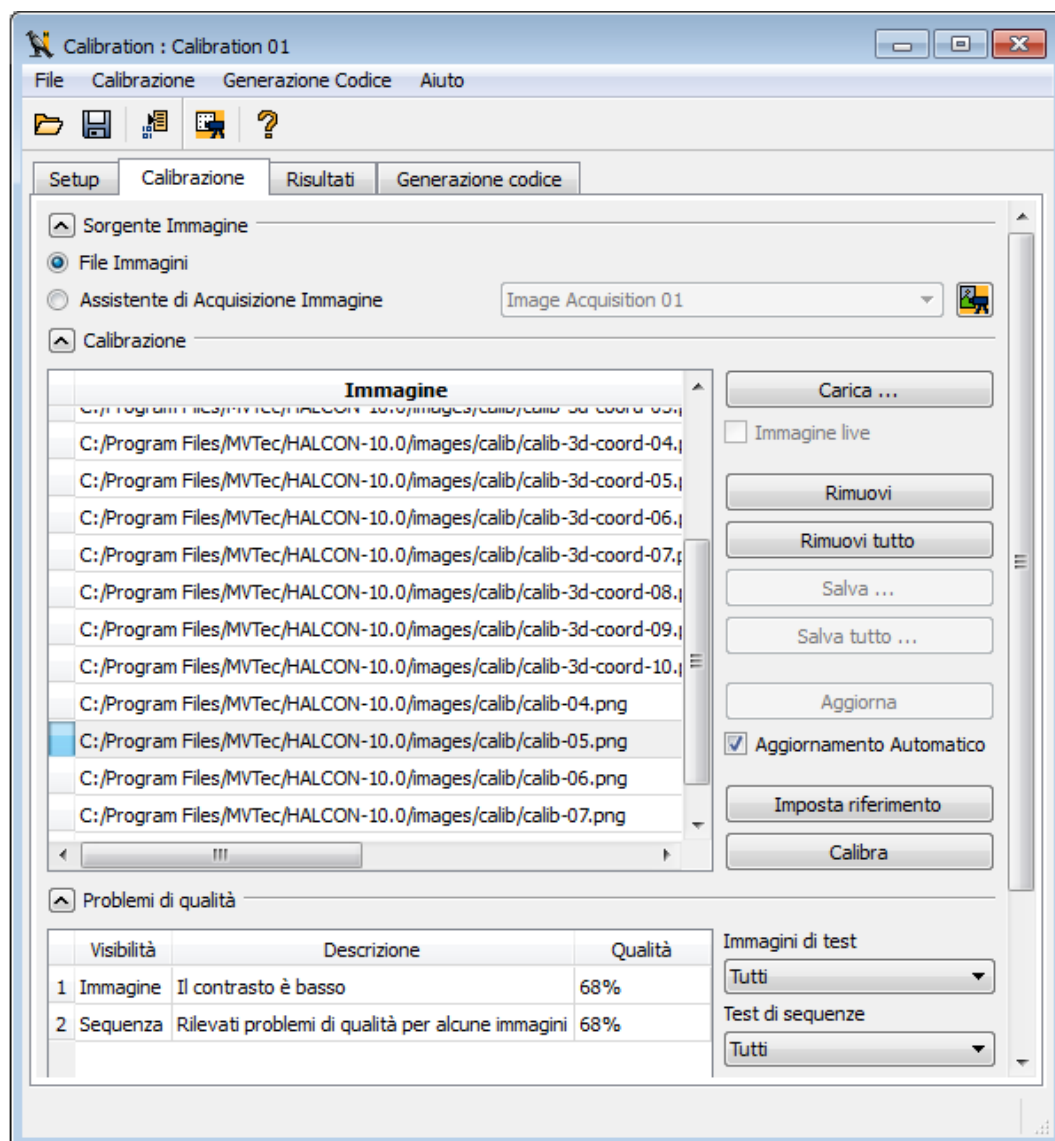


Figura 4-5: *Interfaccia di controllo della qualità delle immagini. In questo caso è selezionata l'immagine di figura 4.4, in cui il contrasto risulta essere basso.*

Una volta acquisito il set di immagini, è possibile procedere con la calibrazione. Bisogna solamente selezionare quale immagine usare come sistema di riferimento; esso viene posizionato nel centro del calibratore. I risultati sono poi mostrati in una schermata come quella di figura 4.6 (paragrafo seguente), in cui in alto è mostrato l'errore medio, vale a dire l'errore di retroproiezione medio, e l'elenco dei parametri intrinseci, delle distorsioni e infine i parametri estrinseci (“Posa della telecamera”).

4.2.1 Confronto con la tecnica di Zhang-LM

Si analizzano i risultati di diverse calibrazioni per poter confrontare il metodo Zhang-LM, con la tecnica di Heikkilä, implementata nelle librerie commerciali HALCON. Lo scopo di tale confronto è lo studio delle prestazioni dell'approccio utilizzato da tali librerie al problema della calibrazione, al fine di verificare se esistono soluzioni alternative al metodo Zhang-LM, attualmente utilizzato in laboratorio, altrettanto robuste e accurate.

Come modello di telecamera si utilizza sempre il modello *pinhole camera*. La parola "modello" nel seguito non indicherà i differenti modelli di telecamera (*pinhole camera*, *affine camera*, etc...), bensì i vari tipi di modelli *pinhole camera* a diverso numero di parametri (ad esempio lineare a 4 parametri o non lineare a 8 parametri, etc...). Si ricorda che le librerie HALCON consentono di selezionare solamente due diversi modelli di telecamere: *divisionale* e *polinomiale*. Il primo considera solamente un parametro di distorsione, mentre il secondo considera 3 parametri di distorsione radiale e due di distorsione tangenziale. L'algoritmo di Zhang-LM implementato in laboratorio invece consente di selezionare diversi modelli a seconda del numero di parametri intrinseci e di parametri di distorsione scelti. Per il confronto ci si limiterà ad utilizzare il modello a 4 parametri ($u_0, v_0, f, K1$) o il modello a 8 parametri ($u_0, v_0, f, K1, K2, P1, P2, K3$). Sono state eseguite diverse prove allo scopo di verificare che i dati della procedura di Heikkilä diano risultati compatibili con quelli di Zhang-LM.

A titolo di esempio si riporta nel seguito una prova di confronto effettuata con una telecamera Dalsa GENIE serie hm1400, con ottica da 13 mm. Essendo l'ottica piuttosto corta con buona probabilità è conveniente tener conto delle distorsioni e dunque il modello di riferimento sarà quello polinomiale, per quanto riguarda la tecnica delle librerie HALCON. Sono comunque state eseguite due calibrazioni, una con modello divisionale e una polinomiale. Per quanto riguarda la tecnica di Zhang-LM, volendo utilizzare un modello simile a quello polinomiale e visto che dal datasheet della telecamera è possibile verificare che i pixel sono quadrati, viene utilizzato un modello a 8 parametri, imponendo il parametro di skew nullo, le

lunghezze focali nelle due direzioni del sensore uguali e considerando tutte le componenti di distorsione. Le dimensioni dei pixel da datasheet sono $7.4 \times 7.4 \mu\text{m}$ e la risoluzione $1400 \times 1024 \text{ pixel}$. L'ottica è tenuta fissa 13 mm . Dai risultati ci si aspetta quindi che le coordinate del centro ottico siano vicine al centro del sensore $u_0 \simeq 700 \text{ pixel}$ e $v_0 \simeq 512 \text{ pixel}$. Inoltre deve risultare che $f_{pix} \simeq \frac{13 \text{ mm}}{7.4 \mu\text{m}} = 1756.75$. Sono state utilizzate 15 immagini per la calibrazione.

Nella figura 4.6 e 4.7 sono rappresentati i risultati del problema di calibrazione con le librerie HALCON rispettivamente nel caso di modello divisionale e polinomiale. Nella tabella 4.1 invece sono riportati i risultati principali delle calibrazioni effettuate rispettivamente utilizzando la tecnica di Zhang lineare, Zhang-LM e le librerie HALCON con modello divisionale e polinomiale.

	ZHANG	ZHANG- LM	HALCON- DIVIS.	HALCON- POLIN.
Tempo [s]	0.3	2	1.5	2.5
Err. Ret. [pix]	1.270	0.066	0.107	0.035
Focale [mm]	12.83	12.97	13.00	12.94
u0 [pixel]	685.45	698.8	683.62	702.52
v0[pixel]	532.98	527.05	533.15	527.64

Tabella 4.1: Confronto tra i principali risultati delle calibrazioni utilizzando il metodo di Zhang lineare, Zhang-Lm, HALCON divisionale e HALCON polinomiale; 15 immagini, dimensione dei pixel $7.4 \mu\text{m}$, risoluzione $1400 \times 1024 \text{ pixel}$, ottica da 13 mm .

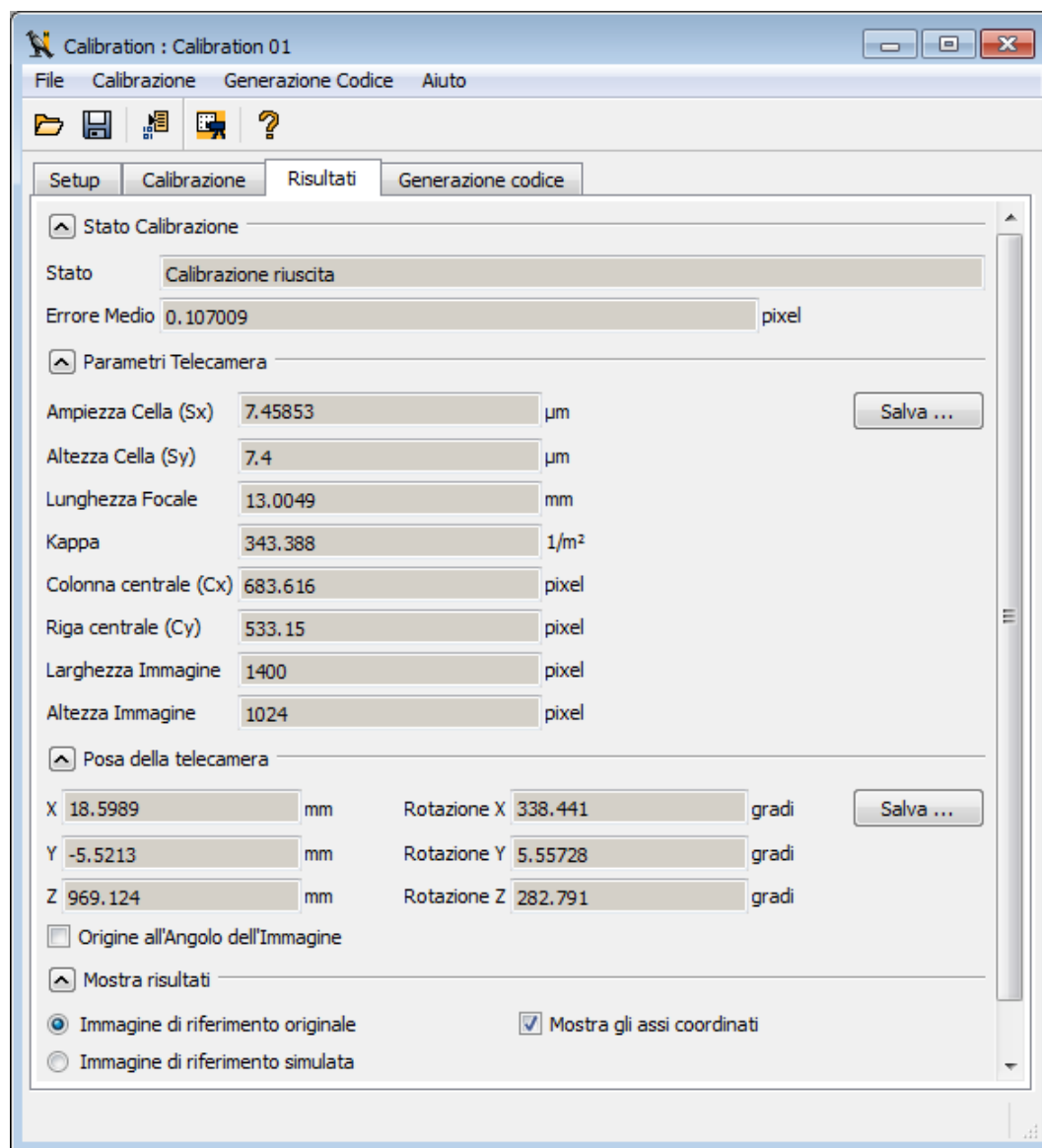


Figura 4-6: Risultati di calibrazione le librerie HALCON con un modello divisionale, 15 immagini, dimensione dei pixel 7.4 μm , risoluzione 1400x1024 pixel, ottica da 13 mm.

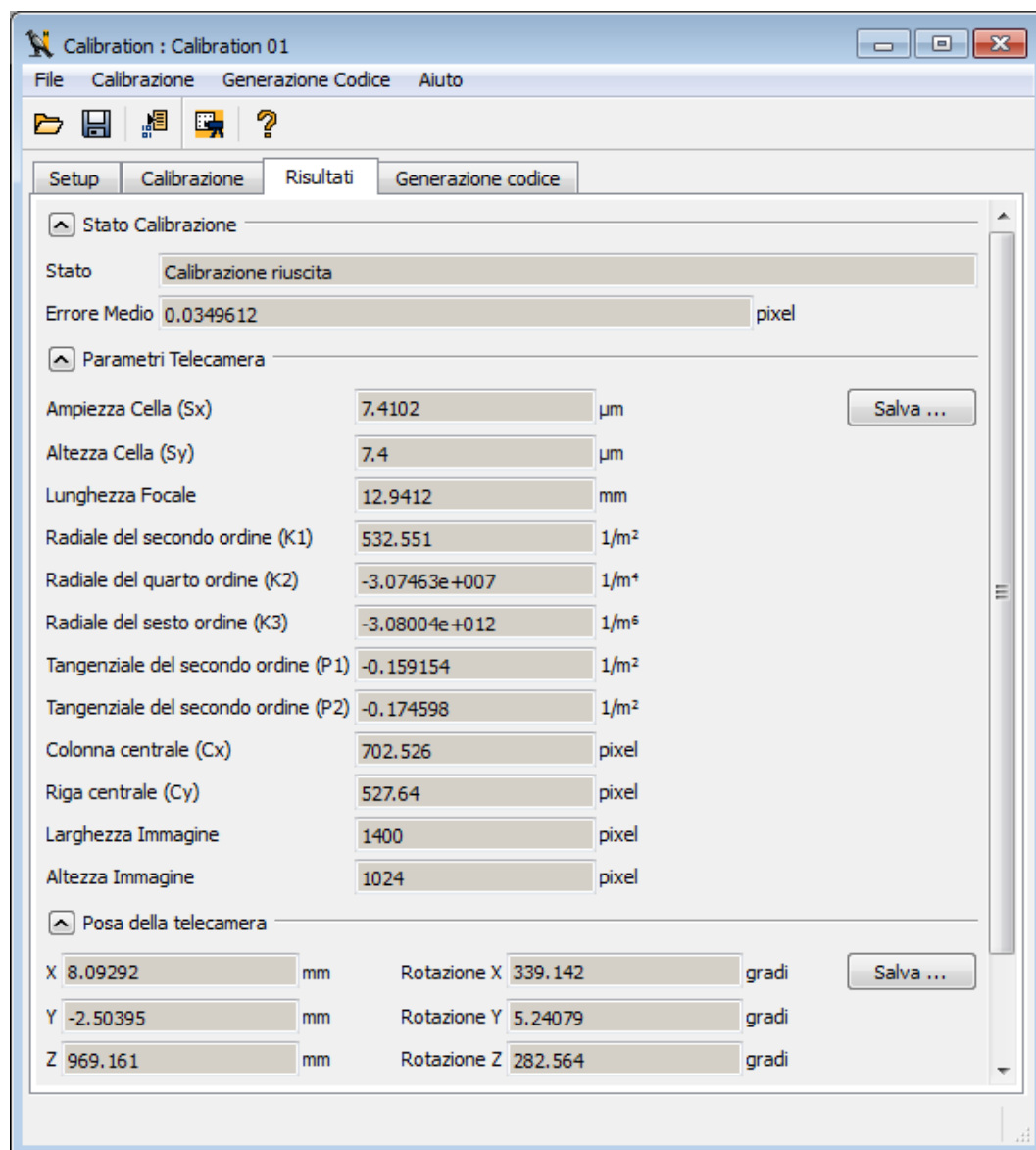


Figura 4-7: Risultati di calibrazione utilizzando le librerie HALCON con un modello polinomiale, 15 immagini, dimensione dei pixel 7.4 μm, risoluzione 1400x1024 pixel, ottica da 13 mm.

Da un confronto tra le diverse tecniche si nota che i risultati di Zhang-LM ottimizzano quelli di Zhang, restituendo un errore di retroproiezione notevolmente inferiore e parametri intrinseci più vicini ai valori attesi. Per quanto riguarda le tecniche delle librerie HALCON, nel caso di modello divisionale l'errore di retroproiezione è molto vicino a quello trovato con Zhang-LM ma leggermente più alto. Anche i parametri intrinseci sono compatibili. Nel caso di modello polinomiale invece i risultati si discostano di pochissimo, trovando un errore di retroproiezione addirittura inferiore e parametri intrinseci molto vicini sia a quelli attesi che a quelli trovati in Zhang-LM. Per quanto riguarda i tempi di esecuzione dell'algoritmo di calibrazione di HALCON, essi sono compatibili con i tempi impiegati dall'algoritmo di Zhang-LM (pochi secondi) e non rappresentano dunque un problema per la calibrazione.

Da questa e altre prove simili effettuate sui diversi algoritmi di calibrazione si può dedurre che le prestazioni dell'algoritmo di Heikkilä contenuto nelle librerie HALCON sono compatibili a quelle dell'algoritmo di Zhang-LM utilizzato come riferimento in letteratura e presso il laboratorio VBLab.

Riassumendo, si può osservare che le due tecniche hanno molte caratteristiche in comune. Entrambe infatti permettono di calibrare senza dover conoscere la posizione del calibratore, poiché il sistema di riferimento mondo è solidale con esso. Questo semplifica notevolmente la procedura di setup di calibrazione, complicando però il problema, poiché i parametri estrinseci variano per ogni immagine acquisita. Si ricorda che bisogna prestare molta attenzione alla fase di acquisizione delle immagini, poiché è necessario che il calibratore venga posto in pose più differenziate possibile, coprendo tutto il sensore.

Il metodo di Zhang richiede almeno l'uso di 3 immagini, anche se con LM si ottengono risultati accurati, specialmente se il modello utilizzato è complesso, con almeno 15 immagini. L'algoritmo di Levenberg-Marquardt converge velocemente a una soluzione ottima solo se inizializzato correttamente, quindi solo se i parametri trovati con la tecnica di Zhang sono vicini alla soluzione ottima. In questo caso l'algoritmo raggiunge accuratissime elevate. Infine è stato deciso di utilizzare un

metodo manuale per trovare le correlazioni fra i punti nelle immagini, utilizzando il software “Feature Linker”. In questo modo si può ottenere un riscontro visivo immediato dei risultati di correlazione, limitando gli errori che possono accadere con procedure automatiche. Chiaramente però un approccio manuale è più lungo e laborioso.

La tecnica contenuta nelle librerie HALCON invece, presenta alcuni vincoli non indifferenti. Per prima cosa per poter calibrare è necessario conoscere la dimensione dei pixel che compongono il sensore con esattezza. Quindi tale tecnica non può sempre essere applicata. Secondariamente HALCON richiede che per calibrare si acquisisca un insieme di immagini con caratteristiche molto esigenti. Il processo di acquisizione delle immagini deve dunque essere fatto con molta attenzione. Questo fattore però può anche essere considerato un vantaggio, in quanto le limitazioni restringenti sulla bontà delle immagini fanno sì che gli errori in fase di acquisizione siano meno frequenti. HALCON limita la scelta del modello di telecamera a due differenti modelli, uno più semplice (chiamato divisionale) e uno più complesso (detto polinomiale). Inoltre HALCON effettua la ricerca delle correlazioni tra il modello del calibratore e i punti nell’immagine in maniera semplice e automatica. Grazie al Calibration Assistant, HALCON si presta ad essere utilizzato anche da utenti meno esperti. Inoltre è stato verificato che l’accuratezza dei risultati di calibrazione calcolati con HALCON è compatibile con quella trovata col metodo di Zhang-LM. Infine HALCON contiene al suo interno delle funzioni che dà la possibilità di tarare in maniera semplice e veloce un sistema stereoscopico a più telecamere. Questi algoritmi consentono di effettuare contemporaneamente la calibrazione delle telecamere che compongono il sistema stereoscopico e la loro registrazione, restituendo un errore di retroproiezione medio tra tutte le telecamere, parametro utilizzato per definire la bontà della calibrazione.

Data la semplicità di utilizzo, la comodità nell’acquisizione delle immagini di calibrazione, la possibilità di utilizzo da parte di utenti meno esperti e soprattutto la compatibilità delle prestazioni di calibrazione con l’algoritmo di Zhang-LM, verranno utilizzate le librerie HALCON per effettuare le calibrazioni descritte nel proseguo del lavoro.

5. Attività sperimentale

Lo scopo di questo capitolo è quello di effettuare una qualificazione metrologica in generale della stereoscopia. Per fare ciò è stato realizzato un sistema di misura a tre telecamere. Per prima cosa viene descritto il set up utilizzato nella campagna sperimentale e il funzionamento dei dispositivi che compongono il sistema di misura con le motivazioni che hanno portato alla loro scelta. Nel seguito vengono descritte nel dettaglio le prove sperimentali che sono state effettuate per caratterizzare metrologicamente il sistema di misura e, da ultimo, vengono riportati i risultati ottenuti. In particolare si è voluto analizzare il legame tra l'errore di retroproiezione 2D medio, indice di bontà della calibrazione comunemente utilizzato, e l'accuratezza delle misure effettuate. Inoltre si è voluto indagare l'influenza del numero di immagini sull'accuratezza della calibrazione, arrivando a definire un numero minimo consigliato. Si è poi analizzato quanto l'esecuzione di misure fuori dall'area di calibrazione comportasse un degrado dell'accuratezza del sistema. Infine è stata sviluppata un'analisi della propagazione dell'incertezza in fase di calibrazione mediante il metodo di Monte Carlo e un'analisi di sensitività della procedura di ricostruzione tridimensionale.

5.1 Setup del sistema

Per quanto riguarda l'attività sperimentale è stato realizzato un sistema ottico stereoscopico a tre telecamere. In commercio esistono numerose tipologie di telecamere digitali. Alcuni parametri e caratteristiche comuni tra i vari modelli sono descritti nel paragrafo 5.1.1. La scelta del dispositivo da utilizzare non è banale e dipende fortemente dalle applicazioni. Il criterio di scelta delle telecamere con caratteristiche più adatte al caso in esame è descritta nel paragrafo 5.1.2. Alle telecamere sono stati poi aggiunti altri componenti hardware e è stato sviluppato un applicativo software per la taratura del sistema e per l'acquisizione delle misure.

5.1.1 Parametri delle telecamere

Le telecamere digitali sono sistemi complessi e, nella maggior parte di esse, è possibile agire direttamente su determinate caratteristiche, fra cui le più importanti sono:

- Frequenza di acquisizione (o *frame rate*)
- Modalità di acquisizione (sincrona/asincrona/triggered)
- Definizione di ROI direttamente in fase di acquisizione
- Tempo di esposizione (o *shutter*)
- Formato e compressione delle immagini

Tutti questi parametri devono essere presi in considerazione per la corretta scelta della telecamera da utilizzare nell'applicazione di interesse. Altre caratteristiche importanti per tale scelta sono la tecnologia del sensore (CCD o CMOS), le sue dimensioni e risoluzione e la modalità del trigger. In un sistema di misura con telecamere bisogna ricordare che la frequenza di acquisizione è data quasi esclusivamente dal tempo di acquisizione (più un ritardo dovuto all'elettronica interna del dispositivo). Telecamere con risoluzioni basse, oltre che costare meno, portano solitamente dei vantaggi in termini di tempi di acquisizione. Chiaramente con basse risoluzioni diminuisce l'accuratezza in fase di misura; è dunque necessario trovare un compromesso tra le due caratteristiche.

L'utilizzo di un sistema triggerato permette di acquisire immagini sincrone da diversi dispositivi, *on demand* o con una precisa frequenza di acquisizione, evitando che esse acquisiscano in istanti non sincroni. In caso non si possa utilizzare un trigger per l'acquisizione, è possibile lasciare il dispositivo in modalità *free run* e impostare la frequenza di acquisizione mediante il parametro *frame rate*. Così facendo però non è possibile sincronizzare più dispositivi tra loro. Nel caso in esame il sistema è composto da più telecamere e in alcuni casi applicativi anche da illuminatori (vedi capitolo 6), ed è quindi molto importante utilizzare un sistema triggerato, poiché le misurazioni vanno eseguite su immagini acquisite in maniera perfettamente sincrona. Le telecamere in esame possono essere comandate sia da un trigger software che da un trigger esterno. Mentre il vantaggio di un segnale generato via software è chiaramente la semplicità, i vantaggi di un segnale esterno generato da hardware sono una maggiore accuratezza e la possibilità di sincronizzare altri dispositivi, come ad esempio illuminatori. Nel nostro caso il segnale di trigger viene generato da un generatore di funzione Velleman PCSU1000 (paragrafo 5.1.3) e viene inviato a tutti i dispositivi connessi, ovvero telecamere e eventuali illuminatori. Anche nel caso in cui il sistema sia triggerato, questo tipo di telecamere possono lavorare in maniera sincrona o asincrona. Nel caso di acquisizione triggerata sincrona (il cui diagramma di acquisizione è rappresentato in figura 5.1) l'arrivo del trigger non fa scattare l'acquisizione che è invece legata al successivo trigger interno alla telecamera. Si osserva inoltre come nuovi impulsi di trigger non abbiano effetto se giungono durante il tempo di esposizione. Il caso asincrono è solitamente quello più usato poiché permette di lasciare al trigger il compito di coordinare l'intero processo di acquisizione. In figura 5.2 è mostrato il diagramma temporale di tale schema: l'esposizione può iniziare dopo un ritardo impostabile dall'arrivo del trigger per permettere eventualmente di sincronizzare l'acquisizione con illuminatori o altri dispositivi.

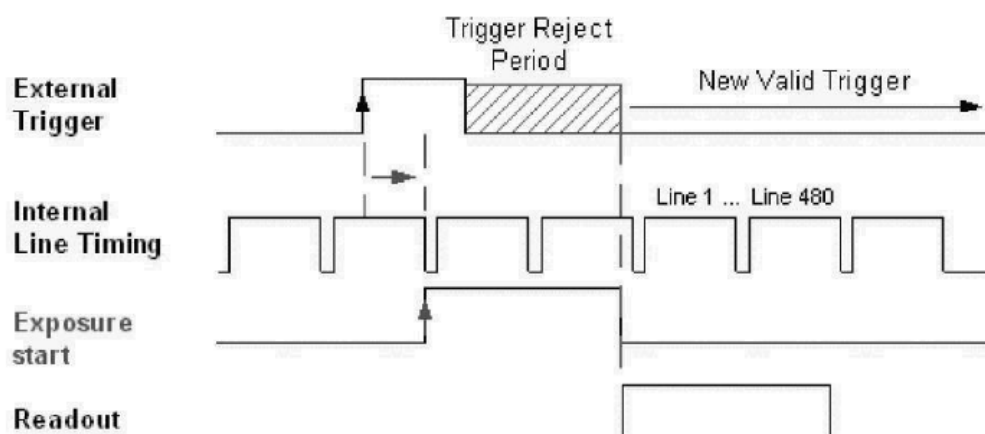


Figura 5-1: Diagramma temporale di un'acquisizione triggerata sincrona.

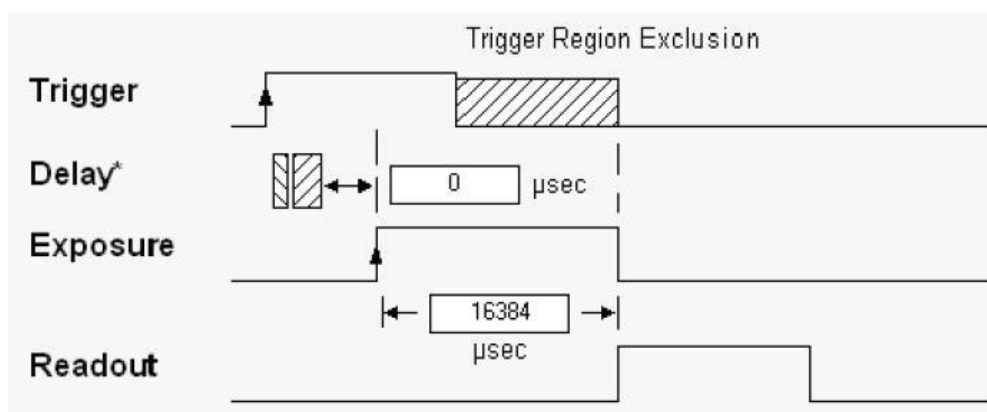


Figura 5-2: Diagramma temporale di un'acquisizione triggerata asincrona.

Il tempo di esposizione (o shutter), vale a dire è il tempo durante il quale l'otturatore della telecamera rimane aperto per permettere alla luce di raggiungere il sensore, è il principale fattore che limita superiormente la frequenza di acquisizione della telecamera. Questo parametro influenza la luminosità dell'immagine acquisita; aumentando il tempo di esposizione l'immagine è più luminosa. Tale caratteristica non è contemplata dal modello *pinhole* (paragrafo 3.2); in esso infatti il passaggio della luce attraverso un foro puntiforme le garantisce una corretta messa a fuoco. Nella realtà invece per ogni impostazione dell'obiettivo (lunghezza focale e apertura del diaframma fissati), esiste un'unica distanza a cui gli oggetti appaiono perfettamente a fuoco; la nitidezza diminuisce gradualmente

spostandosi in avanti (verso il dispositivo) e indietro. La *profondità di campo* è l'intervallo di distanze davanti e dietro al soggetto in cui la sfocatura è impercettibile o comunque tollerabile.

Un buon metodo per aumentare la frequenza di acquisizione della telecamera è la definizione di una *ROI* (Region of interest) che limita la regione di sensore effettivamente acquisita. In tale modo sia l'acquisizione che le operazioni successive risultano molto più rapide. Chiaramente però si perde volume di misura, riducendo la quantità di dati che possono essere analizzati. Tale possibilità viene fornita solitamente per i soli dispositivi digitali basati sui CMOS, nei quali ogni pixel è dotato dell'elettronica di condizionamento del segnale. Contrariamente, i dispositivi a CCD eseguono la scansione sequenziale di tutte le linee che compongono il sensore. In realtà anche con quest'ultimo tipo di sensore è possibile ricavare una ROI ma con vantaggi notevolmente inferiori in termini di diminuzione di informazioni acquisite.

La frequenza di acquisizione effettiva imposta al dispositivo tiene in considerazione i tempi di acquisizione, quelli di elaborazione e il trasferimento delle immagini. E' quindi necessario che la scheda di acquisizione sia in grado di sopportare il carico di informazioni inviate dalla telecamere senza rallentamenti o perdita di informazioni. Questo obiettivo non è sempre facilmente raggiungibile, soprattutto nel caso in cui la risoluzione delle immagini o la frequenza di acquisizione sono molto elevate.

Il tempo di ritardo totale per un sistema di acquisizione comandato da trigger potrà quindi essere calcolato come:

$$t_{delay\ telec.} = t_{delay\ trigger} + t_{esposiz.} + t_{preprocessing} + t_{trasferimento} \quad (5.1)$$

È chiaro che in un sistema stereoscopico è di fondamentale importanza che le telecamere acquisiscano in maniera sincrona.

5.1.2 Scelta delle telecamere

Il criterio di scelta del dispositivo più adatto dipende fortemente dalle applicazioni in cui esso viene utilizzato. Nel sistema stereoscopico in esame si è deciso di utilizzare telecamere UEye UI-5240CP, della casa produttrice *Imaging Development Systems* (IDS) (figura 5.4), le quali specifiche principali sono riassunte in figura 5.3. Questa scelta è stata dettata principalmente dalle specifiche imposte dalle applicazioni che si sono realizzate in questo lavoro e che sono descritte nel capitolo 6.

Sensor	
Sensor Technology	CMOS Mono
Manufacturer	e2v
Resolution (MPix)	1.31
Resolution (h x v)	1280 x 1024
ADC	10 bit
Color depth (camera)	12 bit
Pixel Class	SXGA
Sensor Size	1/1.8
Shutter	
max. fps in Freerun Mode	50.0
Binning Modes	M/C automatic
Subsampling Modes	-
Sensor Model	EV76C560ABT
Pixel size	5.3 μm
Optical Size	6.784 mm x 5.427 mm
Optical sensor diagonal	8.69 mm (1/1.84")

Figura 5-3: Specifiche principali del sensore della telecamera IDS UEye UI-5240CP utilizzata in questo lavoro.



Figura 5-4: *Telecamera IDS UEye UI-5240CP, il modello utilizzato nel sistema stereoscopico in esame.*

Per rispettare le specifiche dello strumento di misura è necessario utilizzare una risoluzione elevata per quanto riguarda le telecamere industriali. Il dispositivo scelto, infatti, presenta un sensore di dimensioni 1280 x 1080 pixel. Chiaramente immagini ad alta risoluzione necessitano più tempo per essere trasferite e occupano più spazio fisico in memoria. Per il formato immagine utilizzato dal dispositivo in esame ogni pixel occupa 8 bit (1 byte); un'intera immagine dunque occuperà circa 1.38 MB. Scegliendo una risoluzione così elevata si è dunque data meno importanza alla frequenza di acquisizione. Nel dispositivo in esame il *frame rate* massimo è infatti 50 Hz se non viene impostata alcuna ROI, ma può aumentare a seconda delle dimensioni della ROI utilizzata. Nel caso in cui è necessario avere un frame rate superiore, è sempre possibile diminuire la ROI aumentando la frequenza di acquisizione.

La telecamera permette di acquisire immagini in *freerun* o in modalità triggerata, permettendo di controllare l'acquisizione con un trigger software o esterno. Nella scelta della telecamera è indispensabile verificare che essa presenti la possibilità di utilizzare un trigger hardware esterno, poiché questo è il metodo più semplice e sicuro per poter sincronizzare più dispositivi differenti (telecamere e illuminatori). La telecamera può essere comandata da un segnale di trigger esterno

di ampiezza compresa tra 5 e 24 V. Inoltre la telecamera consente di acquisire in modalità sincrona o asincrona come descritto nel paragrafo 5.1.1.

La scelta di utilizzare un dispositivo che comunica con il computer attraverso una porta Gigabit Ethernet è dettata principalmente dal fatto che essa consente di utilizzare cavi molto più lunghi rispetto ad altre tecnologie. Questa caratteristica è molto importante soprattutto nelle applicazioni in Galleria del Vento, dove lo strumento viene posizionato in sala prove mentre il computer è collocato in sala controlli. In questo caso è dunque necessario utilizzare un cavo lungo (10-15 m) per la comunicazione tra telecamere e computer.



Figura 5-5: Generatore di funzione e oscilloscopio Velleman PCSU 1000, utilizzato per la generazione del segnale di trigger.

5.1.3 Altri componenti hardware

Per sincronizzare le tre telecamere e id eventuali illuminatori è stata utilizzata un generatore di funzione Velleman PCSU1000 (figura 5.5). In realtà questo dispositivo svolge la funzione di oscilloscopio e generatore di funzione ed è stato scelto in quanto esso comunica con il computer attraverso l'interfaccia USB e può essere comandato direttamente dal software di acquisizione. La funzione oscilloscopio non viene utilizzata, sfruttando solamente la funzione di generatore di funzione. Esso infatti permette di generare un segnale (onda quadra, sinusoidale, dente di sega, ecc..) di ampiezza regolabile (10 Vpp al massimo). Selezionando un

onda quadra di 10 Vpp, impostando il *duty cycle* al 50% del periodo e intervenendo sulla frequenza del segnale è quindi possibile generare un segnale di trigger alla frequenza desiderata. Questo segnale viene inviato alle telecamere e se presenti anche a eventuali illuminatori, così che essi siano tutti comandati dallo stesso segnale e dunque perfettamente sincronizzati tra di loro.

Il computer utilizzato è fornito di 3 differenti schede INTEL Pro1000 che consentono la comunicazione mediante Gigabit Ethernet.

5.1.4 Setup finale

È necessario, a questo punto, trovare la configurazione ottimale dello strumento di acquisizione (dimensione della baseline, angolazione delle telecamere e distanza dall'area di lavoro) in maniera tale da minimizzare l'incertezza di misura e permettere una corretta acquisizione delle misure. Per il posizionamento delle telecamere nello spazio, il supporto utilizzato è riportato in figura 5.6:



Figura 5-6: Sistema di telecamere utilizzato per le misure.

Esso permette, entro certi limiti, di modificare posizione ed orientamento delle tre telecamere. La disposizione delle telecamere, come è in figura 5.6, non risulta ottimale in senso assoluto per quanto riguarda l'incertezza di misura generata dal processo di ricostruzione tridimensionale. Infatti, per far sì che ogni telecamera porti informazioni non ridondanti, sarebbe necessario che esse presentino assi ottici

ognuno ortogonale agli altri due, in maniera tale che una telecamera possa riprendere frontalmente il piano ij , una il jk e l'ultima il ik , come in figura 5.7.

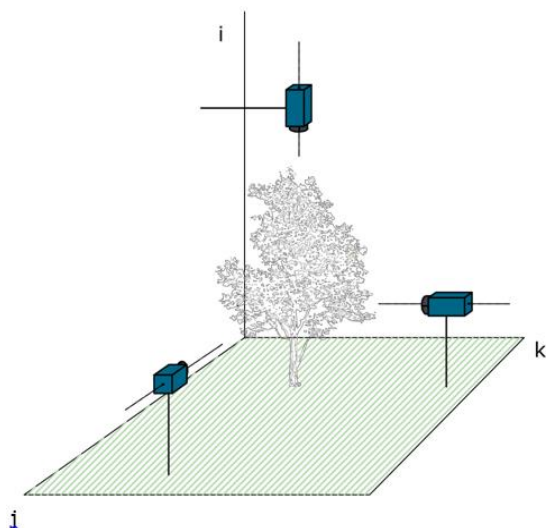


Figura 5-7: *Disposizione ottimale di uno strumento a tre telecamere in cui ogni strumento tende ad essere ortogonale agli altri due.*

Questo perché una telecamera, generando un'immagine in due dimensioni, non è in grado di portare alcuna informazione di profondità. Inoltre, in questo modo, si otterrebbe un errore di ricostruzione tridimensionale uniformemente distribuito lungo le tre direzioni, anziché avere, come accade posizionando le telecamere come in figura 5.6, una ricostruzione 3D della scena in cui una delle coordinate risulta meno accurata rispetto alle altre due. Tuttavia, questo tipo di approccio risulta sconveniente in sede di acquisizione e non è facilmente applicabile. Infatti, mantenendo tale setup si incontrerebbero un gran numero di occlusioni che andrebbe a interferire con le misure. Il setup definitivo utilizzato nelle prove qui descritte è dunque quello di figura 5.6. Da notare che sulle telecamere rappresentate in figura sono montati degli illuminatori e sulle ottiche sono presenti dei filtri; queste componenti sono previste nel sistema in quanto si renderanno utili per alcune applicazioni (capitolo 6) ma non sono stati utilizzati per queste prime prove. Essi verranno poi descritti in dettaglio nel paragrafo 6.1.1.

5.1.5 Software

L'applicativo software è stato sviluppato nell'ambiente di sviluppo Visual Studio 2010 della Microsoft. In particolar modo è stato scelto il linguaggio C# con il framework .Net 4.0, soprattutto perché permette di utilizzare la programmazione ad oggetti che per sua natura è molto versatile, flessibile e consente di gestire ottimamente gli eventi che si susseguono all'interno del programma. Inoltre è da sottolineare la facile mantenibilità del software, la modularità, che consente di scomporre il problema in più semplici sotto-problemi, e la possibilità di riutilizzare facilmente il codice in altri ambiti. Nella potente piattaforma .Net vi sono moltissime librerie che consentono di effettuare parte delle applicazioni ordinarie. Inoltre è un'ottima piattaforma di programmazione poiché consente di integrare facilmente librerie prodotte da terze parti.

Si è deciso di integrare la programmazione in C# con la libreria di image processing HALCON, realizzata da MVTEC. Questa possiede un proprio ambiente di sviluppo, chiamato HDevelop, ed è integrabile con i più comuni linguaggi di programmazione: C, C++, C#, Visual Basic .Net e Delphi. Alcune applicazioni sono quindi state implementate in HDevelop e poi esportate in linguaggio C#, mentre altre più semplici sono state scritte direttamente in C# utilizzando comandi provenienti dalla libreria HALCON.

5.2 Taratura dello strumento

Un requisito necessario per effettuare misurazioni tridimensionali è che il sistema di misura sia tarato. Per la taratura della tripletta di telecamere si è deciso di sviluppare un tool in C#, utilizzando la libreria HALCON. Il metodo di calibrazione utilizzato è dunque quello proposto da Heikkilä, che è stato validato e testato nel capitolo 4, dove si è verificato che l'accuratezza di tale metodo è compatibile con quella di Zhang-LM, anche se esso presenta alcune limitazioni. Nel paragrafo 4.2.1 è stato sottolineato come tali limitazioni siano ininfluenti per le

applicazioni in cui si colloca il sistema di misura in esame. La scelta di utilizzare l'algoritmo di HALCON è stata dettata da due principali motivi. Prima di tutto si è voluto creare un software utilizzabile anche da un utente che non conosca nel dettaglio il processo di calibrazione. HALCON risulta semplice e immediato da utilizzare, essendo un software commerciale creato anche per utenti meno esperti. Inoltre si vuole inserire la procedura di taratura in un pacchetto software unico, utilizzando il linguaggio di programmazione C# e la libreria HALCON. È vero che la taratura della tripletta viene fatta offline a monte delle acquisizioni, ma risulta comunque più comodo e ordinato utilizzare lo stesso linguaggio di programmazione per tutto il pacchetto software. Un'ulteriore motivazione che ha influenzato tale scelta è che il calcolo delle corrispondenze tra i punti del modello del calibratore e i punti immagine in HALCON è effettuata in maniera automatica. Infine HALCON è particolarmente comodo per la taratura di un sistema stereoscopico composto da più dispositivi. Infatti, oltre alla calibrazione di telecamere singole, è possibile eseguire facilmente la taratura di un sistema stereoscopico formato da più dispositivi. È sufficiente selezionare il numero di telecamere che lo compongono e per ognuna di esse acquisire un set di immagini in maniera sincrona con le altre. L'algoritmo di taratura di un sistema stereoscopico di HALCON consente di trovare i parametri intrinseci ed estrinseci di tutte le telecamere del sistema. I parametri estrinseci sono riferiti a uno dei dispositivi del sistema stereoscopico, selezionato dall'utente. La procedura di taratura di una tripletta restituisce un valore di errore di retroproiezione che non sarà più riferito alla singola telecamera, bensì all'intera tripletta (valor medio delle tre). Nel seguito con il termine “*calibrazione*” si intenderà l'intera procedura di taratura di un sistema stereoscopico, costituita dalla calibrazione delle singole telecamere e la loro registrazione.

È stato quindi implementato un tool, la cui interfaccia è mostrata in figura 5.8, che consente di acquisire, tramite i tre dispositivi, una serie di immagini di un calibratore in diverse pose. Le immagini sono acquisite in maniera sincrona per ogni dispositivo e, se il calibratore viene trovato con successo in tutte e tre le immagini, esse vengono salvate. Così è possibile creare l'insieme di immagini da cui l'algoritmo di HALCON ricava i dati per eseguire la taratura delle tre

telecamere. I risultati di taratura sono poi salvati in un file e serviranno per eseguire la ricostruzione stereoscopica.

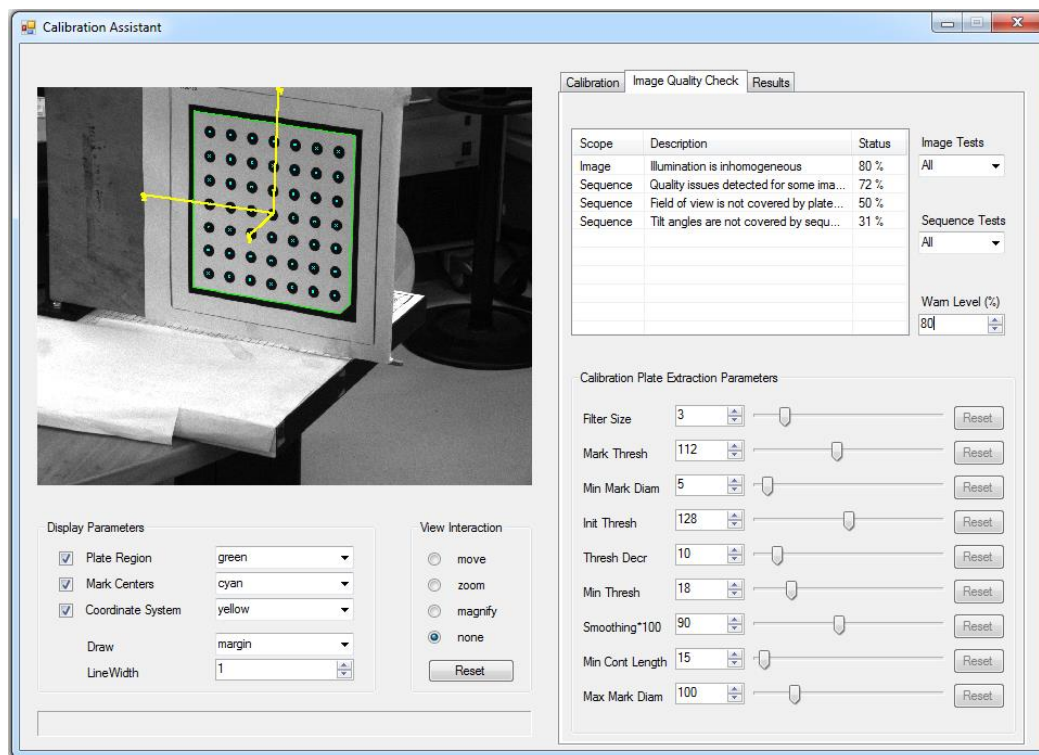


Figura 5-8: Interfaccia del tool di calibrazione inserito nel pacchetto software.

5.3 Studio del legame tra errore 3D e errore 2D

Come descritto nel paragrafo 4.1.3, durante la procedura di calibrazione viene minimizzato l'errore di retroproiezione 2D. Esso rappresenta la distanza tra i punti individuati nelle immagini (i dati di partenza) e i punti del calibratore trasformati sulla base dei parametri di telecamera, quindi retroproiettati. L'errore di retroproiezione mantiene un significato fisico ed è un buon indice della bontà della taratura ma non sempre definisce l'accuratezza dello strumento in termini di errore di ricostruzione tridimensionale, parametro di massimo interesse quando si effettuano misure 3D. Si vuole dunque in questo lavoro determinare quanto l'errore

2D (o di retroproiezione) sia indice della bontà della ricostruzione tridimensionale, vale a dire quanto esso sia legato all'errore 3D.

Per fare ciò sono state effettuate una serie di calibrazioni utilizzando un pattern planare a geometria nota come quello rappresentato in figura 5.10; esso è costituito da una piastra in alluminio sulla quale è stato stampato, con accuratezze inferiori a 0.05 mm, un quadrato di dimensioni 200x200 mm con all'interno 49 marcatori circolari. Esso è stato posto in flangia a un robot antropomorfo ABB IRB1600-6/1.45. In questo modo è possibile dare al calibratore delle pose note con un'accuratezza superiore a quella del sistema di misura e dunque considerata trascurabile. Tutti i produttori di robot devono rispettare la norma UNI EN ISO 9283. Questa è una delle principali norme internazionali relative alla robotica e regola i criteri di prestazioni ed i relativi metodi di test. In questi documenti risulta evidente che le specifiche di ripetibilità e accuratezza debbano essere ottenute in modo statistico, attraverso esperimenti ripetuti, eseguendo precise procedure. Per il robot ABB IRB1600-6/1.45, da datasheet, i dati riguardanti ripetibilità ed accuratezza sono illustrati di seguito

	Statica (mm)	Dinamica (mm)
Ripetibilità	0.02	0.19
Accuratezza	0.04	1.03

Le discrepanze tra caso statico e caso dinamico sono da imputare alle forze di inerzia e alle risonanze meccaniche. Possono infatti influire durante il movimento del robot. Si può tuttavia assumere che, per velocità ridotte del robot, i valori di accuratezza e ripetibilità nel caso dinamico, si avvicinino ai valori statici. In ogni caso per le prove in esame il calibratore verrà movimentato da una posa all'altra e acquisito a robot fermo in potenza, quindi è di interesse solamente il valore di ripetibilità e accuratezza del caso statico.



Figura 5-9: Robot antropomorfo ABB IRB1600-6/1.45 utilizzato per le prove di qualificazione metrologica del sistema stereoscopico.



Figura 5-10: Calibratore planare montato in pinza al robot ABB IRB1600.

Vista la possibilità di avere il calibratore in pose note è stato possibile calcolare due diversi errori tridimensionali: il primo è l'errore sulle distanze tra i marker (chiamato nel seguito *errore 3d sulle distanze*), calcolato come differenza degli scostamenti tra le distanze dei marker di un calibratore in una certa posa ricostruite dal sistema stereoscopico e quelle reali (conosciute con accuratezze trascurabili rispetto a quelle attese dal sistema di misura); il secondo è l'errore dalle posizioni dei marker in diverse pose movimentate mediante il braccio robotico (chiamato nel seguito *errore 3d sulle traslazioni*). Quest'ultimo è dato dallo scostamento medio delle distanze tra i blob misurate in pose differenti con il sistema stereoscopico e le distanze reali. Tali distanze reali sono note poiché è conosciuta la posa in cui si trova il calibratore in pinza al braccio robotico.

Prima di tutto sono state acquisite un numero sufficiente di pose del calibratore da coprire abbondantemente tutto lo spazio inquadrato, utilizzando diverse angolazioni, come spiegato nel paragrafo 4.1.1. In particolare si è deciso di utilizzare 200 diverse pose del calibratore montato in pinza al robot. Con queste si è proceduto ad effettuare diverse calibrazioni. In particolare si sono effettuate calibrazioni utilizzando un numero di immagini da 15 a 100 con passo 5. Il valore minimo di 15 immagini è giustificato dalle prove descritte nel paragrafo 5.4, atte a valutare la bontà della procedura di taratura al variare dal numero di pose utilizzate. Per ognuna di queste sono state fatte 20 diverse calibrazioni scegliendo in maniera casuale terne di immagini tra quelle acquisite. In questo modo le immagini selezionate potrebbero non soddisfare i requisiti necessari alla creazione di un buon set di dati di calibrazione, descritti in dettaglio nel paragrafo 4.1.1. In ogni calibrazione viene restituito l'errore di retroproiezione 2D medio trovato. Vengono dunque calcolati l'errore 3D sulle distanze e l'errore 3D sulle traslazioni. Il primo è calcolato come media degli scostamenti delle distanze tra i marcatori di un calibratore, le quali coordinate tridimensionali sono ricostruite con il sistema stereoscopico, dalle distanze reali (conosciute a priori con accuratezza elevata). Esso viene trovato come media su 20 diverse pose del calibratore. Il secondo invece viene calcolato come segue: vengono selezionate 20 pose del calibratore, note con accuratezza trascurabile rispetto a quella attesa dal sistema di misura, in quanto esso

è montato in pinza al robot; si conoscono dunque le distanze relative tra i marker nelle diverse pose. Mediante il sistema stereoscopico vengono misurate le stesse distanze relative tre più pose e viene calcolato lo scostamento medio tra queste e quelle prese a riferimento dal robot, trovando il valore dell'errore 3D sulle traslazioni. L'errore di retroproiezione restituito dalla procedura di calibrazione viene messo in relazione con i due errori 3D trovati.

Il modello di telecamera utilizzato è quello denominato divisionale (paragrafo 4.2), vale a dire più semplice che tiene conto solamente del primo parametro di distorsione radiale. Le ottiche montate sulle telecamere hanno infatti valori di lunghezza focale superiori a 25 mm, caratteristica che permette di trascurare gli altri parametri di distorsione se l'ottica utilizzata è di buona qualità.

Nelle figure seguenti vengono riportati gli andamenti di entrambi gli errori 3D in funzione dell'errore 2D trovato in fase di taratura. Per questione di leggibilità dei grafici non sono stati considerati i punti con errore di retroproiezione maggiore di 2 pixel, poiché in questi casi si hanno calibrazioni totalmente errate e quindi errori 3D molto elevati (superiori a qualche millimetro).

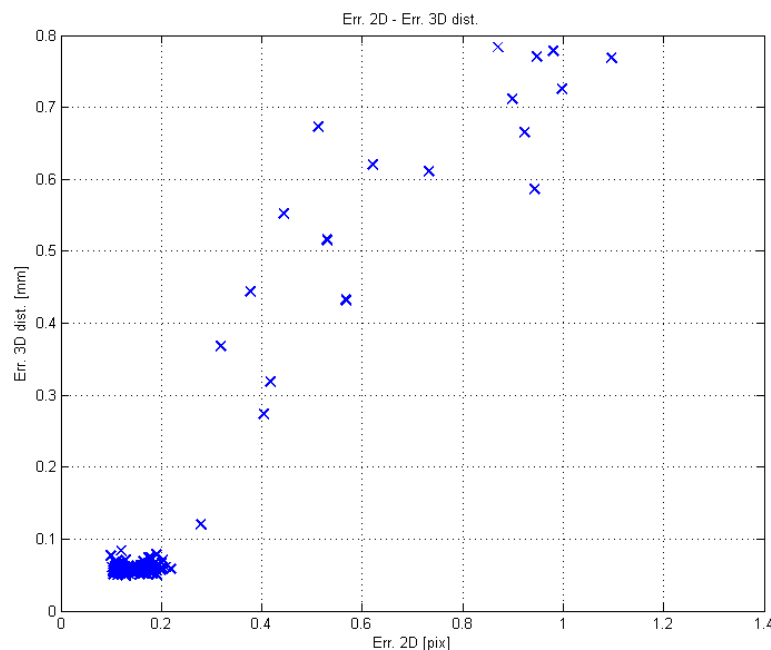


Figura 5-11: Errore 3D sulle distanze al variare dell'errore di retroproiezione 2D.

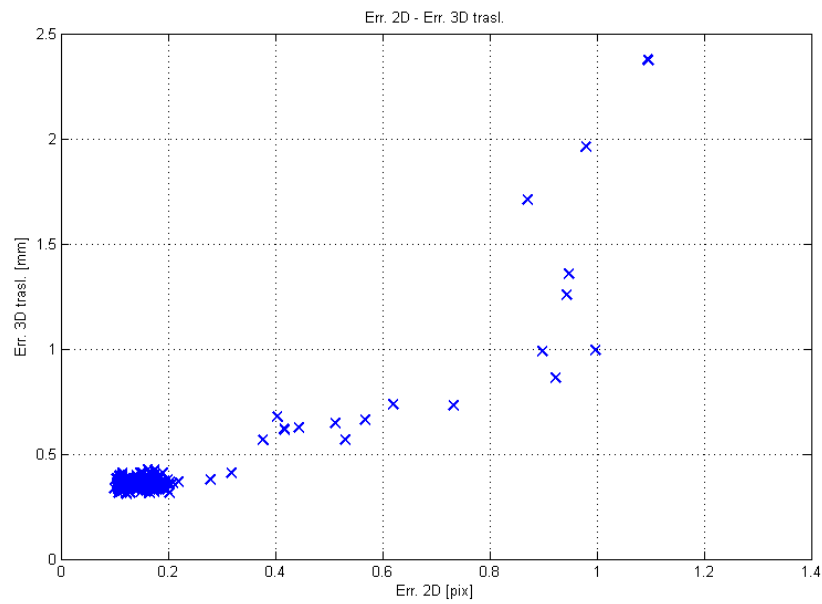


Figura 5-12: Errore 3D sulle traslazioni al variare dell'errore di retroproiezione 2D.

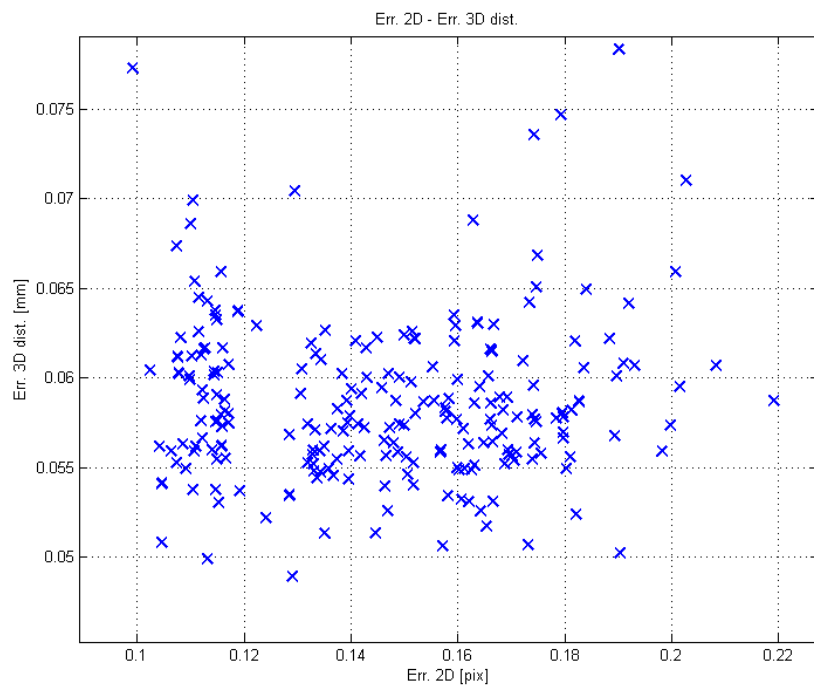


Figura 5-13: Errore 3D sulle distanze al variare dell'errore di retroproiezione 2D (particolare per errori 2D bassi).

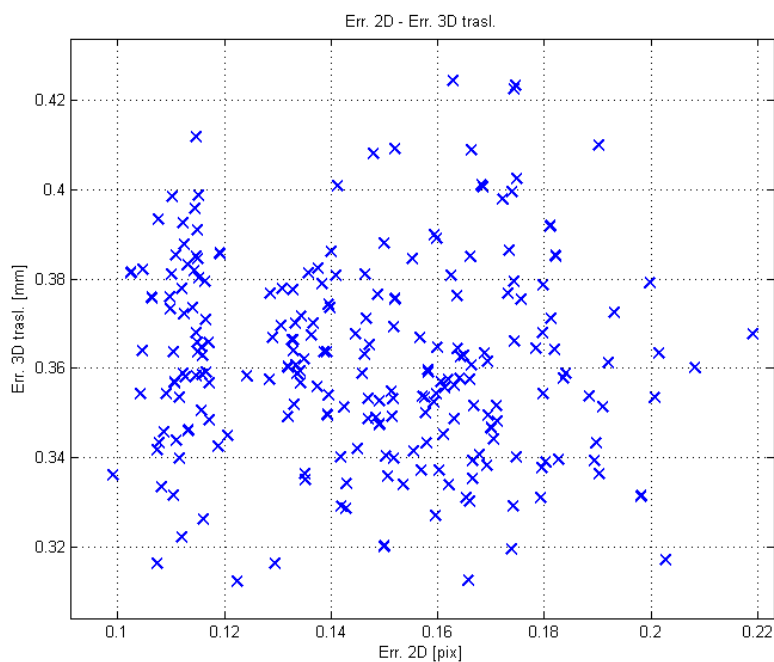


Figura 5-14: *Errore 3D sulle traslazioni al variare dell'errore di retroproiezione 2D (particolare per errori 2D bassi).*

Come si vede dalle figure 5.11 e 5.12 l'errore di retroproiezione è un buon parametro per definire se la calibrazione è stata effettuata correttamente. Al crescere dell'errore di retroproiezione restituito dalla procedura di calibrazione infatti sia l'errore sulle distanze che quello sulle traslazioni crescono in maniera considerevole.

Osservando le figure 5.13 e 5.14, che mostrano il legame tra errore 2D e 3D per valori di errore di retroproiezione molto bassi (< 0.22 pixel), si nota che è presente una fase di transizione in cui al crescere dell'errore di retroproiezione 2D, l'errore sulle misure 3D non cresce. Questo implica che per valori modesti dell'errore di retroproiezione la calibrazione ha dato risultati ottimali e esiste un intervallo di valori di errore 2D in cui l'accuratezza della calibrazione rimane ottimale. In base ai test effettuati si può concludere che per valori di errore di retroproiezione superiori a 0.22 pixel, al crescere dell'errore 2D peggiora l'accuratezza del sistema. Infatti, come evidenziato dai grafici, l'errore 3D calcolato aumenta in modo proporzionale all'errore 2D.

Sono state effettuate altre prove utilizzando differenti configurazioni del sistema stereoscopico ottenute modificando la distanza dal robot, il campo inquadrato, la posa relativa delle telecamere e la lunghezza focale delle ottiche. Tutte le prove svolte hanno confermato i risultati descritti sopra. Si è deciso dunque, per motivi di leggibilità, di riportare solamente i dati raccolti in una singola campagna sperimentale nella quale il sistema stereoscopico presenta una configurazione considerata standard per l'utilizzo in Galleria del Vento.

5.4 Efficacia della calibrazione al variare del numero di pose utilizzate

La procedura di calibrazione implementata nelle librerie HALCON prevede l'utilizzo di almeno tre diverse pose del calibratore per poter eseguire l'algoritmo di minimizzazione dell'errore di retroproiezione. Il produttore delle librerie però consiglia di utilizzare almeno 10 differenti pose del calibratore per garantire risultati accurati. Si vuole con queste prove valutare come varia l'errore di retroproiezione 2D e l'errore 3D al variare del numero di pose del calibratore utilizzate.

Per fare ciò vengono effettuate varie calibrazioni acquisendo un numero diverso di pose compreso tra 3 e 100. Le pose del calibratore utilizzate per effettuare le varie calibrazioni sono state volutamente scelte senza prestare troppa attenzione agli accorgimenti necessari a raccogliere un buon set di dati, in termini di orientamento e posizione del calibratore (paragrafo 4.1.1). In questo modo si vuole caratterizzare al meglio il processo di taratura, che non sempre viene eseguito da utenti esperti. In tutte le immagini acquisite però il calibratore deve essere ben correttamente a fuoco, ben visibile, interamente inquadrato da tutte le telecamere ed illuminato in modo omogeneo. In particolare per ogni numero di pose utilizzate vengono fatte 20 diverse calibrazioni con pose casuali. Viene riportato nel seguente grafico l'andamento dell'errore di retroproiezione ottenuto nelle 20 diverse calibrazioni per ogni numero di pose utilizzato.

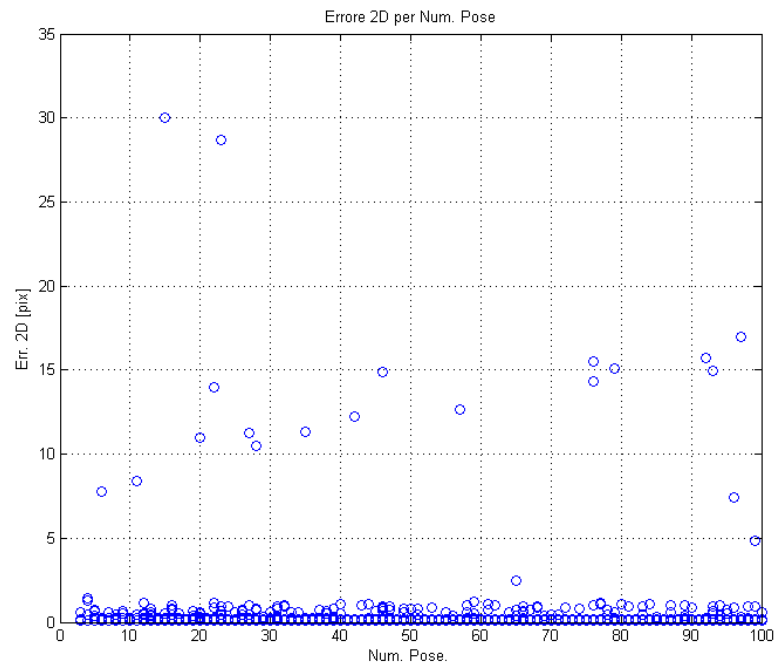


Figura 5-15: Andamento dell'errore di retroproiezione per calibrazioni con diverso numero di pose.

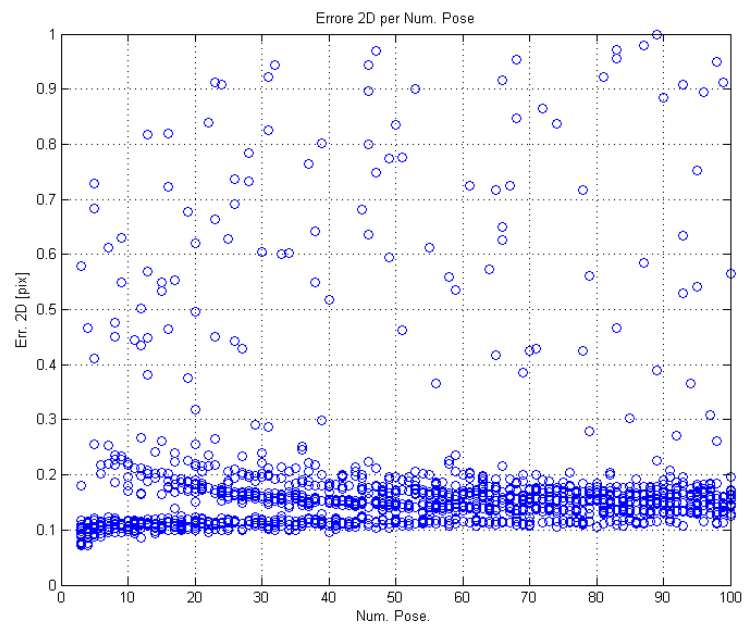


Figura 5-16: Andamento dell'errore di retroproiezione per calibrazioni con diverso numero di pose (risultati per errore 2D < 1 pixel).

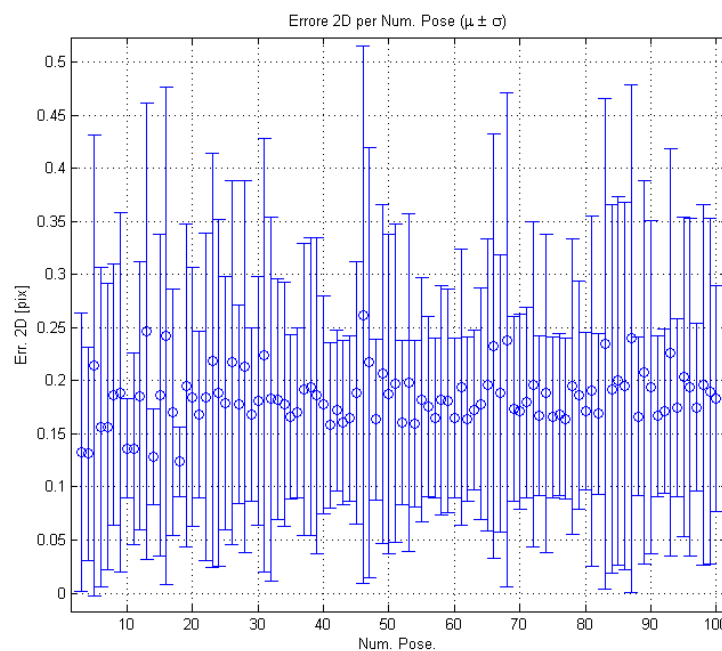


Figura 5-17: Andamento dell'errore di retroproiezione per calibrazioni con diverso numero di pose (risultati per errore 2D < 1 pixel), rappresentato come valore medio \pm la deviazione standard.

Dalla figura 5-15 l'errore di retroproiezione non sembra essere correlato con il numero di pose utilizzato. Inoltre esso può assumere valori anche molto alti indicando una calibrazione che non è andata a buon fine (errore 2d > 1 pixel). Vi sono infatti casi in cui per qualsiasi numero di pose utilizzato qualche calibrazione ha restituito valori di errore 2D elevati. In questi casi il set di immagini di calibrazione, che è stato scelto in maniera casuale, non è ottimale ai fini della calibrazione (ad esempio potrebbero esserci numerose pose molto simili e vicine tra di loro e poche molto differenti, oppure le pose scelte occupavano quasi tutte solamente un'area delle immagini, ecc...). Questi sono i casi in cui sia l'errore 2D che l'errore 3D crescono in maniera evidente, come mostrato nelle figure 5.11 e 5.12.

In figura 5-16 è riportato lo stesso grafico ma filtrando le calibrazioni in cui l'errore di retroproiezione ha dato valore maggiore di 1 pixel, vale a dire eliminando i casi in cui la calibrazione ha sicuramente restituito valori non accettabili. In figura

5-17, invece, ancora gli stessi dati sono riportati utilizzando una notazione grafica differente, vale a dire rappresentando la media dell'errore di retroproiezione trovato per ogni numero di pose (cerchio blu) e le bande di incertezza rappresentate dagli intervalli definiti da \pm la deviazione standard. Da questi grafici si nota che per calibrazioni con un numero di pose inferiori a 15 può succedere che il valore dell'errore di retroproiezione risulti essere addirittura migliore della media. Questo fatto è causato da una mancanza di informazioni nella procedura di calibrazione. Calibrando infatti con pochissime immagini è possibile che esse risultino avere pose tutte molto simili tra loro. In questo caso l'algoritmo di ottimizzazione dell'errore di retroproiezione riesce a calcolare una configurazione dei parametri di calibrazione che minimizza l'errore 2D in maniera consistente solamente per quanto riguarda quella posa. Quello che accade dunque è che nonostante l'errore 2D restituito sia molto basso, i parametri di calibrazione risultano essere ottimali solamente per quella particolare posa, ma non sono generalmente corretti. Andando ad analizzare i grafici che mostrano l'andamento degli errori 3D (sulle distanze e sulle traslazioni) al variare del numero di pose (figura 5-18 e 5-19), si nota infatti che per un numero di pose inferiore a 15 l'errore 3D non è mai sotto il valor medio (0.1 mm per l'errore 3D sulle distanze e 0.4 mm per l'errore 3D sulle traslazioni) ma piuttosto presenta valori molto più alti della media. I grafici degli errori 3D sono stati realizzati filtrando le calibrazioni che restituiscono un errore di retroproiezione maggiore di 1 pixel. Queste infatti danno errori 3D molto elevati in quanto la calibrazione è risultata completamente errata e avrebbero reso i grafici poco leggibili. Da queste considerazioni si può intuire come nel caso di utilizzo di meno di 15 pose è possibile incorrere in false buone calibrazioni, che restituiscono un valore di errore di retroproiezione molto basso e dunque sembrano essere state eseguite correttamente, ma in realtà portano a commettere errori in fase di misura (errori 3D) molto elevati. I set di dati acquisiti per effettuare queste calibrazioni dunque non rispecchiano i requisiti descritti nel capitolo 4.1.1.

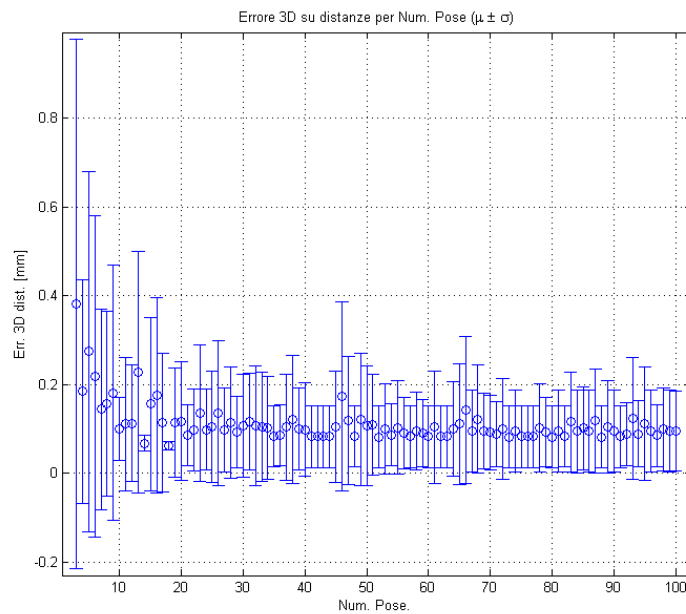


Figura 5-18: Andamento dell'errore 3D sulle distanze, rappresentato come valore medio \pm la deviazione standard, per calibrazioni con diverso numero di pose (risultati per errore 2D < 1 pixel).

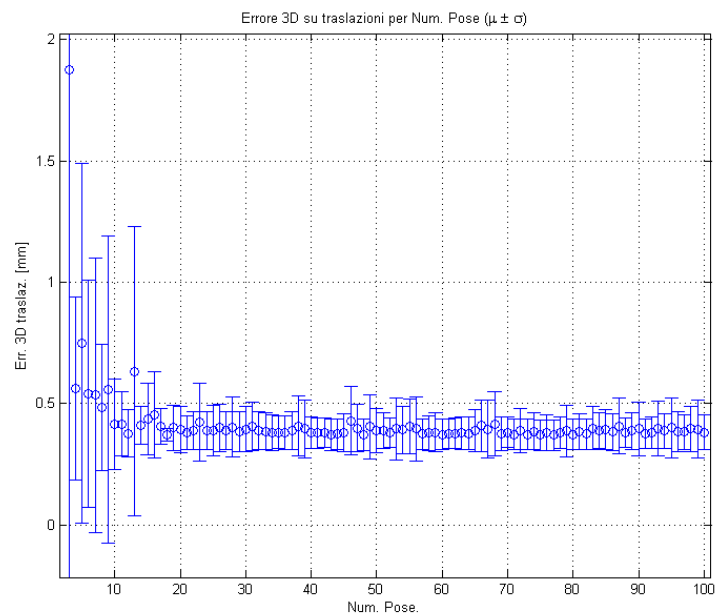


Figura 5-19: Andamento dell'errore 3D sulle traslazioni, rappresentato come valore medio \pm la deviazione standard, per calibrazioni con diverso numero di pose (risultati per errore 2D < 1 pixel).

Riassumendo quindi è possibile dire che per effettuare una calibrazione correttamente è conveniente utilizzare almeno 15 pose differenti. Sopra questo valore è possibile affermare che l'errore di retroproiezione restituito dalla procedura di calibrazione definisce la bontà della calibrazione e garantisce una buona accuratezza in fase di misura. Utilizzando meno di 15 immagini invece si rischia di incorrere in false buone calibrazioni, vale a dire calibrazioni che restituiscono un errore 2D molto basso ma errori 3D elevati, dunque un basso grado di accuratezza di misura.

5.5 Studio dell'accuratezza di misura all'esterno dal volume di calibrazione

Nonostante sia buona norma calibrare tutto il volume di lavoro in cui si prevede di effettuare le misure tridimensionali, questo risulta difficile in alcune applicazioni particolari. Ad esempio per alcune applicazioni in Galleria del Vento può risultare difficile prevedere la posizione che assumerà il modello sottoposto alla forza del vento. Altre volte invece lo spazio in cui il misurando andrà a muoversi è coperto dall'oggetto stesso e quindi non si ha la possibilità di calibrare l'intero volume di misura. Per questo motivo in questo paragrafo si vuole approfondire come varia l'accuratezza quando vengono effettuate delle misure fuori dallo volume calibrato, vale a dire l'area che racchiude tutte le coordinate tridimensionali dei marker del calibratore che sono stati considerati in fase di calibrazione.

Per eseguire queste verifiche sono state acquisite le pose del calibratore planare mostrate in figura 5.20. In particolare il sistema di riferimento ha l'origine nel centro del calibratore nella prima posa acquisita; gli assi x e y sono complanari al piano del calibratore, mentre l'asse z è perpendicolare ad esso ed è negativo in direzione del sistema stereoscopico. Tra le pose acquisite, 80 hanno la coordinata z del centro del calibratore compresa tra -100 mm e $+100$ mm (in blu nella figura 5.20). Di queste 80 solo 60 sono state utilizzate per effettuare la calibrazione del sistema stereoscopico, mentre le altre 20 vengono utilizzate per calcolare l'errore 3D sulle distanze. Si ricorda che si intende errore 3D sulle distanze la media degli scostamenti delle distanze tra i marker del calibratore ricostruiti col sistema stereoscopico dalle distanze reali. Inoltre sono state acquisite altre 20 pose fuori dallo spazio utilizzato per calibrare, quindi con la coordinata Z del centro maggiore di 300 mm o minore di -300 mm (in rosso nella figura 5.20). Comunque tutti i calibratori nelle pose selezionate in questa prova hanno una coordinata z del centro compresa tra -500 mm e 500 mm in modo da non uscire dalla profondità di campo delle tre telecamere. In questo modo tutte le immagini utilizzate nella prova risultano a fuoco. Nella figura 5.21 sono riportati i risultati della prova; in

particolare sull'ascissa è riportata la coordinata z del centro del calibratore ricostruito e sulle ordinate l'errore 3D di ricostruzione dello stesso. Questo ripetuto per 20 prove nello spazio di calibrazione (in blu) e 20 fuori da esso (in rosso).

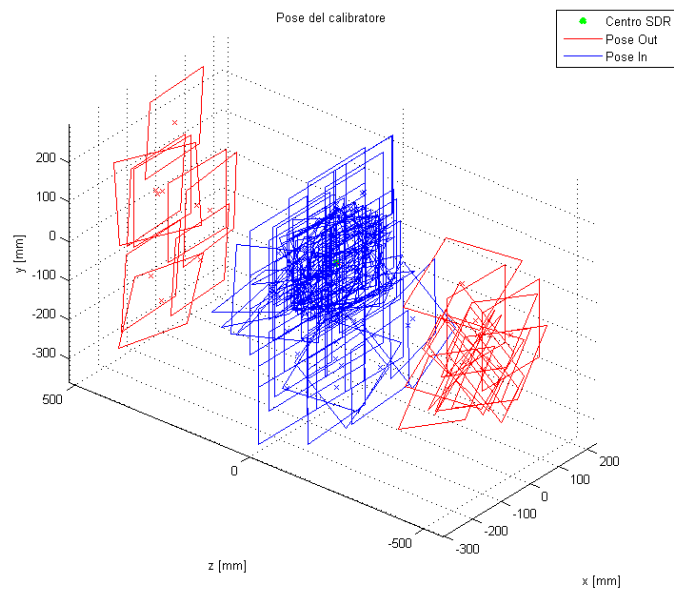


Figura 5-20: Pose dei calibratori acquisiti. In blu le pose all'interno del campo di calibrazione e in rosso quelle all'esterno.

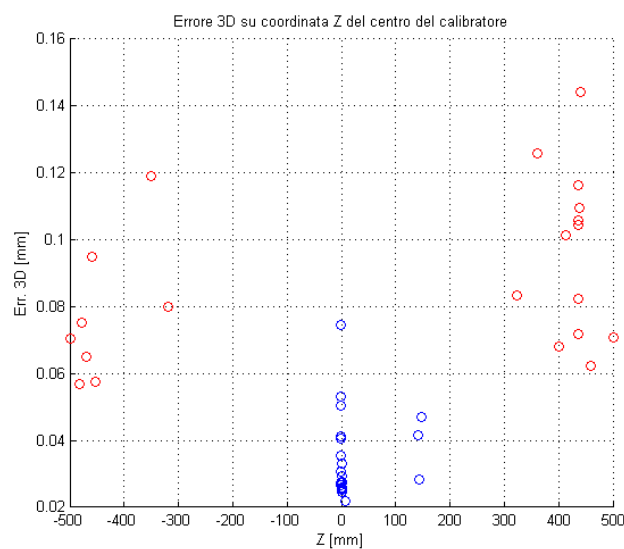


Figura 5-21: Errore 3D sulle distanze tra i marker del calibratore a seconda della coordinata Z del suo centro. In blu i risultati nella zona di calibrazione, in rosso quelli fuori dalla zona di calibrazione.

Come si vede dalle analisi effettuate l'errore 3D è più alto fuori dal campo di calibrazione (punti rossi). In particolare è stata calcolata la media degli errori 3D di tutte le pose fuori spazio di calibrazione e di tutte le pose al suo interno. I risultati sono i seguenti:

$$Err. 3D_{dentro} = 0.04 \text{ mm}$$

$$Err. 3D_{fuori} = 0.09 \text{ mm}$$

I dati quindi mostrano che è effettuare misure fuori dallo spazio di calibrazione (nei limiti del campo inquadrato e delle profondità di campo date dalle ottiche utilizzate), diminuisce l'accuratezza delle misure. È dunque buona norma definire quale sarà il campo in cui si andranno ad effettuare le misure e preoccuparsi di tarare il sistema stereoscopico muovendo il calibratore nell'intero spazio di misura.

5.6 Studio della propagazione dell'incertezza mediante il metodo Monte Carlo

Si vuole studiare la propagazione dell'incertezza di misura durante il processo di calibrazione nel quale viene utilizzato l'algoritmo di Heikkilä. Per fare ciò si utilizza il metodo Monte Carlo, alla quale la GUM (*“Guide to the expression of uncertainty in measurement”*) dedica un addendum sull'analisi della propagazione delle incertezze [20].

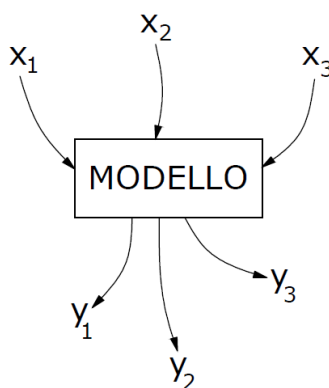


Figura 5-22: Modello deterministico caratterizzato dal set di ingressi x_i e dal set di uscite y_i .

Il metodo Monte Carlo è un metodo numerico basato su procedimenti probabilistici, usato in statistica per la risoluzione di problemi di varia natura, che presentano difficoltà analitiche non altrimenti o difficilmente superabili. Prende nome dal casinò di Monte Carlo, simbolo del gioco d'azzardo per antonomasia [21]. Concettualmente il metodo si basa sulla possibilità di eseguire, utilizzando numeri casuali, un campionamento di una distribuzione di probabilità assegnata, $F(X)$; ossia sulla possibilità di generare una sequenza di eventi $X_1, X_2, \dots, X_n, \dots$, distribuiti secondo la $F(X)$. In pratica, invece di servirsi di un campione di numeri effettivamente estratti a caso, si ricorre a una sequenza di numeri ottenuti con un processo iterativo ben determinato; tali numeri vengono detti pseudo-casuali giacché, pur non essendo casuali, hanno proprietà statistiche analoghe a quelle dei veri numeri casuali. Il metodo Monte Carlo viene utilizzato per trovare le soluzioni di problemi matematici, che possono avere molte variabili e che non possono essere risolti facilmente, per esempio il calcolo integrale. In questi casi infatti può risultare difficile, se non impossibile, studiare la propagazione dell'incertezza derivando le equazioni che descrivono il modello del sistema. Il metodo Monte Carlo consente di studiare la propagazione dell'incertezza senza invertire le equazioni del modello, se questo è deterministico [22]. Risulta quindi molto conveniente quando il modello in analisi è complesso, non lineare o include più di due parametri incerti. L'efficienza di questo metodo aumenta rispetto agli altri metodi quando la dimensione del problema cresce. Una simulazione conta, tipicamente, un numero molto elevato di iterazioni del modello.

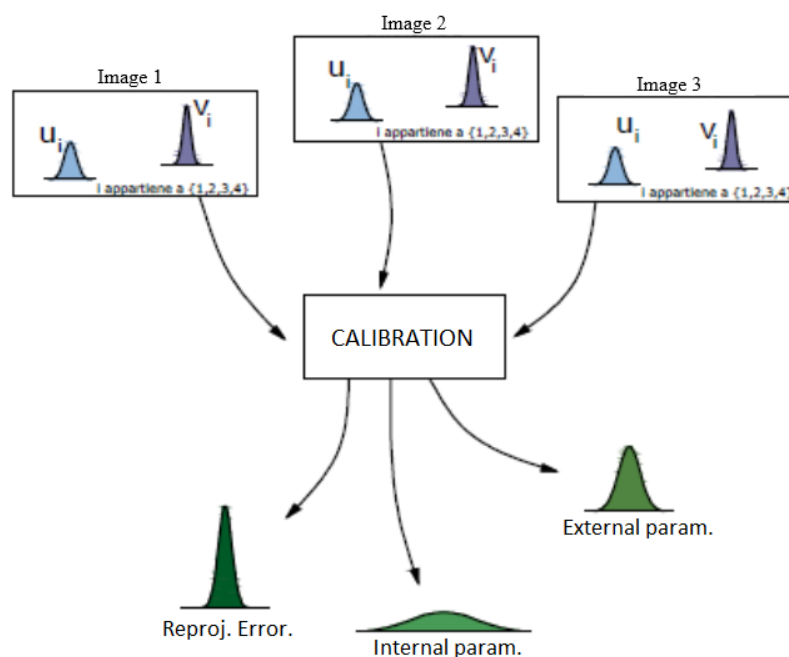


Figura 5-23: Schematizzazione della simulazione Monte Carlo per la propagazione stocastica dell'incertezza.

Nel caso della calibrazione si vuole utilizzare il metodo Monte Carlo per studiare la propagazione delle incertezze sulla stima di diversi parametri intrinseci ed estrinseci. Prima di tutto è necessario stimare l'incertezza della procedura di acquisizione delle immagini con il sistema stereoscopico. Per fare ciò si acquisisce un numero sufficientemente grande di immagini del calibratore (500 immagini) in posizione fissa, dalle quali è possibile stimare la distribuzione delle coordinate immagine (u, v) dei vari marcatori. Per ognuna di queste, mediante un algoritmo di blob detection, vengono estratti i centri dei 49 marker in coordinate immagine (u, v). Dato che il sistema è composto da tre telecamere si potranno così creare 294 vettori (49 punti \times 2 coordinate \times 3 immagini) composti da 500 elementi (numero delle diverse immagini acquisite con il sistema stereoscopico). Ad ogni vettore viene sottratto il proprio valor medio e viene calcolata la funzione di densità di probabilità e la deviazione standard. Si ottengono così 294 valori di deviazione standard; questi valori sono tra loro tutti molto simili ed in particolare il valore più piccolo riscontrato è stato 0.020 pixel mentre quello massimo 0.036 pixel con un valor medio di tutte le deviazioni standard di 0.026 pixel. Nella figura 5.24 vengono

rappresentate le funzioni di densità di probabilità (PDF) per ogni coordinata di ogni immagine; in questo caso le u e v di ogni blob sono state considerate tutte insieme così da avere $2 \times 3 = 6$ vettori da rappresentare. In viola viene rappresentato la linea che identifica il valor medio mentre in azzurro, giallo e verde rispettivamente le bande di incertezza estesa al 68%, 95%, 99.7% ($\pm\sigma$, $\pm 2\sigma$, $\pm 3\sigma$), approssimando la distribuzione a una gaussiana standard.

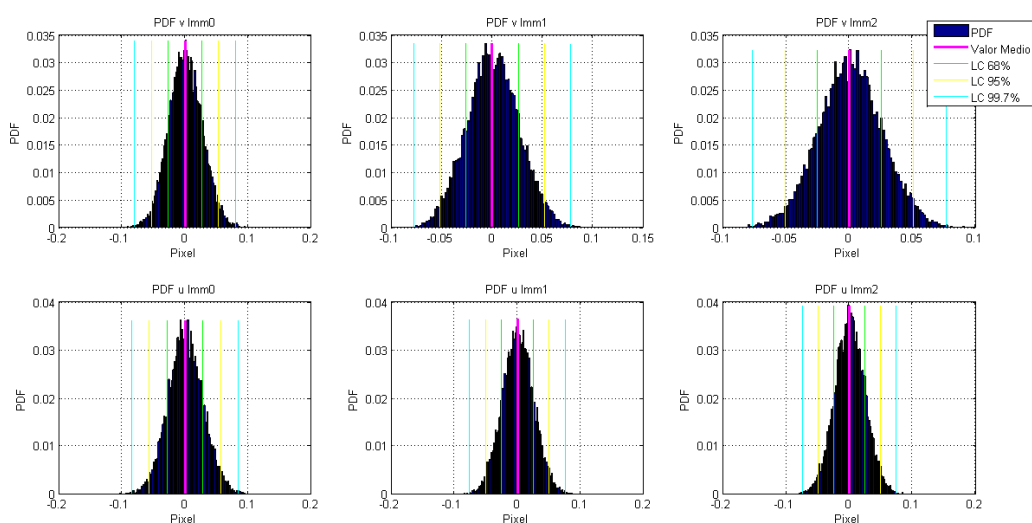


Figura 5-24: Funzione di densità di probabilità delle coordinate (u, v) dei blob di un calibratore planare fermo, nelle 3 diverse immagini.

Già dalla figura è facile intuire che i dati acquisiti sono approssimabili mediante una distribuzione gaussiana. Per verificare questa affermazione viene eseguito il test di Kolmogorov-Smirnov [23], che è un test non parametrico che consente di verificare la forma della distribuzione campionaria. Esso è molto utilizzato in quanto rappresenta l'alternativa non parametrica al test della t -Student [16] e quando si ha un'ipotesi parametrica di distribuzione gaussiana, come nel nostro caso, l'efficienza del test è circa il 95% per piccoli campioni e diminuisce leggermente per campioni più grandi. Esso è più potente rispetto altri test non parametrici (ad esempio il test della mediana o il test del chi quadrato [16]) e per campioni piccoli si ritiene essere più potente rispetto al test di Wilcoxon-MannWhitney [24]. Con il test di Kolmogorov-Smirnov si vuole quindi verificare che le distribuzioni campionarie raffigurate in Figura 5-24 siano approssimabili con

una gaussiana. Il test restituisce valore positivo per tutte le distribuzioni; nel seguito sono riportate le distribuzioni campionarie con la relativa funzione di distribuzione cumulativa (CDF) in blu e la CDF di una gaussiana standard. Si nota come la curva in rosso sia sovrapponibile a quella in blu per tutte le coordinate in tutte le immagini.

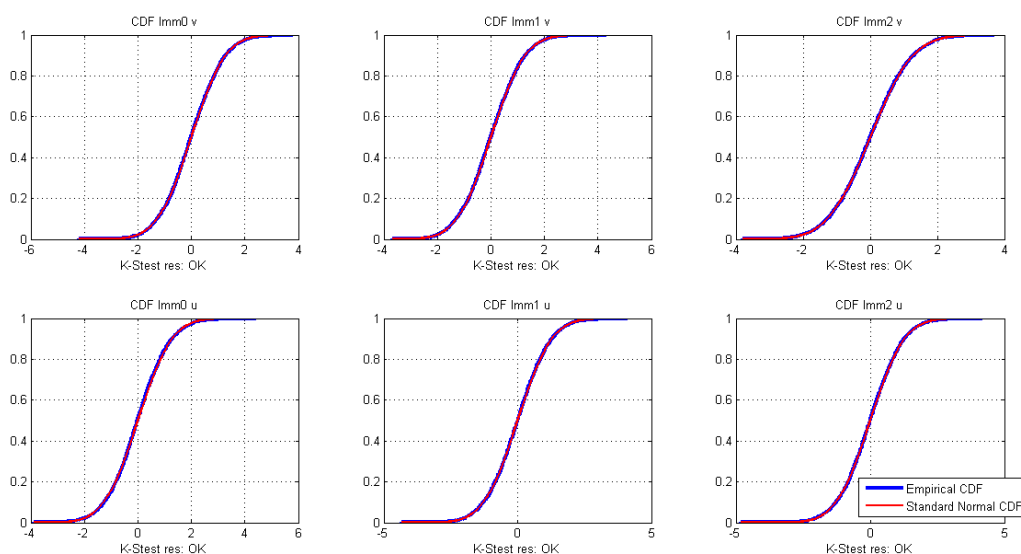


Figura 5-25: Funzione di distribuzione cumulativa delle coordinate (u , v) dei blob di un calibratore planare fermo, nelle 3 diverse immagini (in blu). In rosso funzione cumulativa di una normale standard. Il risultato del test di Kolmogorov-Smirnov è sempre positivo.

Quindi è possibile concludere che l'acquisizione dei punti di un calibratore da parte del sistema stereoscopico è soggetta a un errore di acquisizione che è approssimabile con una gaussiana. In particolare in queste condizioni di misura (posizione delle telecamere, posizione del calibratore, luminosità dell'ambiente, ecc..) l'errore in fase di acquisizione dei blob di un calibratore è approssimabile mediante una gaussiana con deviazione standard pari a 0.026 pixel. Le stesse prove sono state ripetute per diverse pose del calibratore e hanno portato agli stessi risultati discussi in precedenza.

Una volta stabilita la funzione di densità di probabilità alla quale è affetta la procedura di acquisizione di marker, si è passati alla scrittura di un algoritmo che mediante il metodo Monte Carlo consente di stimare la propagazione dell'incertezza durante la procedura di calibrazione dei parametri intrinseci della

telecamera. In pratica l'algoritmo prende in ingresso 40 diverse pose di un calibratore e ad ognuna delle coordinate dei centri dei marker trovati in ogni immagine viene sommato un valore preso in maniera casuale da una distribuzione normale standard a media nulla e deviazione standard 0.026 pixel. Viene calibrato il sistema stereoscopico con questi dati in ingresso e i dati in uscita (parametri intrinseci, estrinseci e errore 2D) vengono salvati. Questa procedura viene ripetuta per un numero n di volte elevato (nel nostro caso $n=1000$) così da poter avere in uscita n valori di tutti i parametri della telecamera. Da questi valori è possibile stimare la funzione di distribuzione che possono assumere i diversi parametri della calibrazione data una certa distribuzione dei parametri in input (coordinate dei marker).

La lunghezza focale delle telecamere utilizzate in questo caso è circa 33 mm, un'ottica così lunga è praticamente esente da distorsioni ottiche e quindi il modello della telecamera è semplificabile, trascurando la maggior parte dei contributi di distorsione ottica, in particolar modo i contributi di distorsione tangenziale e il secondo e terzo contributo radiale. La procedura di calibrazione di HALCON prevede la possibilità di scegliere due diversi modelli di telecamera: il modello chiamato "*divisionale*" che prevede solamente un parametro K che tiene conto di tutte le distorsioni, e un modello chiamato "*polinomiale*" che invece presenta tre componenti radiali $K1$, $K2$, e $K3$ e due componenti tangenziali $P1$ e $P2$. Il secondo modello è più complesso del primo ma, come verrà mostrato in seguito, non sempre complicare il modello porta a risultati migliori. Sapendo dunque di avere un'ottica performante e con lunghezza focale elevata possiamo decidere di usare per la calibrazione il modello divisionale.

Le distribuzioni trovate mediante Monte Carlo vengono poi sottoposte al test di Kolmogorov-Smirnov per verificare la possibilità di essere approssimate a una distribuzione normale standard. Il risultato di tale test è risultato positivo per tutti i parametri intrinseci. Allo stesso modo anche l'errore di retroproiezione presenta una distribuzione approssimabile mediante una gaussiana. Nella figure seguenti sono riportate le distribuzioni di alcuni dei parametri trovati con le relative distribuzioni cumulative e il risultato del testi di Kolmogorov-Smirnov.

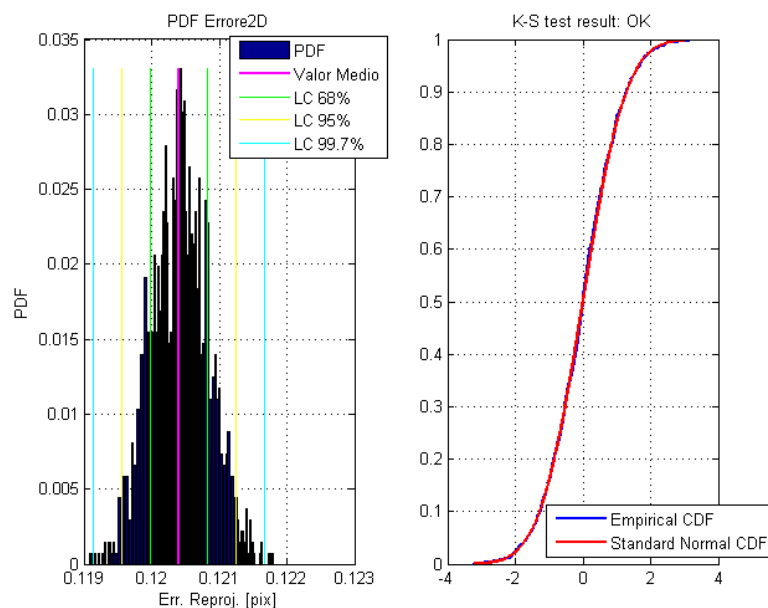


Figura 5-26: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del parametro Errore di Retroproiezione (Errore2D)

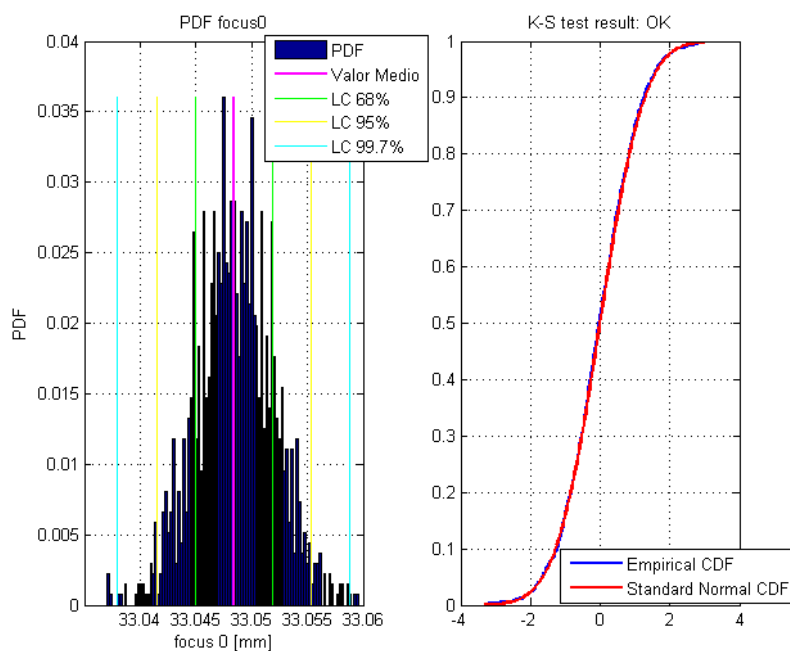


Figura 5-27: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del parametro fuoco della prima telecamera (focus0)

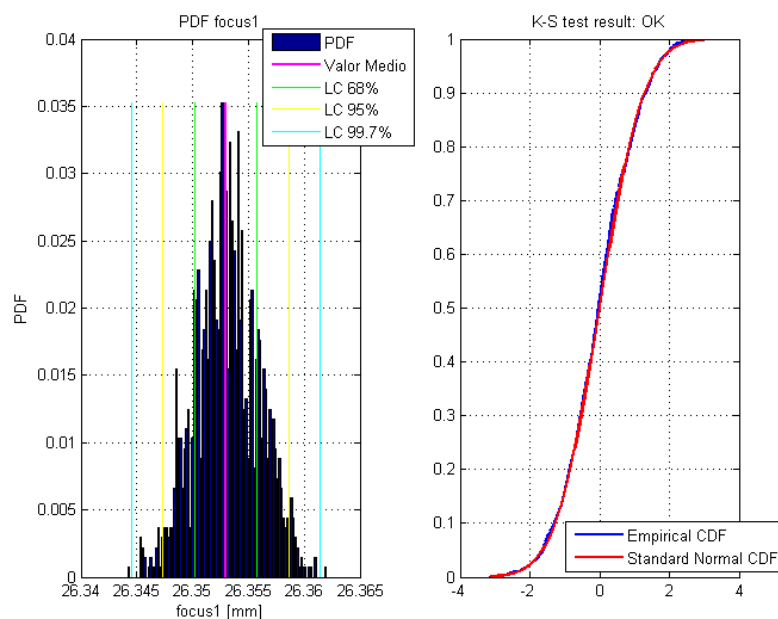


Figura 5-28: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del parametro fuoco della seconda telecamera (*focus1*)

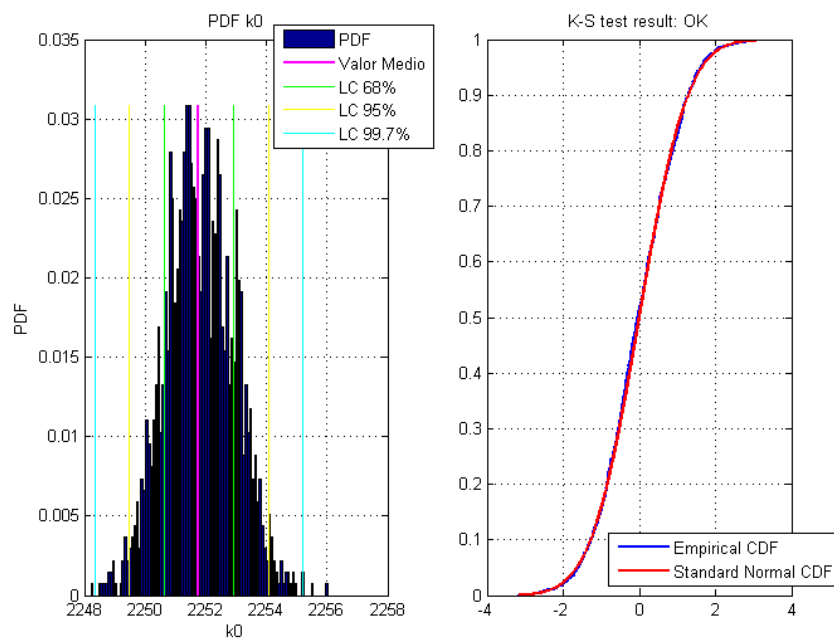


Figura 5-29: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del parametro di distorsione della prima telecamera (*k0*)

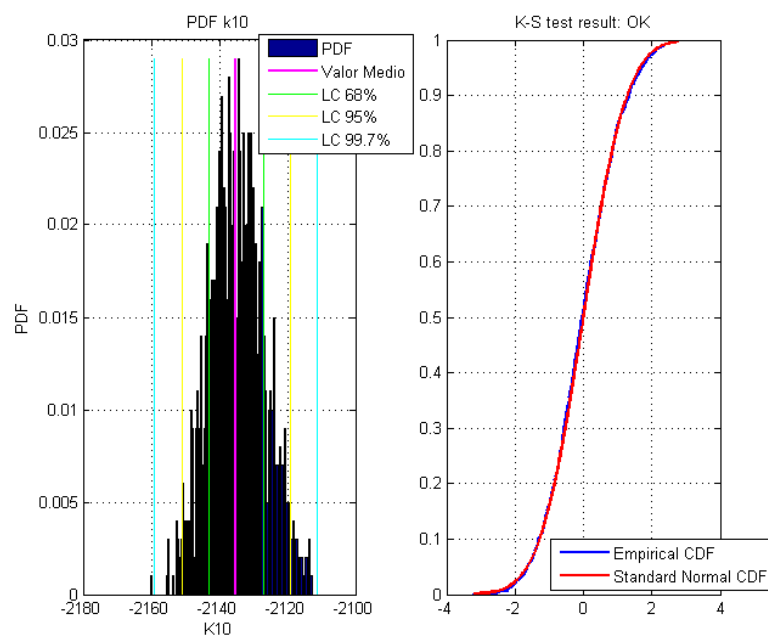


Figura 5-30: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del primo parametro di distorsione radiale della prima telecamera (k_{10})

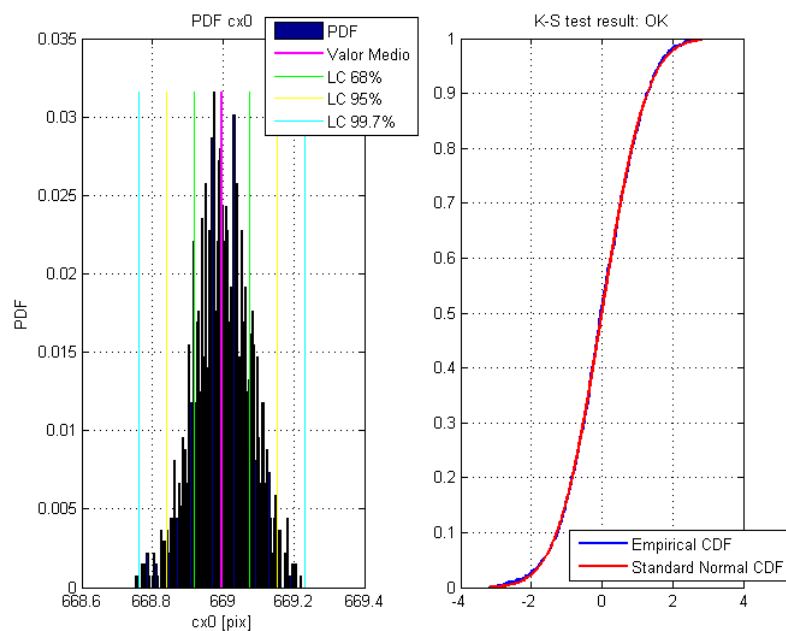


Figura 5-31: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov della coordinata orizzontale del centro ottico della prima telecamera (cx_0)

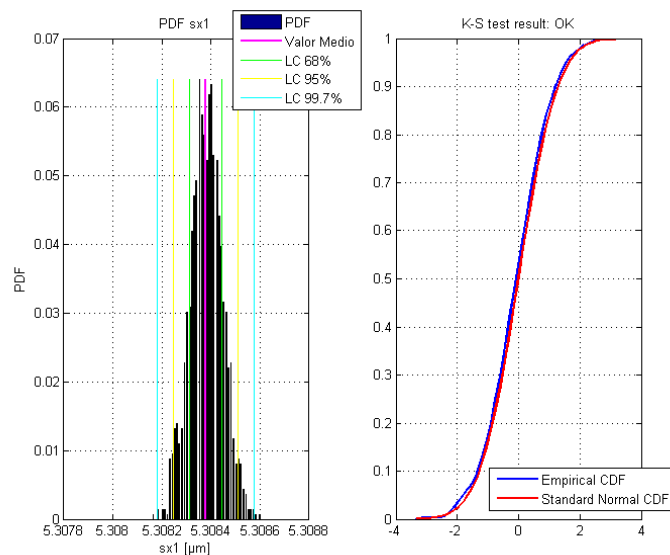


Figura 5-32: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del parametro dimensione orizzontale del sensore della seconda telecamera ($sx1$)

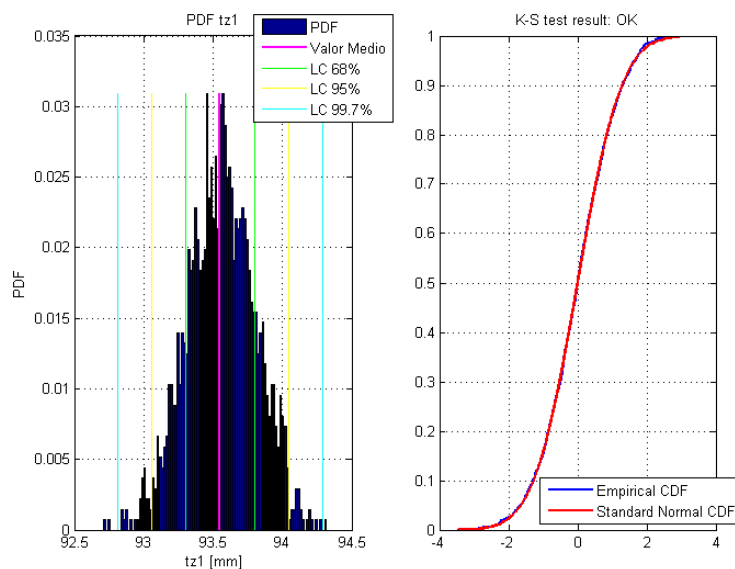


Figura 5-33: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del parametro che definisce la traslazione lungo l'asse z della seconda telecamera rispetto alla prima ($tz1$)

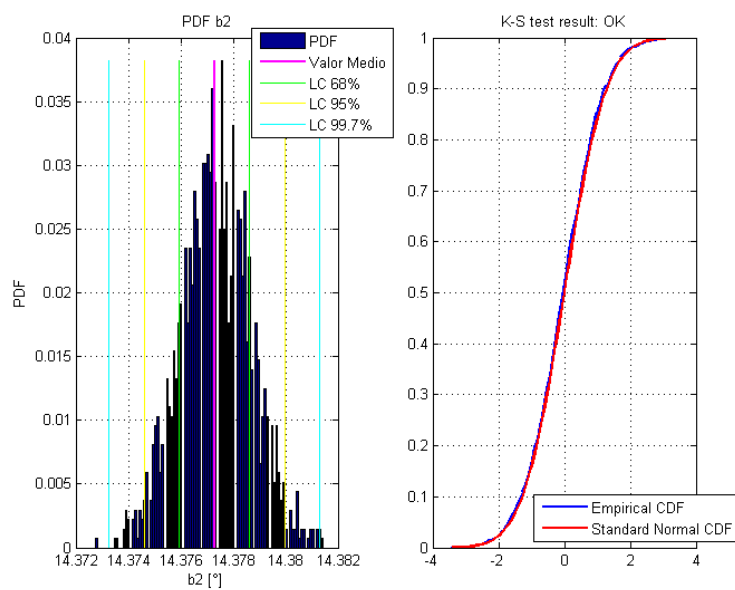


Figura 5-34: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del secondo angolo di Eulero che definisce la rotazione della terza telecamera rispetto alla prima (b_2)

	<i>Parametri</i>	<i>Mean</i>	<i>Std</i>	<i>Test K-S res.</i>
	Reproj. Error [pix]	0.1204	0.0004	OK
INT. IMAGE 0	focus0 [mm]	33.048	0.003	OK
	cx0 [pix]	668.996	0.078	OK
	cy0 [pix]	491.632	0.079	OK
	k0 [1/m ²]	2251.7	1.1	OK
	sx0 [μm]	5.30914	0.00006	OK
INT. IMAGE 1	focus1 [mm]	26.353	0.003	OK
	cx1 [pix]	656.30	0.12	OK
	cy1 [pix]	486.1	0.1	OK
	k1	1855.3	1.9	OK
	sx1 [μm]	5.30838	0.00006	OK
INT. IMAGE 2	focus2 [mm]	30.232	0.003	OK
	cx2 [pix]	657.8	0.1	OK
	cy2 [pix]	507.19	0.09	OK
	k2	2131.0	1.5	OK
	sx2 [μm]	5.30914	0.00007	OK
EXT. IMAGE 1	tx1 [mm]	-429.157	0.035	OK
	ty1 [mm]	-215.98	0.02	OK
	tz1 [mm]	93.5	0.2	OK
	a1 [°]	356.015	0.001	OK
	b1 [°]	7.620	0.001	OK
	g1 [°]	359.6909	0.0003	OK
EXT. IMAGE 1	tx2 [mm]	-903.32	0.08	OK
	ty2 [mm]	-40.644	0.009	OK
	tz2 [mm]	102.72	0.28	OK
	a2 [°]	359.225	0.001	OK
	b2 [°]	14.377	0.001	OK
	g2 [°]	4.3357	0.0003	OK

Tabella 5.1: Medie e deviazioni standard di errore di retroproiezione, parametri intrinseci ed estrinseci delle tre telecamere risultanti dal metodo di Monte Carlo, con relativo risultato del test di Kolmogorov-Smirnov. In questo caso la calibrazione è stata fatta usando un modello divisionale.

I risultati trovati sono mostrati nella tabella 5.1, che riporta l'errore di retroproiezione e i parametri intrinseci ed estrinseci di ogni telecamera con il loro valor medio e la deviazione standard trovati mediante il metodo di Monte Carlo. Inoltre nell'ultima colonna viene riportato per ogni parametro il risultato del test di Kolmogorov-Smirnov. Dai risultati si vede come l'incertezza introdotta in fase di acquisizione dei blob non introduce un'incertezza elevata nei risultati di calibrazione. In particolare l'errore di retroproiezione risultante ha una deviazione standard di 0.0004 pixel, un valore molto basso. Guardando l'incertezza delle tre focali si vede come la loro deviazione standard sia sempre 0.003 mm e assume un valore pari a circa 0.01% del valor medio. Anche gli altri parametri assumono valori di dispersione piuttosto bassi. Da questi risultati possiamo concludere che l'incertezza introdotta in fase di acquisizione delle immagini di calibrazione non porta ad avere incertezze sui parametri intrinseci elevate.

Come spiegato sopra le calibrazioni sono state effettuate utilizzando il modello più semplice, che considera un solo parametro di distorsione. Un errore che spesso si compie è quello di utilizzare un modello più completo per avere un errore di retroproiezione più basso possibile. Così facendo però il modello non rappresenta la realtà; sono presenti alcuni parametri che incidono sull'algoritmo di ottimizzazione durante la calibrazione ma che non hanno un vero e proprio significato fisico, poiché dovrebbero in realtà avere un valore nullo. A dimostrazione di ciò viene eseguito un Monte Carlo, nel quale le calibrazioni sono effettuate utilizzando un modello polinomiale. In questo caso le distribuzioni dei parametri di calibrazione in uscita al modello Monte Carlo non presentano più nella maggior parte dei casi la possibilità di essere approssimate mediante distribuzione gaussiana standard.

In particolare nelle tabella 5.2 e 5.3 sono riportati i risultati di tale prova. Nell'ultima colonna è possibile vedere se la distribuzione trovata per ogni parametro di calibrazione può essere approssimata con una gaussiana standard secondo il metodo di Kolmogorov-Smirnov. Sono riportati anche i valori di media e deviazione standard del campione rilevato, che nei casi in cui la distribuzione non può essere approssimata con una gaussiana non hanno alcuna valenza statistica.

Parametri intrinseci				
	Parametri	Mean	Std	Test K-S res.
	Reproj. Error [pix]	0.1137	0.0005	OK
INT. IMAGE 0	focus0 [mm]	33.0621	0.0044	OK
	cx0 [pix]	650.9	1.9	Not OK
	cy0 [pix]	513.8	1.3	Not OK
	k10 [1/m ²]	-2135	7.9	OK
	k20 [1/m ³]	3.8E+06	1.3E+06	OK
	k30 [1/m ⁴]	-72.4E+10	6.3E+10	OK
	p10 [1/m ²]	0.231	0.025	Not OK
	p20 [1/m ²]	-0.289	0.016	Not OK
	sx0 [μm]	5.3090	0.0001	OK
INT. IMAGE 1	focus1 [mm]	26.357	0.003	OK
	cx1 [pix]	647.5	1.7	Not OK
	cy1 [pix]	496.45	0.97	Not OK
	k11 [1/m ²]	-1816	13	OK
	k21 [1/m ³]	20.1E+06	3.3E+06	OK
	k31 [1/m ⁴]	-28.7E+11	2.4E+11	OK
	p11 [1/m ²]	0.106	0.022	Not OK
	p21 [1/m ²]	-0.120	0.012	Not OK
	sx1 [μm]	5.3086	0.0001	OK
INT. IMAGE 2	focus2 [mm]	30.2193	0.0046	Not OK
	cx2 [pix]	634.5	2.9	Not OK
	cy2 [pix]	526.8	2.1	Not OK
	k12 [1/m ²]	-2030.1	9.9	OK
	k22 [1/m ³]	1.4E+06	1.8E+06	OK
	k32 [1/m ⁴]	-64.7E+10	9.4E+10	OK
	p12 [1/m ²]	0.311	0.037	Not OK
	p22 [1/m ²]	-0.247	0.052	Not OK
sx2 [μm]	5.3091	0.0002	Not OK	

Tabella 5.2: Medie e deviazioni standard di errore di retroproiezione, parametri intrinseci delle tre telecamere risultanti dal metodo di Monte Carlo, con relativo risultato del test di Kolmogorov-Smirnov. In questo caso la calibrazione è stata fatta usando un modello “polinomiale”.

Parametri estrinseci				
	Parametri	Mean	Std	Test K-S res.
EXT. IMAGE 1	tx1 [mm]	-428.89	0.05	OK
	ty1 [mm]	-216.32	0.03	OK
	tz1 [mm]	93.93	0.38	Not OK
	a1 [°]	356.09	0.02	Not OK
	b1 [°]	7.68	0.03	Not OK
	g1 [°]	359.721	0.002	Not OK
EXT. IMAGE 2	tx1 [mm]	-902.51	0.16	Not OK
	ty1 [mm]	-40.98	0.02	Not OK
	tz1 [mm]	107.34	0.81	Not OK
	a1 [°]	359.24	0.05	Not OK
	b1 [°]	14.29	0.04	Not OK
	g1 [°]	4.39	0.01	Not OK

Tabella 5.3: Medie e deviazioni standard di errore di retroproiezione, parametri estrinseci delle tre telecamere risultanti dal metodo di Monte Carlo, con relativo risultato del test di Kolmogorov-Smirnov. In questo caso la calibrazione è stata fatta usando un modello “polinomiale”.

Come si può vedere alcune distribuzioni risultanti, tra le quali quelle rappresentanti le coordinate del centro ottico c_x e c_y , i parametri di distorsione tangenziale p_1 e p_2 , la larghezza del pixel s_x (si ricorda che l'altezza del pixel invece è un parametro fisso e non ottimizzato in HALCON) e i parametri estrinseci, non sono approssimabili con una gaussiana. Vengono riportate a titolo di esempio nelle figure seguenti alcune delle distribuzioni trovate, con la relativa funzione di distribuzione cumulativa e test di Kolmogorov-Smirnov.

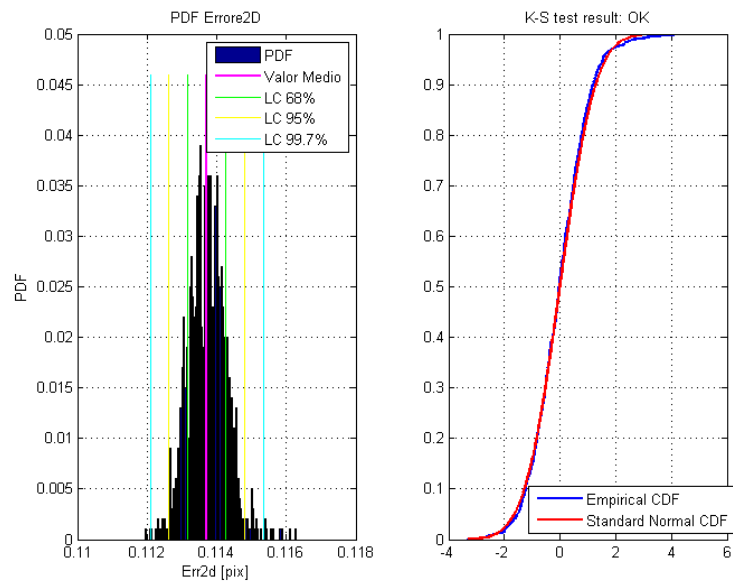


Figura 5-35: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del parametro Errore di Retroproiezione (Errore2D)

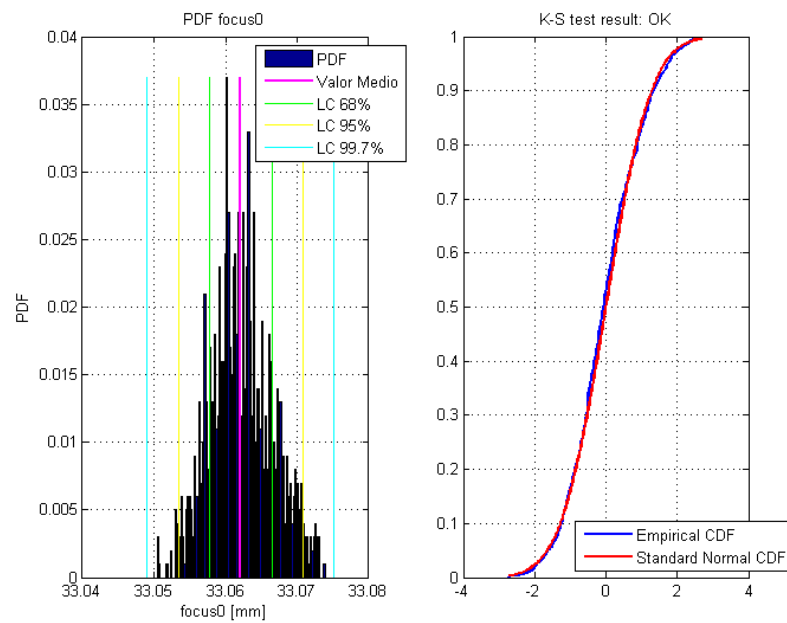


Figura 5-36: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del parametro fuoco della prima telecamera (focus0)

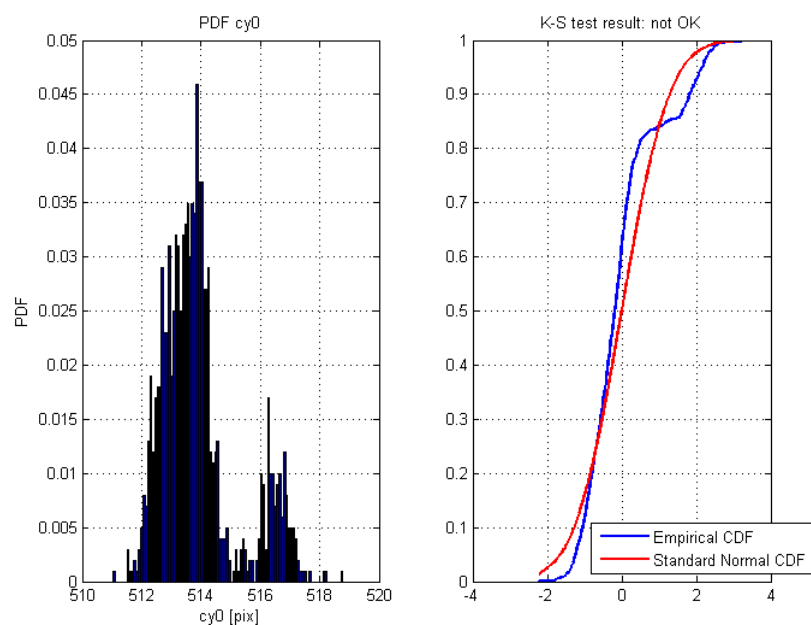


Figura 5-37: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov della coordinata verticale del centro ottico della prima telecamera (cy_0)

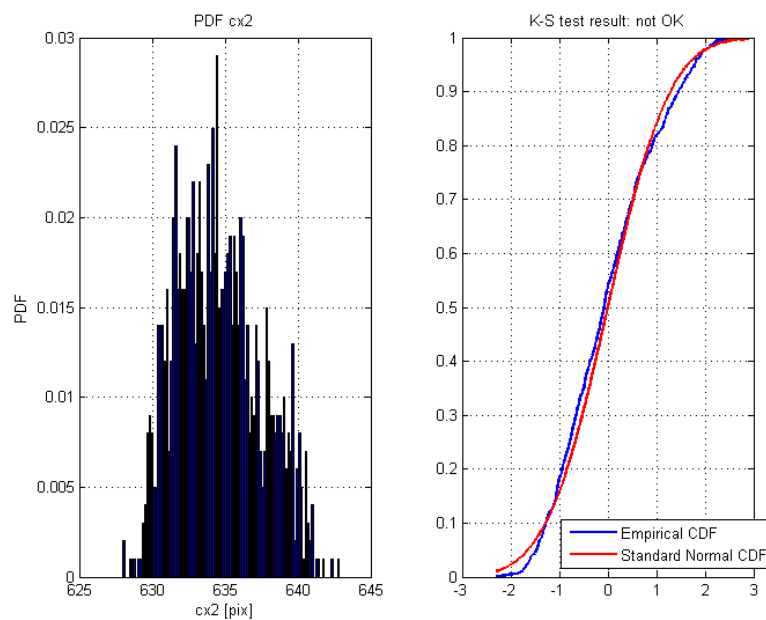


Figura 5-38: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov della coordinata orizzontale del centro ottico della terza telecamera (cx_2).

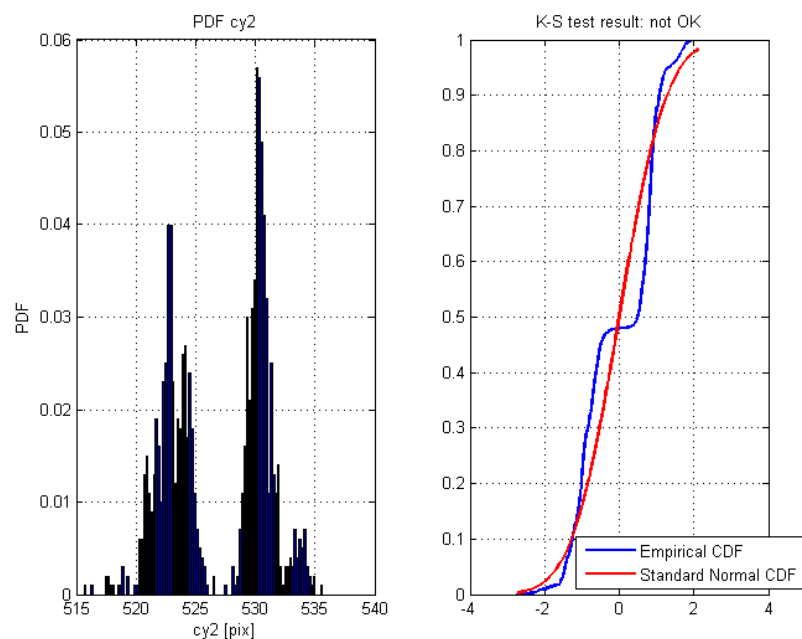


Figura 5-39: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov della coordinata verticale del centro ottico della terza telecamera ($cy2$).

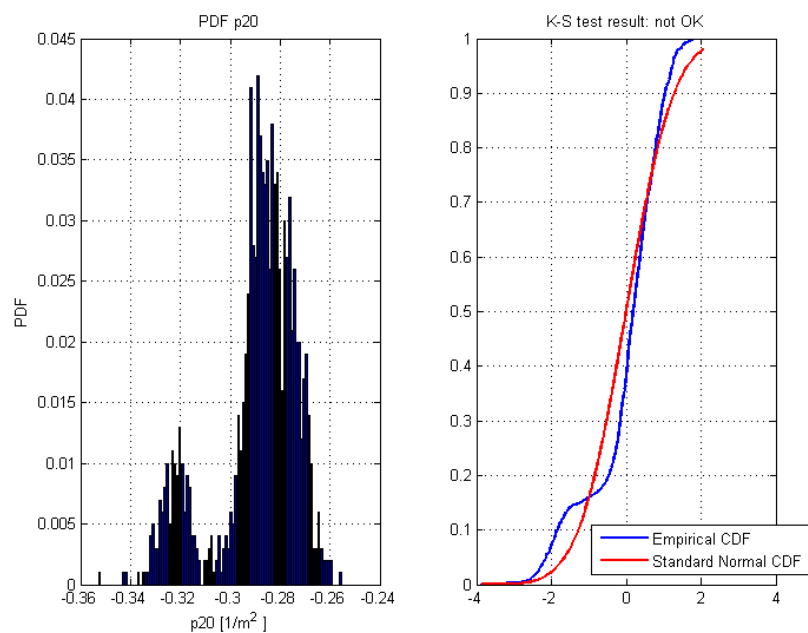


Figura 5-40: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del secondo parametro di distorsione tangenziale della prima telecamera ($p20$)

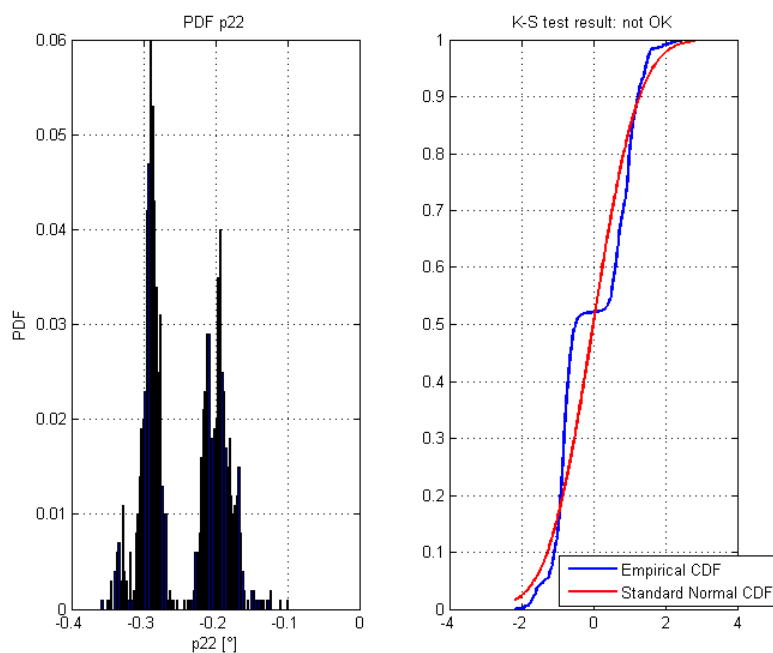


Figura 5-41: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del secondo parametro di distorsione tangenziale della terza telecamera (p_{22})

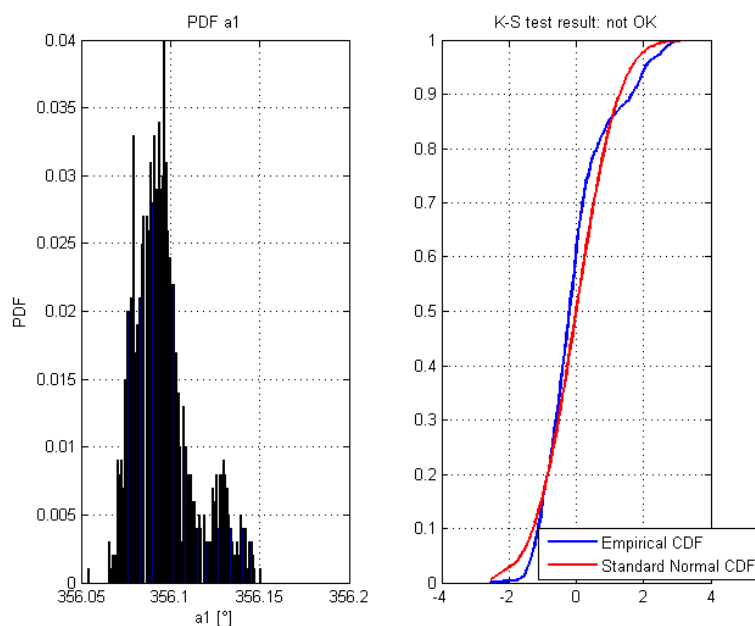


Figura 5-42: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del primo angolo di Eulero che rappresenta la rotazione della seconda telecamera rispetto alla prima (a_1)

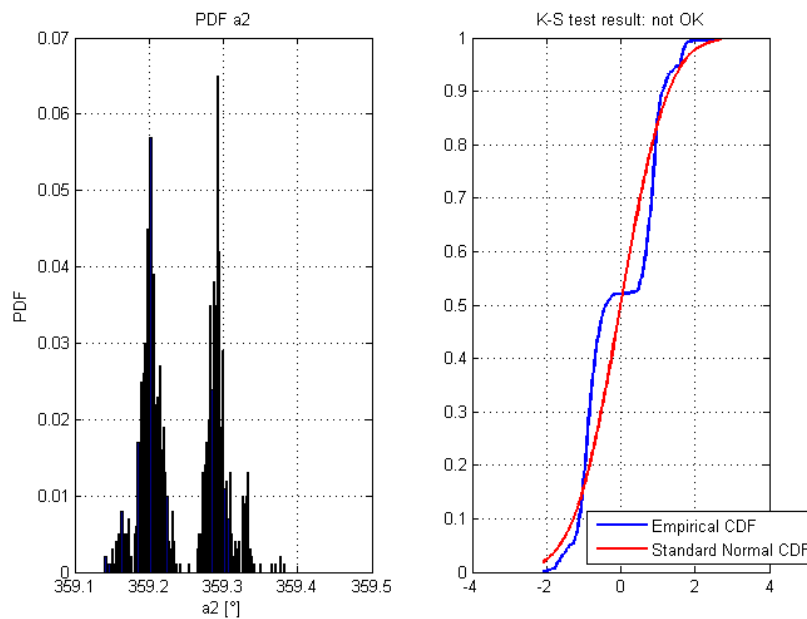


Figura 5-43: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del primo angolo di Eulero che rappresenta la rotazione della terza telecamera rispetto alla prima (a_2)

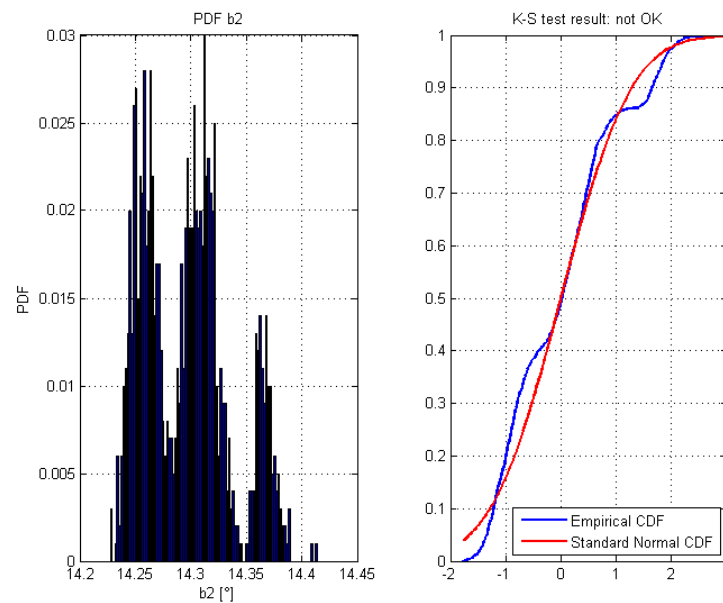


Figura 5-44: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del secondo angolo di Eulero che rappresenta la rotazione della terza telecamera rispetto alla prima (b_2)

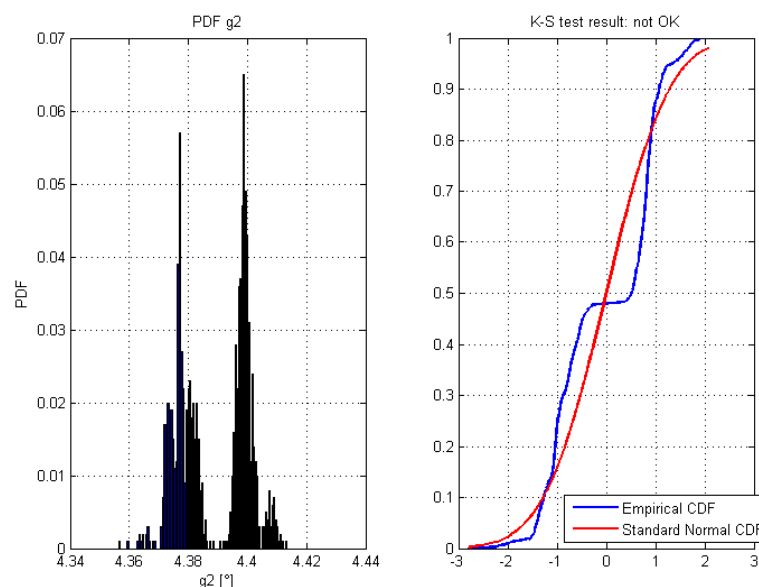


Figura 5-45: Funzione di distribuzione di probabilità (PDF), distribuzione cumulativa (CDF) e risultato del test di Kolmogorov-Smirnov del terzo angolo di Eulero che rappresenta la rotazione della terza telecamera rispetto alla prima (g_2).

In qualche caso esse presentano una forma a “gaussiana a due o tre picchi”, come quelle nelle figure 5.44 e 5.45. Altre invece presentano degli andamenti troppo dispersi per essere approssimabili a una gaussiana, come in figura 5.38. Queste distribuzioni sono dovute a una sovra-parametrizzazione del problema di calibrazione. Si stanno introducendo infatti dei parametri in più, come le distorsioni tangenziali e la seconda e terza radiale, che non rappresentano fedelmente la realtà in quanto potrebbero essere trascurati utilizzando come in questo caso un’ottica a focale lunga. L’algoritmo di minimizzazione presente in calibrazione assegna a questi parametri dei valori e, spostando il centro ottico delle varie immagini, minimizza l’errore di retroproiezione medio. L’algoritmo riesce così a raggiungere minimi locali diversi tra di loro, in cui i parametri perdono la loro fisicità e sono compensati dallo spostamento del centro ottico. I risultati mostrano distribuzioni gaussiane con più picchi in quanto i minimi locali possibili che l’algoritmo raggiunge potrebbero essere più di uno.

In questo paragrafo sono stati riportati per brevità e leggibilità solamente i risultati ottenuti dalla prova descritta sopra. Essa è stata ritenuta più significativa

delle altre in quanto il sistema stereoscopico presenta una configurazione considerata standard per l'utilizzo in Galleria del Vento. Altre prove con configurazioni differenti sono state effettuate ottenendo risultati del tutto riconducibili a quelli mostrati sopra. In particolare modificando il set di immagini utilizzate per la calibrazione i risultati sono molto simili a quelli sopra, con valor medi e deviazioni standard dei parametri di calibrazione compatibili con quelli in tabella 5.1. Modificando la posizione delle telecamere è stato necessario modificare leggermente la focale dell'ottica e dunque i valor medi trovati sono differenti, specialmente i valori relativi alle focali e ovviamente gli estrinseci. Però confrontando le deviazioni standard delle focali esse comunque assumono un valore che è sempre inferiore al 0.05% del valore medio.

Da queste prove si può dunque concludere che l'incertezza introdotta dall'algoritmo di acquisizione dei centri dei blob dà origine in fase di taratura del sistema stereoscopico a incertezze sui parametri di calibrazione molto basse. Per verificare il fatto che questi valori di incertezza sui parametri di calibrazione possano essere considerati minimi viene eseguita un'analisi di sensitività del processo di ricostruzione tridimensionale nel prossimo paragrafo.

5.7 Analisi di sensitività

Grazie all'analisi della propagazione dell'incertezza effettuata nel paragrafo precedente, è stato possibile stimare le incertezze che si hanno in fase di taratura del sistema stereoscopico, dovute al propagarsi dell'incertezza di acquisizione dei blob del calibratore. Per verificare che i valori di incertezza trovati nel paragrafo precedente possano essere considerati trascurabili viene qui effettuata un'analisi di sensitività. Questa analisi consiste nel valutare gli effetti sui risultati forniti da un modello (ossia dalla funzione che lo descrive analiticamente) indotti da modifiche nei valori delle variabili di ingresso. Quindi nel nostro caso verranno modificati alcuni parametri di calibrazione e verranno effettuate alcune ricostruzioni tridimensionali di pose di un calibratore planare. Così facendo si potrà valutare lo

scostamento dei valori delle coordinate che avviene modificando i parametri di calibrazione.

I valori che i parametri di calibrazione potrebbero assumere sono compresi tra un valore minimo e uno massimo:

$$\mu_x - 3\sigma_x < x < \mu_x + 3\sigma_x$$

Dove μ_x è il valor medio del parametro e σ_x è la sua deviazione standard. Si è deciso di utilizzare un intervallo di $\pm 3\sigma_x$ poiché in tale intervallo ricadono il 99.7% delle probabilità dei possibili valori che il parametro può assumere. Si riporta qui a titolo esemplificativo un'analisi di sensitività effettuata con il sistema stereoscopico nelle stesse condizioni della prova descritta nel paragrafo precedente. Dunque i parametri di media e deviazione standard sono deducibili dalla tabella 5.1.

Come prima prova viene modificato il valore della focale della prima telecamera (“*focus0*”). In particolare esso potrà assumere valori compresi tra x_{min} e x_{max} così definiti:

$$x_{min} = 33.048 - 3 * 0.003 = 33.039 \text{ mm}$$

$$x_{max} = 33.048 + 3 * 0.003 = 33.057 \text{ mm}$$

Vengono dunque imposti in ingresso al modello di ricostruzione tridimensionale 13 valori di focale della prima telecamera compresi tra x_{min} e x_{max} intervallati di un valore pari a metà della deviazione standard. Per ogni valore imposto di focale vengono ricostruite 30 pose differenti del calibratore. Le coordinate tridimensionali ricostruite vengono sottratte del valore che assumono con la focale della prima telecamera imposta al suo valor medio (33.048mm) e viene calcolato il valore assoluto di tale differenza, trovando così il valore di scostamento delle coordinate 3D dal valore atteso. Per ognuna delle 30 pose poi viene effettuata la media degli scostamenti di ciascuna coordinata. Così facendo è possibile rappresentare l'andamento degli scostamenti delle coordinate tridimensionali ricostruite al variare del parametro di focale della prima telecamera. Il risultato è mostrato in figura 5.46.

Per prima cosa è possibile osservare che la coordinata z (quella che rappresenta la profondità, quindi perpendicolare al sensore della telecamera) raggiunge il valore

di scostamento maggiore. Questo dimostra che la coordinata z è quella che presenta il maggior contributo di incertezza. Variando anche di poco un parametro della calibrazione, essa varia in maniera più significativa rispetto alle altre. È possibile inoltre osservare che lo scostamento massimo raggiunto nel peggiore dei casi è inferiore al decimo di millimetro, che è comunque uno scostamento piuttosto modesto. Questo implica che l'incertezza sulla focale della prima telecamera stimata mediante il metodo Monte Carlo produce una diminuzione dell'accuratezza delle misure stimabile nel case peggiore in uno scostamento inferiore al decimo di millimetro.

Ripetendo la prova intervenendo sui valori di focale delle altre due telecamere è possibile ottenere i risultati di figura 5.47 (per quanto riguarda la focale della seconda telecamera) e figura 5.48 (per la focale della terza telecamera). Si vede dal confronto tra i tre grafici come la seconda telecamera presenti scostamenti massimi della coordinata z inferiori rispetto alle altre, mentre gli scostamenti delle coordinate x e y sono tra loro compatibili. Questo è dovuto al fatto che la telecamera centrale, a differenza delle altre due, presenta una posizione frontale rispetto alla scena, ossia che gli assi x e y sono fondamentalmente paralleli al piano del sensore mentre l'asse z è perpendicolare ad esso. Per questo la telecamera centrale ha sensibilità massima sulla ricostruzione tridimensionale lungo le coordinate x e y e minima lungo z . Le due telecamere laterali invece hanno una sensibilità in z maggiore di quella centrale. Gli scostamenti lungo la coordinata z sono dunque maggiori rispetto alla telecamera centrale; gli scostamenti lungo x e y sono invece compatibili tra tutte le telecamere.

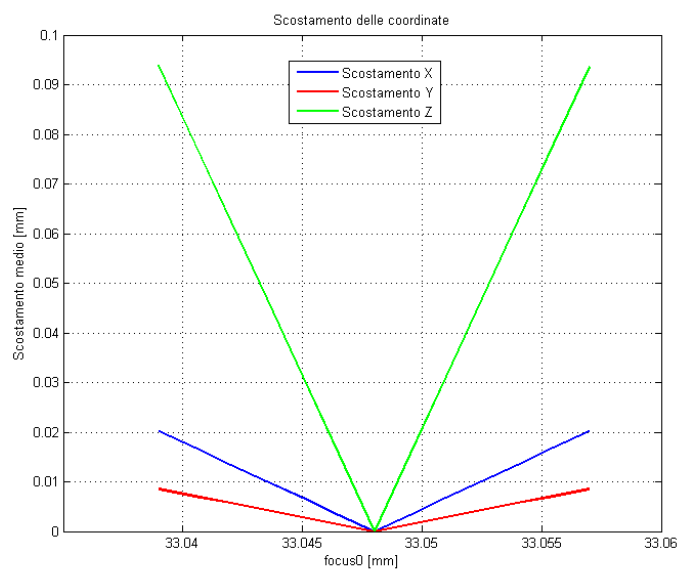


Figura 5-46: Scostamento medio del valore delle coordinate ricostruite al variare della focale della prima telecamera (focus0)

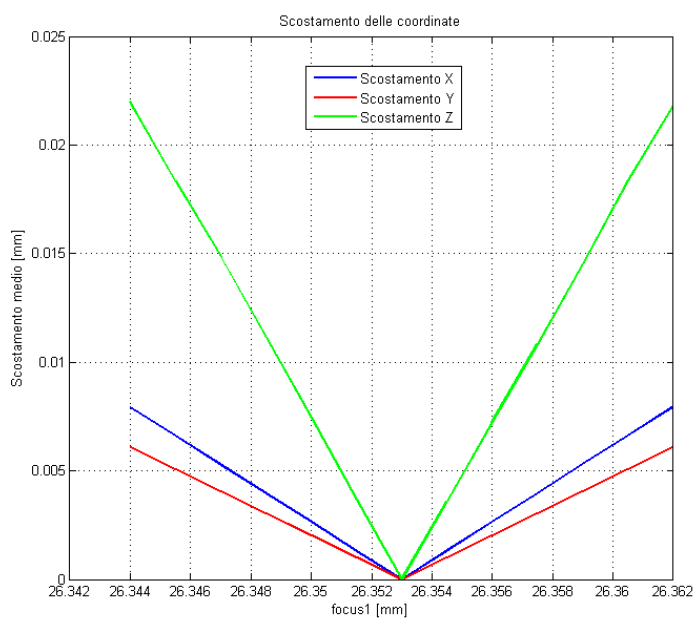


Figura 5-47: Scostamento medio del valore delle coordinate ricostruite al variare della focale della seconda telecamera (focus1)

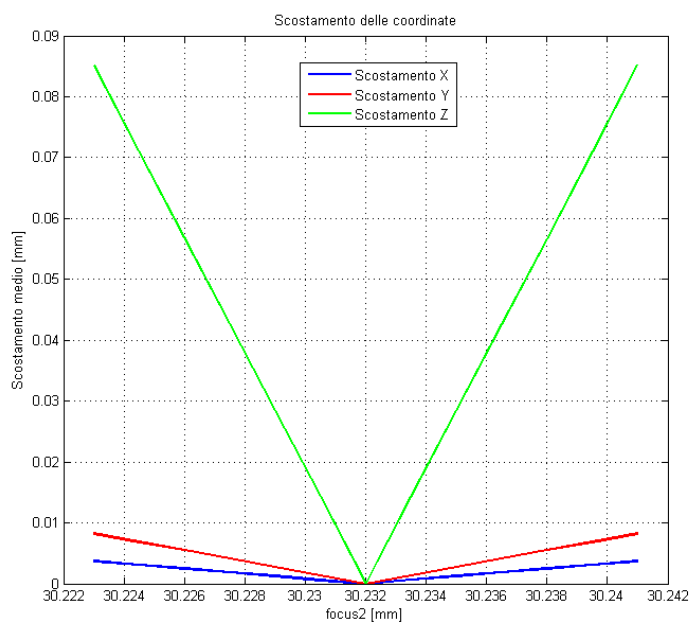


Figura 5-48: *Scostamento medio del valore delle coordinate ricostruite al variare della focale della terza telecamera (focus2)*

Le prove effettuate sulle focali sono state ripetute per tutti i parametri delle tre telecamere. Per questioni di spazio e leggibilità vengono riportati nei grafici di seguito solamente alcuni dei risultati più significativi.

In particolare nella figura 5.49 sono rappresentati i risultati relativi all'analisi di sensitività effettuata modificando il valore del centro ottico orizzontale della seconda telecamera, mentre i risultati nella figura 5.50 sono stati trovati modificando il centro ottico verticale della stessa. Si nota come la coordinata z presenta gli scostamenti maggiori mentre le coordinate x e y presentano valori molto diversi a seconda del parametro che viene modificato. Se viene modificata la coordinata del centro ottico in direzione orizzontale (cx_0) allora la coordinata orizzontale x avrà scostamenti maggiori di y; se invece viene modificata la coordinata verticale del centro ottico (cy_0), allora sarà la coordinata verticale y ad avere scostamenti maggiori. Lo stesso vale anche per le altre due telecamere.

Dall'analisi di sensitività si è visto che i parametri intrinseci danno il maggior contributo di incertezza nelle misure, visto che gli scostamenti trovati raggiungono in qualche caso valori fino a tre decimi di millimetro. Per quanto riguarda gli

estrinseci invece è stato notato che modificando i parametri di traslazione e rotazione gli scostamenti variano sempre meno di mezzo decimo di millimetro, quindi sono molto bassi. A titolo esemplificativo vengono riportati i grafici dell'analisi di sensitività svolta modificando il parametro di traslazione lungo x della prima telecamera e il secondo angolo di Eulero che definisce la rotazione della seconda telecamera rispetto alla prima.

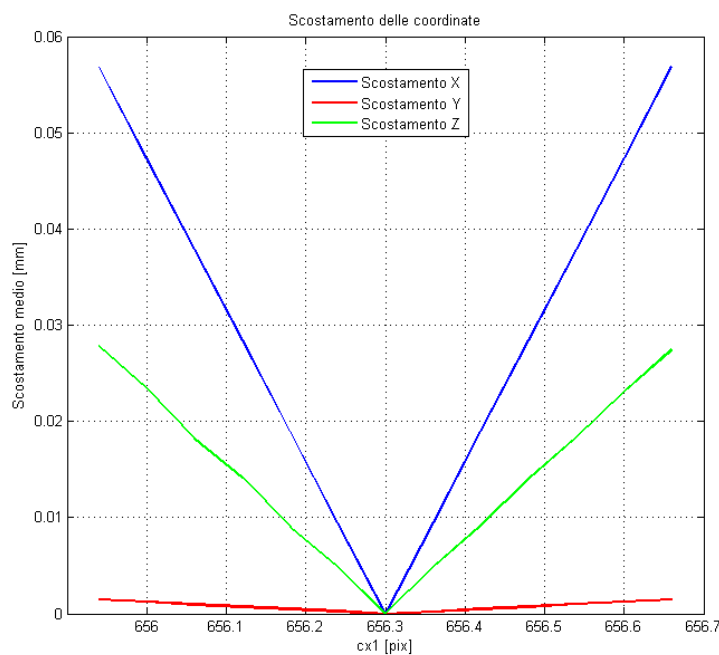


Figura 5-49: *Scostamento medio del valore delle coordinate ricostruite al variare della coordinata orizzontale del centro ottico della seconda telecamera (cx1)*

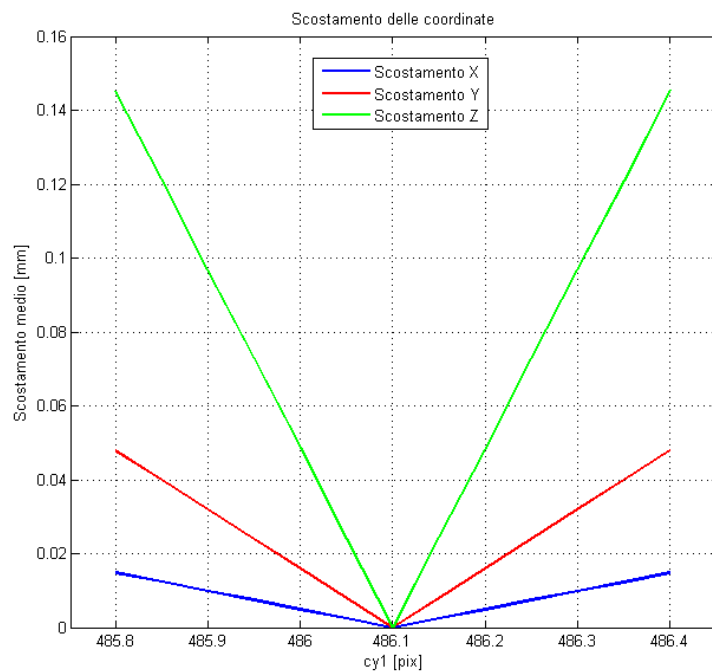


Figura 5-50: Scostamento medio del valore delle coordinate ricostruite al variare della coordinata orizzontale del centro ottico della seconda telecamera ($cy1$)

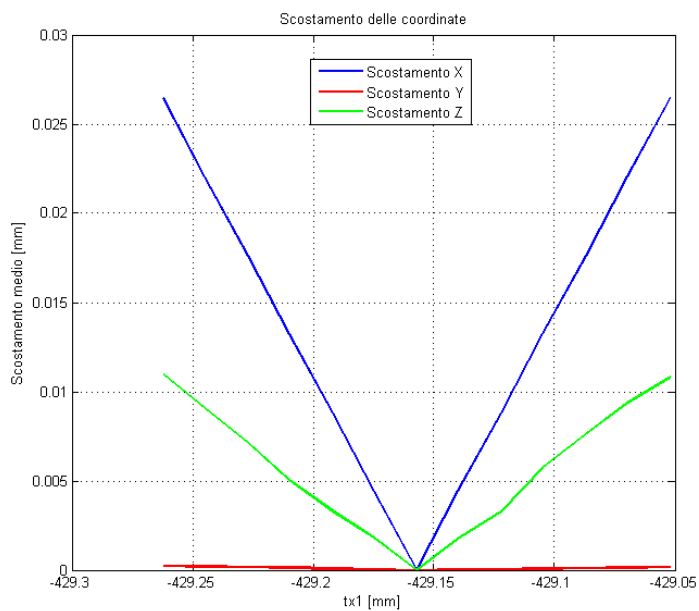


Figura 5-51: Scostamento medio del valore delle coordinate ricostruite al variare del parametro estrinseco di traslazione della seconda telecamera ($tx1$)

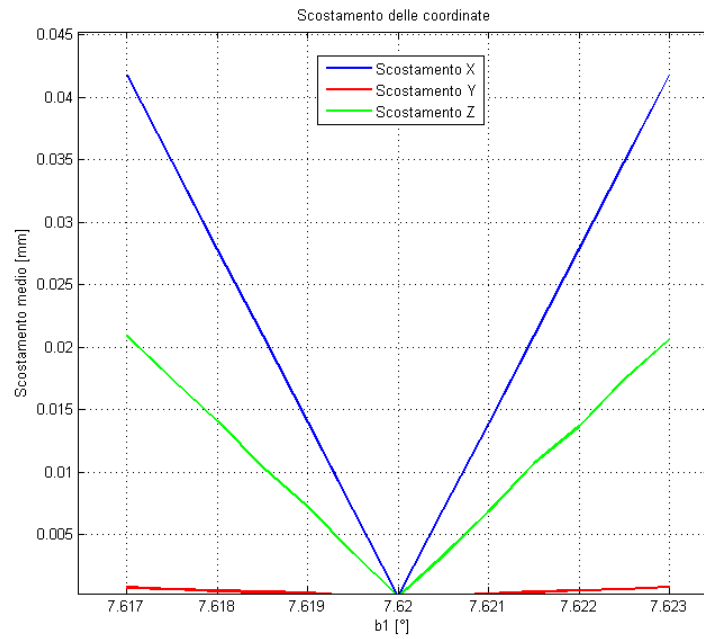


Figura 5-52: *Scostamento medio del valore delle coordinate ricostruite al variare del secondo angolo di Eulero che definisce la rotazione della seconda telecamera rispetto alla prima (b1)*

6. Applicazioni innovative

In questo capitolo vengono riportate le tre principali applicazioni innovative che si sono sviluppate in questo lavoro. Nonostante la diversità dei problemi affrontati, essi sono stati risolti mediante lo sviluppo di sistemi ottici stereoscopici. In particolare i casi studio considerati sono: misure di spostamento in galleria del vento, robot decoratore e bin-picking. La prima applicazione consiste nel misurare gli spostamenti di alcuni punti notevoli di oggetti sottoposti alla vena fluida del vento. Nella seconda un sistema stereoscopico viene utilizzato per riprendere posizione e orientamento di un tool nelle mani di un decoratore in modo da far ripetere delle traiettorie a un robot antropomorfo. La terza invece consiste nella presa di pezzi posti in un cassone alla rinfusa da parte di un robot. Tutte le soluzioni sono state validate con opportuni esperimenti e i risultati principali sono riportati in questo capitolo.

6.1 Misure in Galleria del Vento

Questa applicazione nasce dalla necessità di Galleria del Vento del Politecnico di Milano di avere uno strumento di misura di spostamenti 3D e forme tridimensionali senza contatto. In questa applicazione è indispensabile poter minimizzare l'effetto di carico sulla struttura in esame sottoposta a vena fluida del vento, in modo da non modificare le sue caratteristiche fisiche. Risulta dunque importante l'utilizzo di un sistema di misura ottico non a contatto.

In questa applicazione è di fondamentale importanza utilizzare un'illuminazione controllata; in questo modo le misure non risultano influenzate dalla luce ambiente ed è possibile effettuare misurazioni in qualsiasi condizione luminosa. Come viene descritto nel seguito con il sistema realizzato si potrà misurare dalla forma di vele di modelli di imbarcazioni agli spostamenti di modelli di grosse strutture civili in scala. Le applicazioni sono dunque molto diverse tra loro; questo comporta che il misurando possa assumere varie dimensioni e il posizionamento delle telecamere dipenda fortemente dagli ingombri e dal tipo di prova effettuato. Da queste considerazioni consegue che il sistema deve avere caratteristiche di flessibilità e scalabilità. Questo giustifica la scelta della stereoscopia come tecnica di ricostruzione tridimensionale.

Nel seguito, dopo la descrizione del setup del sistema realizzato, verranno riportate brevemente alcune delle campagne sperimentali effettuate in Galleria del Vento.

6.1.1 Setup del sistema

Nelle applicazioni di Galleria del Vento le telecamere non possono disturbare la vena fluida del vento e quindi il corretto posizionamento di esse è dettato dal miglior compromesso tra inquadratura dell'oggetto da misurare, massimizzazione dell'accuratezza del sistema, facilità di montaggio, assenza di disturbi aereodinamici. In alcune particolari acquisizioni questo lavoro non è sempre facile e comporta uno studio molto dettagliato del setup. Il sistema utilizzato per questa

applicazione è quello descritto nel capitolo 5. Vengono utilizzate le stesse telecamere ma con l'aggiunta di un sistema di illuminazione controllata.

Più l'ambiente è controllato nelle sue caratteristiche più le prove e le misure effettuate saranno ripetibili e meno soggette ad errori. Anche se il sistema di misura può essere usato in condizioni di luce naturale, in alcune applicazioni può essere richiesta robustezza e accuratezze elevate ed è dunque opportuno prevedere un sistema di illuminazione controllata.

Si è dunque deciso di utilizzare tre illuminatori infrarossi coassiali rispetto all'ottica delle telecamere e di posizionare marker catarifrangenti sull'oggetto da misurare. Viene utilizzato un filtro interferenziale centrato sulla lunghezza d'onda dell'illuminatore in modo tale da tagliare le lunghezze d'onda indesiderate. Esso viene posto sull'ottica delle telecamere così da escludere il campo del visibile mettendo in evidenza solamente i marker posti sul misurando. Questa tecnica consente prima di tutto di limitare notevolmente i disturbi causati dalla luce naturale che possono rendere più difficile l'individuazione dei marker. Inoltre questo metodo può essere utilizzato in qualunque applicazione, ponendo sul misurando un numero di marker opportuno. Il numero e la posizione di marker deve essere scelta in modo intelligente e dipende fortemente dalla struttura del misurando e da ciò che si vuole misurare.

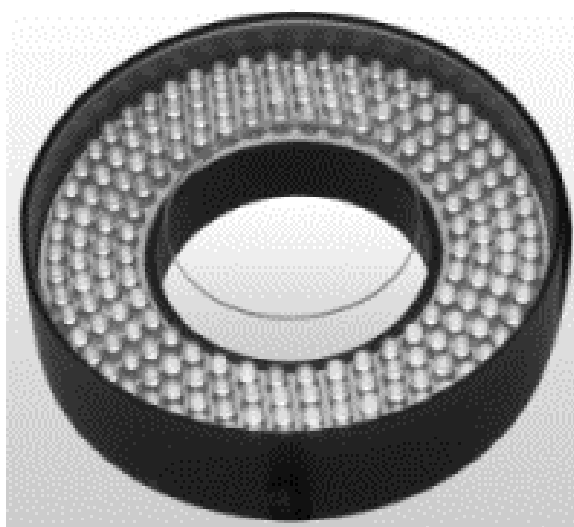


Figura 6-1: *Illuminatore circolare BMT.*



Figura 6-2: *Controllore Gardasoft PP420, utilizzato per la gestione degli illuminatori.*

Il sistema di misura è provvisto dunque di tre illuminatori coassiali BMT modello BDRL-IR110/58 a lunghezza d'onda di 850 nm (figura 6.1), controllati da Gardasoft PP420 (figura 6.2) che consente di gestire separatamente i tre illuminatori. In particolare il controllore, comandato via Ethernet, permette di utilizzare gli illuminatori in maniera continua, con un trigger software o con un trigger esterno. Utilizzando i tre illuminatori in modalità trigger esterno, con lo stesso trigger che comanda anche l'acquisizione delle telecamere si possono sincronizzare i sei dispositivi. Inoltre il controllore consente di impostare il livello di luminosità e l'intervallo in cui l'illuminatore rimane acceso dopo l'arrivo del segnale di trigger.

Inoltre vengono utilizzati sulle ottiche filtri interferenziali passa banda con banda passante centrata intorno a 850 nm. In questo modo si limitano i disturbi, eliminando la luce a lunghezza d'onda differente da quella degli illuminatori.

6.1.2 Taratura del sistema

Per quanto riguarda la calibrazione delle telecamere, in Galleria del Vento è stata utilizzata la tecnica di Heikkilä, implementata nelle librerie di HALCON (paragrafo 4.2). In ogni campagna sperimentale prima di effettuare le misure dell'oggetto, sono state effettuate delle prove atte a verificare la taratura del sistema. Come è stato discusso nel paragrafo 5.3, l'errore di retroproiezione è un buon indice

della bontà della calibrazione ma non sempre definisce l'accuratezza dello strumento in termini di errore di ricostruzione tridimensionale. Si è voluto dunque valutare l'accuratezza del sistema prima di ogni campagna sperimentale. Per fare ciò è possibile ad esempio ricostruire un calibratore a blob catarifrangenti, di cui si conoscono le distanze tra i marker con accuratezza elevata. In figura 6.3 sono mostrate le vedute di tale calibratore dalle tre diverse telecamere durante alcune prove. Con l'algoritmo di blob detection vengono trovati i marker del calibratore e ad ognuno di essi viene assegnato un numero, in modo da identificarlo rispetto agli altri, come mostrato in figura 6.4.

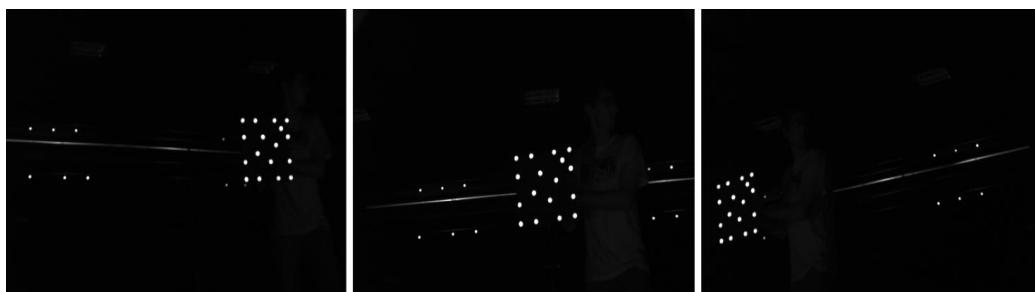


Figura 6-3: Esempio di immagini acquisite per la ricostruzione del calibratore infrarossi.

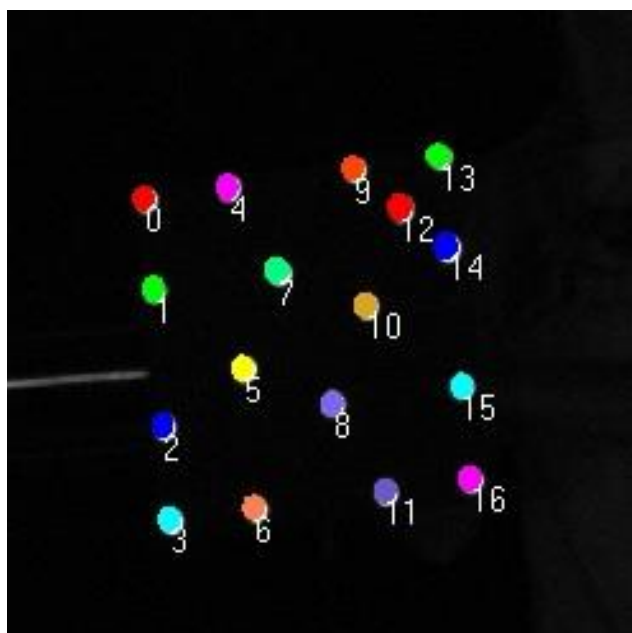


Figura 6-4: Numerazione assegnata ai marker del calibratore infrarossi dall'algoritmo di blob detection.

Vengono quindi calcolate le coordinate tridimensionali dei blob del calibratore. Un buon metodo per stimare la bontà della ricostruzione (e dunque anche della calibrazione), è il confronto fra le distanze tra i blob e le rispettive distanze reali. Alcuni risultati relativi a una di queste prove sono riportati nella tabella 6.1.

DISTANZA PUNTI	Valore Reale [mm]	Valore da ricostruzione [mm]	Scostamento Abs [mm]
0 - 4	80	80.1495	0.1495
4 - 9	120	120.978	0.978
9 -13	80	80.3641	0.3641
9 - 12	56.56854	55.6174	0.95114
9 -14	113.137	112.38	0.757
0 - 3	280	280.637	0.637
0 - 16	395.9798	395.069	0.9108
0 - 1	80	80.9197	0.9197
13 - 16	280	279.151	0.849
15 - 16	80	80.1821	0.1821

Tabella 6.1: Confronto tra le distanze tra alcuni marker trovate dalla ricostruzione e le distanze reali.

Dalla tabella 6.1 si nota che lo scostamento dalle distanze reali non supera mai il millimetro. Questa prova quindi mostra che in questo caso lo strumento è calibrato e l'accuratezza trovata ricostruendo il calibratore è accettabile per l'applicazione (inferiore al millimetro).

Nel caso in cui l'oggetto in esame abbia una struttura rigida è anche possibile verificare il rumore della misura effettuando un'acquisizione del misurando in condizione stazionaria. A titolo esemplificativo vengono nel seguito riportati i risultati di una di queste prove svolte su una struttura civile in condizioni stazionarie. Vengono qui riportati gli spostamenti nelle tre dimensioni di alcuni marker posti sull'oggetto fermo (figura 6.5 e 6.6).

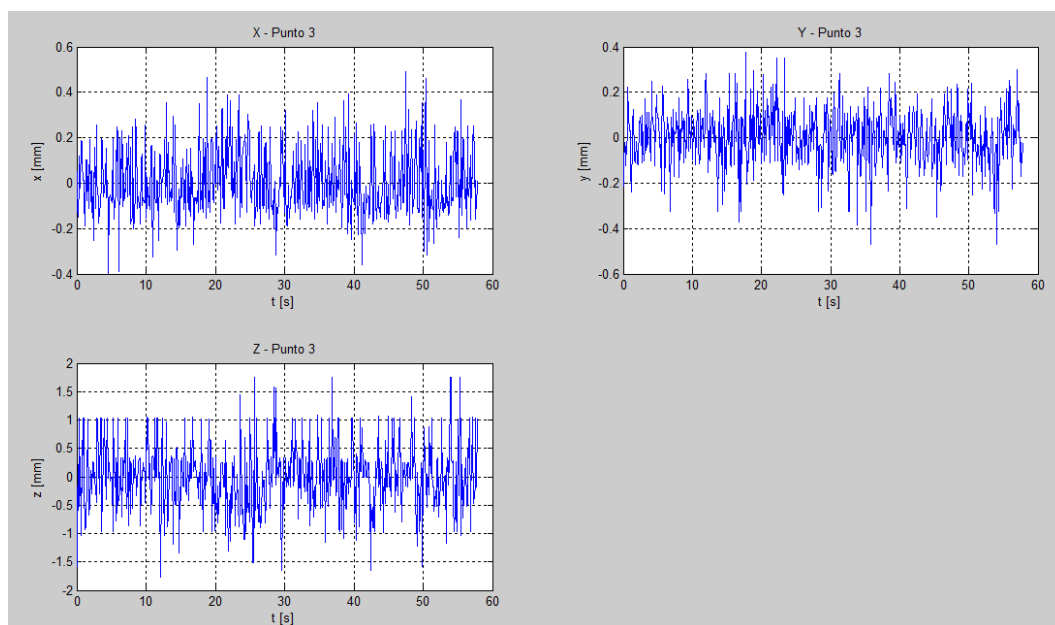


Figura 6-5: *Spostamento lungo i tre assi del punto 3.*

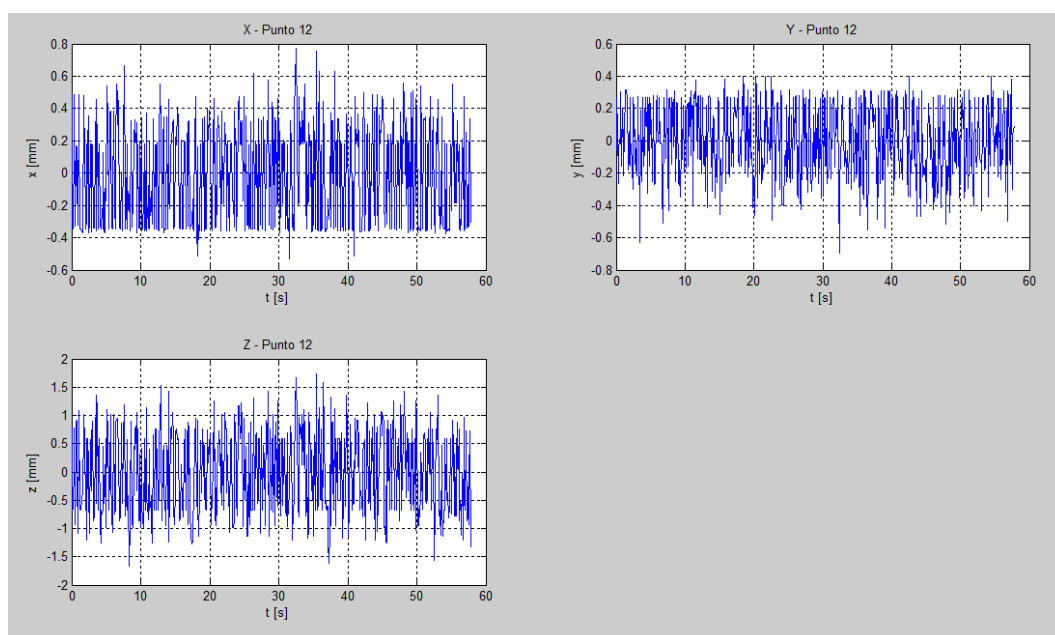


Figura 6-6: *Spostamento lungo i tre assi del punto 12.*

Dall'analisi degli spostamenti si nota che in queste prove le oscillazioni massime su x e y sono nell'ordine di ± 0.5 mm, mentre quelle su z , sul quale si hanno sempre scostamenti maggiori, sono dell'ordine di ± 1.5 mm. L'incertezza maggiore in questo caso è quindi presente sull'asse z ed è inferiore a ± 1.5 mm.

6.1.3 Misura della forma di vele

Presso la Galleria del Vento del Politecnico di Milano vengono spesso provate vele per imbarcazioni con lo scopo di capire come alcune regolazioni di esse possano influire sulla forza di spinta dell'imbarcazione a diverse angolazioni e velocità di vento. Tali vele vengono provate su un modellino in scala come quello riportato nella figura 6.7. Esso è posizionato su una bilancia a sei assi che misura le forze scambiate a terra dall'imbarcazione. Questa bilancia è posizionata su un tavolo rotante che dà la possibilità di simulare diverse angolazioni di vento.



Figura 6-7: Misura della forma della randa di un modello di imbarcazione a vela presso la Galleria del vento

Fino ad oggi l'unica conoscenza relativa alla forma della vela viene dal velaio che attraverso dei software di simulazione realizza quella che viene chiamata *design shape*, vale a dire una forma teorica tridimensionale della possibile forma assunta dalla vela in navigazione. La *design shape* però rappresenta un caso particolare delle infinite possibili configurazioni che la vela può assumere nella realtà durante la navigazione. La forma della vela reale viene chiamata *flying shape* e a differenza della prima, essa è strettamente correlata alle performance dell'imbarcazione. Per questo è indispensabile conoscerne le caratteristiche. La combinazione di una serie di parametri tra cui forza e direzione del vento, regolazione delle vele, interferenze

di altri componenti sulla vena fluida del vento, porta ad avere un infinito numero di diverse possibili flying shape per la stessa vela.

Il sistema stereoscopico realizzato è stato utilizzato in alcune campagne sperimentali per la ricostruzione tridimensionale della forma di vele di un modello di imbarcazione presso la Galleria del Vento del Politecnico di Milano. In particolare vengono qui riportati alcuni risultati di una campagna sperimentale effettuata con North Sails e commissionata dal Offshore Racing Congress (ORC). In questa campagna sperimentale la forma geometrica della randa è stata acquisita con il sistema stereoscopico.

Un approccio basato sulla stereoscopia è sicuramente molto accurato e consente di ricavare informazioni sulla forma delle vele anche in tempo reale. Se i blob vengono disposti sulle linee di forma i parametri geometrici di interesse possono essere calcolati in tempo reale. I grossi svantaggi di questo metodo sono dati dalle occlusioni che possono esserci quando viene misurato un oggetto non propriamente rigido come possono essere le vele per andature portanti, quali gennaker e spinnaker. Per risolvere questa problematica bisognerebbe complicare notevolmente il sistema introducendo diverse telecamere. Inoltre il sistema acquisisce solamente i punti notevoli definiti dai marker catarifrangenti, utili per calcolare i principali parametri geometrici, ma non ricostruisce interamente la forma della vela. Essa dovrà poi essere ricostruita a posteriori partendo dai punti acquisiti. Altro aspetto negativo dell'approccio qui descritto è il fatto che esso è stato pensato e studiato per funzionare all'interno della Galleria del Vento, dove un filtro infrarossi permette di ridurre al minimo i disturbi causati dalla luce ambiente. Questo non vale per ambienti esterni, dove la scena è fortemente illuminata dal sole. Esso presenta infatti forti contributi di luce a tutte le frequenze d'onda, infrarossi compresi, rendendo inutile il filtraggio sulle ottiche e non permettendo ai marker catarifrangenti di essere chiaramente distinti nella scena.

Per questi motivi in parallelo al sistema di acquisizione stereoscopico è stato sviluppato, qualificato metrologicamente e testato in questa campagna sperimentale anche un approccio basato su tecnologia a tempo di volo [25], [27], [28]. In particolare tale approccio consiste nell'utilizzo di un laser scanner 2D movimentato

mediante un motore (figura 6.8). Tale sistema acquisisce l'intera nuvola di punti della vela (e non solo) e quindi necessita una più complessa elaborazione a posteriori dei dati. La nuvola di punti acquisita va infatti pulita, filtrata e vanno trovati i bordi della vela e le linee di forma sulla quale si calcolano i parametri geometrici di interesse. Questo sistema è stato utilizzato per effettuare la ricostruzione tridimensionale delle flying shape di gennaker e spinnaker nelle prove in Galleria del Vento. Da alcuni esperimenti esso è risultato molto robusto anche per quanto riguarda le acquisizioni di forme di vela su scala reale in ambiente esterno. Infatti esso è stato utilizzato nell'ambito del Lecco Innovation Hub [26], [27] per la misura della flying shape di vele reali di una barca a vela interamente sensorizzata costruita all'interno del Politecnico di Milano. Sul sistema di scansione 3D delle vele in scala reale è stata anche depositata una richiesta di brevetto [29].



Figura 6-8: Laser scanner (patent application n° PD2014A000249) utilizzato per la scansione delle vele in Galleria del Vento e sulla barca laboratorio LIH (Lecco Innovation Hub).

Per quanto riguarda la randa, acquisita col sistema stereoscopico, sono stati posizionati una serie di blob catarifrangenti sulle linee di forma della vela (figura 6.9). In questo modo è possibile ricostruire la forma tridimensionale di alcune sezioni della vela di particolare interesse. Per ognuna di queste sezioni vengono calcolati una serie di parametri geometrici che vengono comunemente utilizzati dai velai per definire la forma geometrica della vela (6.10).

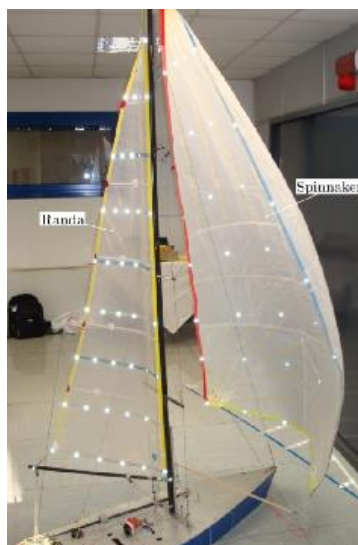


Figura 6-9: Modello di imbarcazione markerizzata per le acquisizioni della forma della vela con il sistema stereoscopico.

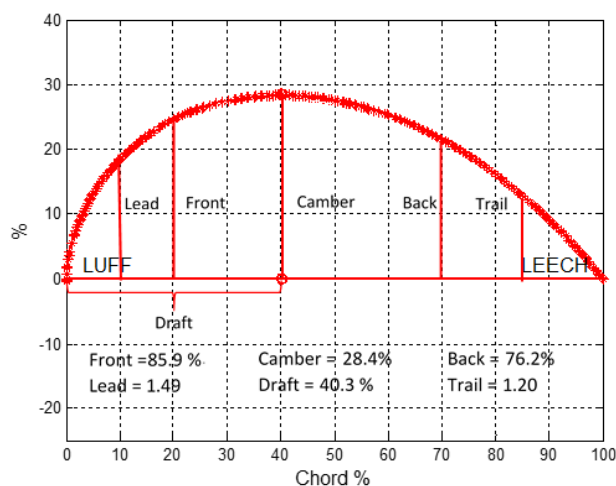


Figura 6-10: Esempio di parametri geometrici calcolati per ogni linea di forma della vela.

La design shape (forma teorica della vela) è stata realizzata e studiata da North Sails, usando il software North Design Suite (NDS), per estrarre i parametri geometrici di essa. Questi parametri sono stati confrontati con quelli della flying shape, acquisiti dal sistema stereoscopico. Solamente a titolo di esempio vengono riportati i risultati di alcune acquisizioni con vento a 95° e settaggi delle vele tali da avere la massima spinta tangenziale (driving force) della barca. Nella figura 6.11 e 12 sono riportate la design shape (griglia in blu) e la flying shape (in rosso) di randa

(chiamata sinteticamente nei grafici seguenti “main”) e gennaker (chiamato invece “A2max”), ricostruite con entrambi i sistemi di acquisizione tridimensionale.

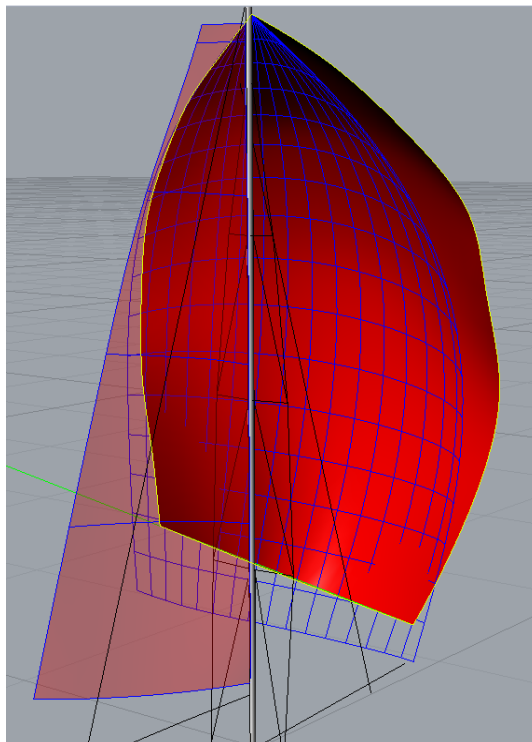


Figura 6-11: Confronto tra ricostruzione tridimensionale della design shape (griglia blu) e della flying shape (superficie rossa) – vista laterale -

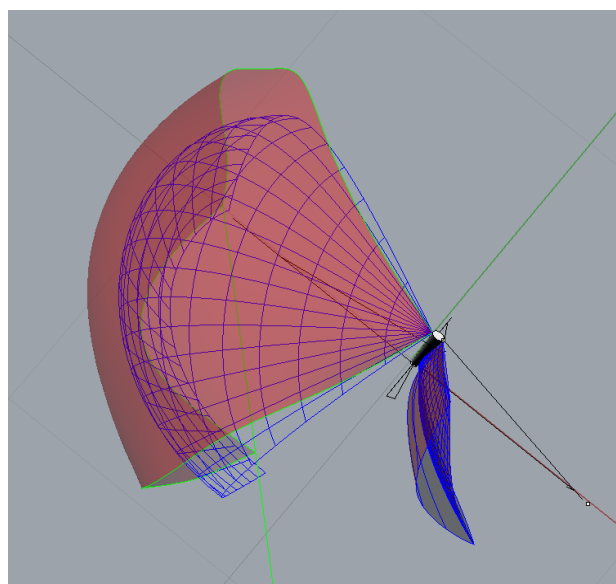


Figura 6-12: Confronto tra ricostruzione tridimensionale della design shape (griglia blu) e della flying shape (superficie rossa) – vista dall’alto -

I parametri che derivano dallo studio della forma delle vele sono diversi; nel seguito vengono riportati alcuni risultati su camber (distanza massima tra la linea di forma e la sua corda in percentuale sulla corda stessa) e twist (angolo tra la corda della linea di forma e la linea di mezzeria dell'imbarcazione). Questi due parametri sono i più importanti per quanto riguarda lo studio delle forze aerodinamiche sulle vele. Inoltre vengono analizzate anche due curve verticali chiamate LUFA e LUSW, che sono rispettivamente la proiezione della curva dell'inferitura (spigolo di ingresso del vento) su un piano definito dagli angoli della vela e la proiezione dell'inferitura su un piano perpendicolare al precedente e definito dai due angoli di mura e testa.

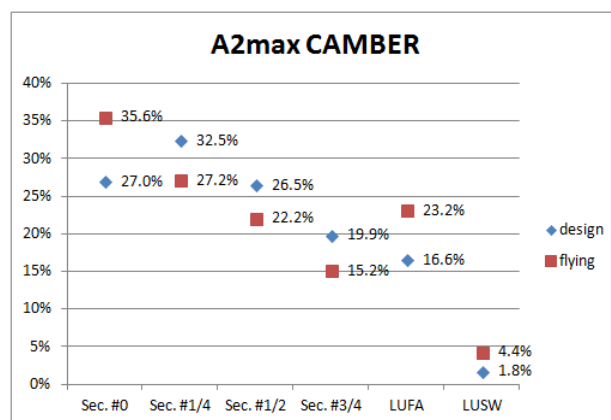


Figura 6-13: Gennaker, parametro Camber (design vs flying shape)

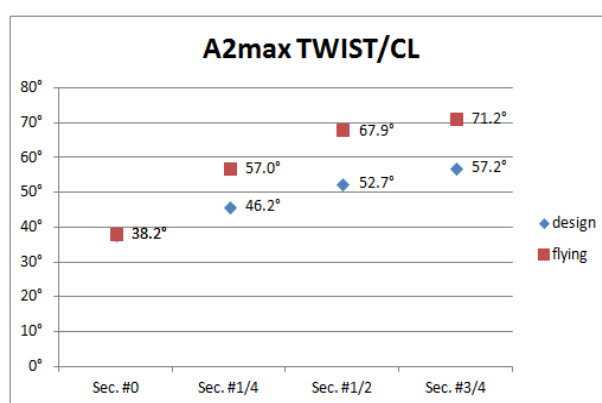


Figura 6-14: Gennaker, parametro Twist (design vs flying shape)

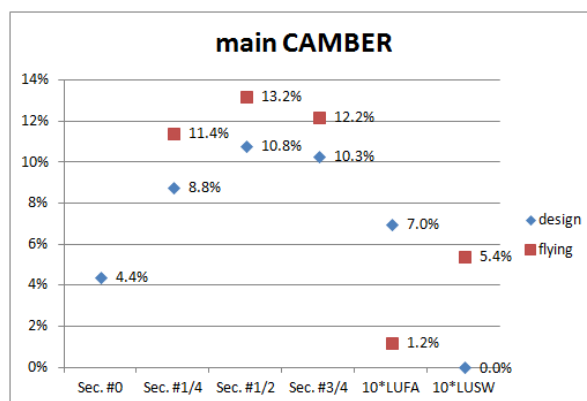


Figura 6-15: Randa, parametro Camber (design vs flying shape)

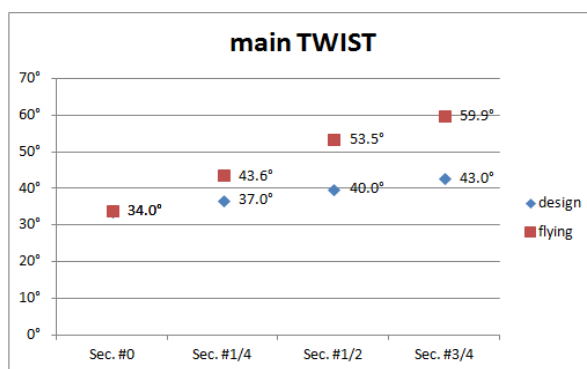


Figura 6-16: Randa, parametro Twist (design vs flying shape)

Entrambe le vele mostrano delle differenze sostanziali tra la design shape e la flying shape. Questo è intuibile analizzando i grafici di camber orizzontale e verticale e di twist. Per quanto riguarda il gennaker A2max esso presenta una forma diversa dalla design shape in quanto le mura sono più alte e il twist è maggiore. Questo è dovuto a una regolazione tale da ricercare la spinta tangenziale massima con le condizioni di vento attuali. Anche la randa presenta un comportamento diverso tra design e flying shape. L'albero risulta infatti più dritto di quanto ipotizzato nella design shape e di conseguenza la curva dell'inferitura reale risulta meno curva che nella design. Inoltre è interessante il confronto tra la forma delle vele ricostruita di figura 6.11 e 6.12 in 3d e un'immagine scattata dalle telecamere di Galleria del Vento (figura 6.17); è facile intuire come le irregolarità delle forme delle vele, specialmente delle inferiture in questo caso, sono state acquisite correttamente dai sistemi di misura.

Il confronto tra design e flying shape risulta essere molto interessante poiché offre al velaio la possibilità di verificare quale sia la forma reale della vela sottoposta a una certa azione del vento e con definite regolazioni. Questo è molto difficile da valutare solamente con la design shape poiché essa è un'approssimazione di come potrebbe essere la forma geometrica della vela, ma non rispecchia a pieno la realtà, influenzata da numerose cause concomitanti.



Figura 6-17: Immagine dall'alto dell'imbarcazione durante le prove in galleria del vento.

Oltre al confronto tra design e flying shape è molto interessante studiare come si modifica la flying shape a seconda delle regolazioni effettuate e confrontarla con le misure di forza acquisite dalla bilancia. Nella figura 6.18 sono riportate le immagini raffiguranti 4 diversi set di regolazioni utilizzati. Partendo con la vela in posizione “step 0”, essa viene poi regolata fino ad arrivare allo “step 1” e così via fino allo “step 3”. Le forme della vela ricostruite dalle acquisizioni effettuate col sistema stereoscopico sono invece riportate nella figura 6.19, dove sono evidenti le 4 diverse regolazioni della randa.

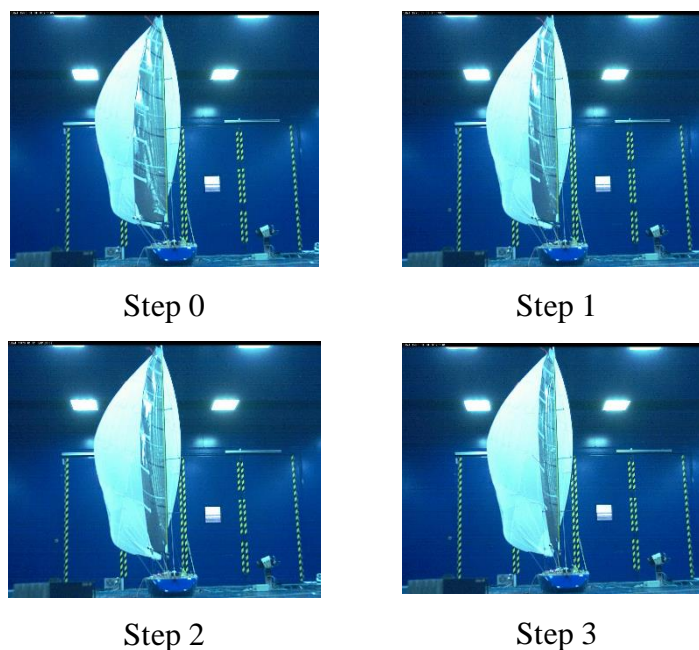


Figura 6-18: Immagini dell'imbarcazione con i diversi settaggi.

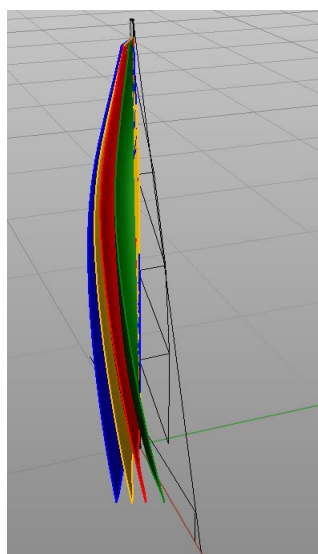


Figura 6-19: Ricostruzione tridimensionale della randa con i diversi settaggi.

L'effetto del parametro chiamato twist, che definisce la torsione della vela, è ben evidente nella figura 6.20. Esso è strettamente correlato alle regolazioni effettuate: infatti tra lo step 0 e 1 viene cazzata la scotta della randa ma essendo l'angolo del boma alto, questa regolazione riduce il twist nella parte bassa della vela più che nella parte alta. Dallo step 1 a 2 il carrello viene spostato più verso il vento (verso il lato destro della barca); questa regolazione riduce in maniera uguale il

twist su tutta la randa. Cazzando la scotta della randa nuovamente viene tensionata la balumina (spigolo di uscita del vento) riducendo il twist della parte alta più che della parte bassa della randa.

Il diagramma delle forze (figura 6.21) invece mostra una leggera riduzione della spinta motrice (tangenziale) dell'imbarcazione dovuta a un aumento della forza laterale. Probabilmente questo effetto è legato al fatto che nello step 4 una parte maggiore della vela risulta essere in stallo, vale a dire che presenta angoli di incidenza al vento troppo alti con una conseguente diminuzione della portanza.

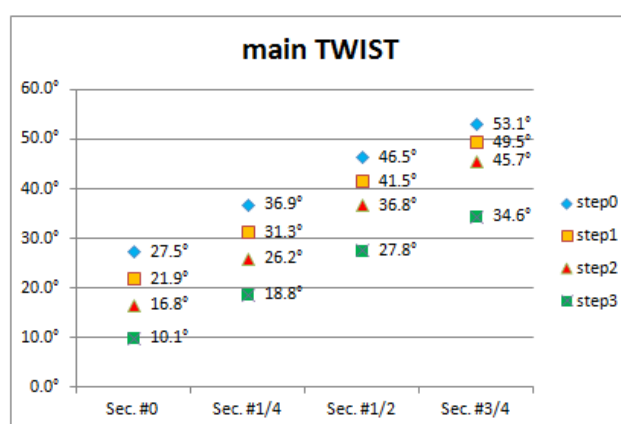


Figura 6-20: Valori di twist con diversi settaggi della randa

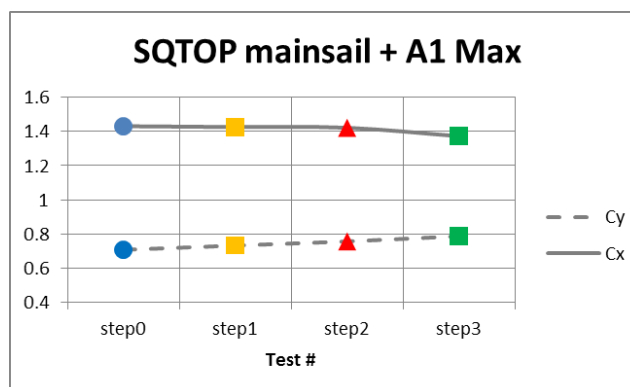


Figura 6-21: Forza di tangenziale e laterale per i diversi settaggi

Le stesse analisi sono state fatte su diversi tipi di randa e in parallelo su diversi gennaker e spinnaker utilizzando il laser scanner. Per ogni coppia di vele le prove sono state ripetute per diversi angoli di incidenza del vento. Dai dati si è potuto constatare che entrambi i sistemi di misura tridimensionale acquisiscono con buona

accuratezza le forme delle vele. Il sistema stereoscopico presenta accuratze e velocità di rielaborazione dei dati migliori ma il setup risulta più complicato e non può funzionare in ambienti esterni. Esso inoltre non consente di acquisire un numero altissimo di punti sulla vela (definiti dai marker catarifrangenti) e risulta poco robusto con vele per andature di portanti a causa di possibili occlusioni.

6.1.4 Misura di strutture civili

Il sistema di misura tridimensionale stereoscopico realizzato per la Galleria del Vento del Politecnico di Milano è stato inoltre utilizzato per diverse campagne di prove in cui viene studiato il comportamento aereodinamico di strutture civili sottoposte ad azione fluida del vento. In questo tipo di campagne sperimentali sono spesso utilizzati degli accelerometri posti sulla struttura in esame su particolari punti che si vogliono analizzare. Questi accelerometri spesso modificano le caratteristiche fisiche e aeroelastiche dei modelli misurati poiché introducono componenti di masse spesso non trascurabili. Con il sistema stereoscopico descritto nel precedente paragrafo è possibile effettuare le stesse misure senza introdurre forti effetti di carico sul misurando. Andranno solamente posizionati dei blob catarifrangenti (con massa e forma trascurabili per questo tipo di applicazioni) nei punti della struttura che si intendono misurare.

Diverse prove sono state fatte per qualificare lo strumento di misura, utilizzandolo in parallelo a una serie di accelerometri, in modo da poterne verificare l'affidabilità. I dati restituiti dagli accelerometri sono stati integrati due volte nel dominio delle frequenze (così da poter eliminare le componenti in bassa frequenza dovute all'accelerazione di gravità), trovando l'andamento temporale dello spostamento. Questo è stato confrontato con gli andamenti temporali misurati dal sistema stereoscopico. Ovviamente per effettuare queste verifiche è necessario che il sistema di riferimento in cui sono espresse le misure degli accelerometri e quello del sistema stereoscopico siano il più possibile allineati. Inoltre i due diversi sistemi di acquisizione sono indipendenti uno dall'altro e dunque i segnali trovati con l'accelerometro e con il sistema stereoscopico non sono sincronizzati e vanno

dunque correlati attraverso un algoritmo di crosscorrelazione, in modo da essere paragonabili.

A titolo di esempio vengono qui mostrati due segnali messi a confronto acquisiti rispettivamente da un accelerometro e dal sistema di misura. Questa prova è stata effettuata per studiare il movimento verticale di una struttura civile posta sotto azione del vento. Viene riportato in figura 6.22 l'andamento verticale di uno dei marker catarifrangenti e il segnale ricavato dalle misure di un accelerometro posto in corrispondenza di esso. Già da una prima analisi dei due segnali si nota che essi sono piuttosto simili, con oscillazioni che differiscono in ampiezza di circa 1 mm. Per un'analisi più dettagliata si è ricavato lo spettro in frequenza dei due segnali, mostrato in figura 6.23. Da questo è facile verificare che i due segnali hanno entrambi frequenza propria pari a 1.6 Hz. A tale frequenza corrisponde nel segnale misurato dal sistema stereoscopico un modulo di 4.7 mm, mentre nel segnale ricavato dall'accelerometro un modulo di 4.37 mm. Lo scostamento della misura in questo caso è dunque circa mezzo millimetro.

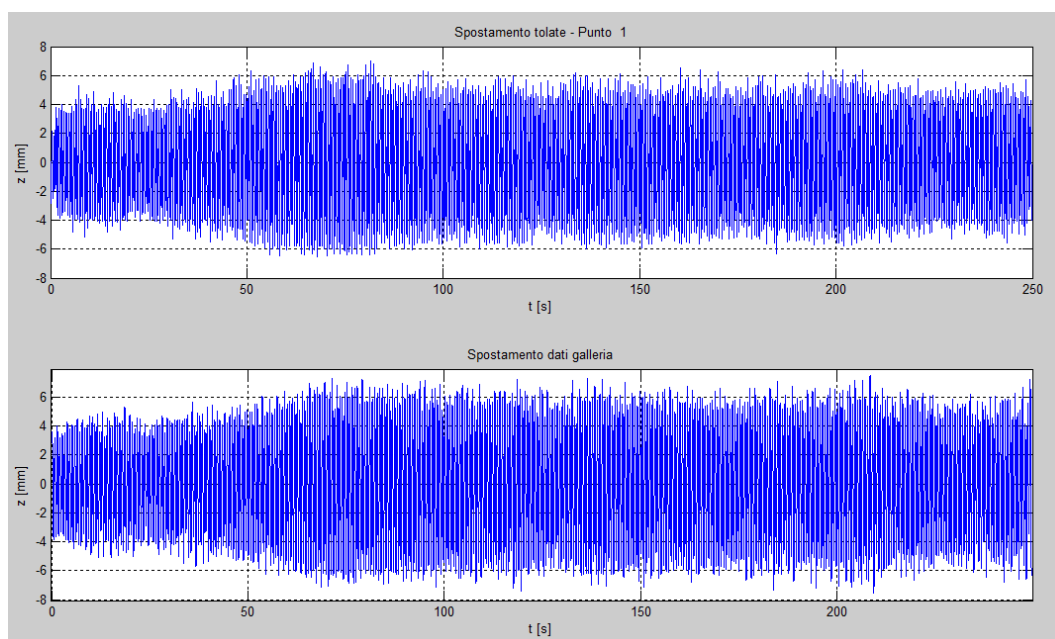


Figura 6-22: Confronto tra lo spostamento del modello misurato dal sistema stereoscopico e dall'accelerometro.

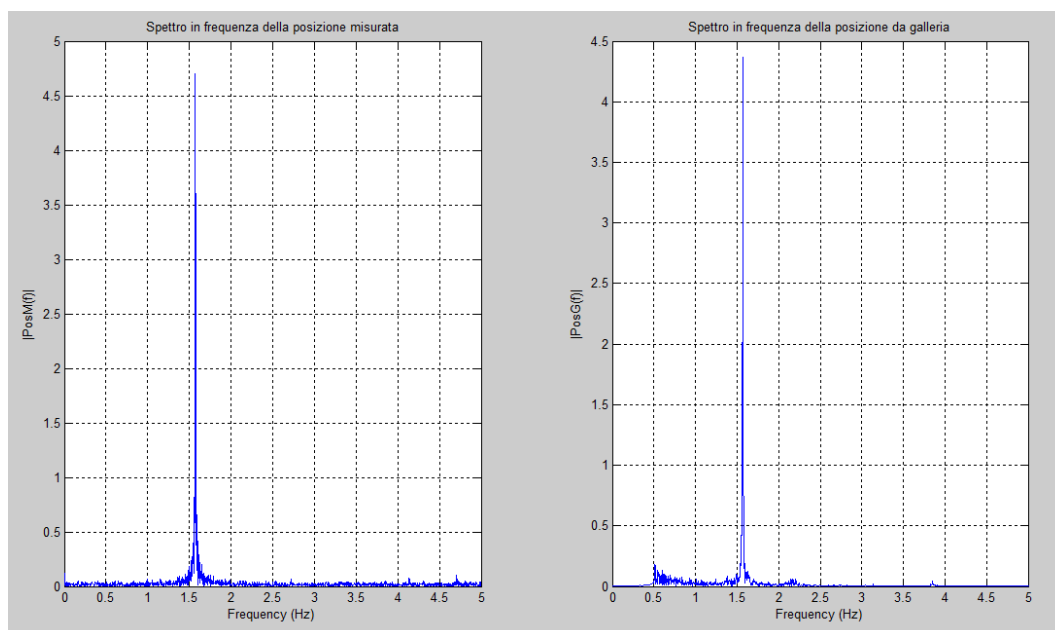


Figura 6-23: Spettro in frequenza dello spostamento del modello misurato dal sistema stereoscopico e dall'accelerometro.

Diverse campagne sperimentali sono state fatte utilizzando il sistema di misura stereoscopico per lo studio aeroelastico di ponti o sezioni di ponti. Ad esempio la figura 6.24 si riferisce alle prove eseguite sul modello di un famoso ponte. Il sistema è stato utilizzato anche per lo studio di una struttura che serve per aumentare la stabilità di uno dei palazzi del nuovo complesso CityLife di Milano (figura 6.25).

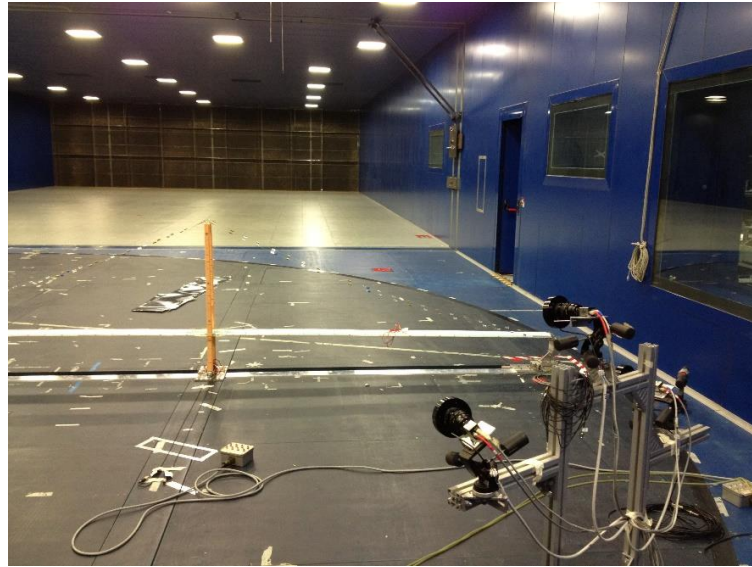


Figura 6-24: *esempio di applicazione del sistema stereoscopico alle misure di spostamento del modello di un modello di ponte in Galleria del Vento.*



Figura 6-25: *esempio di applicazione del sistema stereoscopico alle misure di spostamento del modello di una grossa struttura civile in Galleria del Vento.*

6.2 Robot pittore

Ricorrere all'utilizzo di robot è sempre più comune nella gran parte degli ambienti produttivi. Determinate lavorazioni possono essere pericolose per l'operatore a causa di condizioni ambientali sfavorevoli oppure possono presentare specifiche di processo troppo stringenti perché se ne possa occupare un operatore umano: si richiedono gradi di ripetibilità e accuratezza elevati oppure un'alta velocità di esecuzione. Per questo tipo di lavorazioni è indicato l'utilizzo di un manipolatore robotico in modo da migliorare prestazioni statiche e dinamiche. Nonostante la robotica sia una disciplina ormai matura, la ricerca è ancora fortemente impegnata nel trovare strategie innovative per permettere al manipolatore di interfacciarsi direttamente al mondo circostante. È possibile affiancare al robot una serie di sensori in grado di fornire segnali utili per una sua programmazione automatica.

Nell'ottica di incrementare i volumi produttivi e le prestazioni si vuole valutare la possibilità di affiancare un manipolatore robot ad un processo di decorazione a mano di ceramiche. Sarà necessario eseguire la classica decorazione manuale una sola volta per ogni tipologia di decoro affinché il sistema di telecamere ne possa acquisire le informazioni necessarie a permetterne la riproduzione "robotizzata". In questo caso i vantaggi sarebbero considerevoli: linee in grado di funzionare otto ore/turno, tre turni/giorno, ogni giorno. Lo scopo è quello di ottenere un prodotto finito molto simile al "lavorato a mano", come qualità, ma che presenti costi di produzione paragonabili al serigrafato, quindi alla piastrella stampata. Lo studio di fattibilità, per questo tipo di applicazione robot, viene ad essere necessario, dato il suo aspetto innovativo ed allo stesso tempo inusuale. È infatti fortemente sconsigliato ricorrere ad un metodo di programmazione convenzionale. La necessità di riprodurre fedelmente i movimenti del braccio umano, con la loro ricchezza e fluidità, impedisce l'utilizzo di una tecnica di programmazione punto a punto. Da qui l'idea di sviluppare una nuova tecnica di generazione delle traiettorie robot al fine di risultare competitivi sul mercato per qualità e velocità di esecuzione, a costi ridotti. Si vuole quindi studiare la possibilità di applicare un'innovativa

tecnica di programmazione dinamica e automatica basata sulla visione artificiale ed in particolare sulla stereoscopia. Il progetto è stato svolto in collaborazione con ABB Robotica Italia [32].

Nonostante le specifiche siano dettate da questa particolare applicazione e per un determinato setup della cella produttiva, si è scelto di sfruttare un sistema di misura che fosse in grado di adattarsi a tutte le diverse possibili esigenze in termini di area di lavoro ed ingombri. Lo scopo è quindi quello di creare un prodotto che sia personalizzabile e in grado di soddisfare diverse richieste; esso non si deve limitare a risolvere la specifica problematica della riproduzione del movimento di un decoratore, ma più in generale deve occuparsi di acquisire, trattare e riprodurre un qualsiasi movimento di un utensile generato da un braccio umano. Sono quindi richieste caratteristiche di flessibilità e scalabilità del sistema di misura ed esso deve ovviamente essere in grado di effettuare misurazioni dinamiche. Queste considerazioni hanno portato alla scelta della stereoscopia come metodologia di misura.

6.2.1 Setup e taratura del sistema

Lo strumento di misura scelto per questa applicazione è il medesimo a tre telecamere descritto nel paragrafo 5.1 e utilizzato anche per le misure in Galleria del Vento (paragrafo 6.1).

La posizione relativa delle telecamere è stata dettata da alcuni vincoli dovuti alla particolare applicazione. L'artista per poter effettuare la decorazione del pezzo necessita di potersi muovere su almeno tre lati della piastrella. Sul quarto lato vengono posizionate le telecamere. Se fossero posizionate in modo tale che gli assi ottici fossero allineati con gli assi del sistema di riferimento, il decoratore rischierebbe di oscurarle durante l'acquisizione rendendo i risultati inservibili. È perciò necessario mantenere tutte le telecamere dallo stesso lato dell'area di lavoro, facendo in modo che gli assi ottici siano il meno allineati possibile. Il supporto per le telecamere viene mantenuto ad una distanza di 110 cm dal piano di lavoro e la barra orizzontale, alla quale vengono fissate le telecamere, è situata ad un'altezza

di 80 cm dal piano. La barra sulla quale sono montate le telecamere misura anch'essa 80 cm. È stato deciso di montare le due telecamere esterne agli estremi della barra, in maniera tale da rendere più ampia possibile la baseline del sistema stereoscopico che, in questo modo, viene a coincidere con la lunghezza del supporto. La telecamera centrale è montata a 40 cm da ognuna delle altre due telecamere e ad un'altezza, rispetto alle altre due, 25 cm maggiore. È necessario inoltre far in modo che gli assi ottici delle tre telecamere convergano sull'area di lavoro. Per fare ciò, le telecamere laterali vengono ruotate attorno alla verticale di circa 25 gradi. Inoltre vengono ruotate di circa 35 gradi attorno al supporto orizzontale. La telecamera centrale, invece, non viene ruotata attorno alla verticale essendo posizionata frontalmente al piano di lavoro. Viene però ruotata di circa 45 gradi attorno all'orizzontale.

Sul pennello utilizzato dal decoratore vengono posizionati alcuni marker catarifrangenti (figura 6.27), le quali posizioni nel tempo sono misurabili col sistema stereoscopico. Conoscendo le coordinate 3D dei marker, mediante un algoritmo di registrazione rigida basato su SVD (*Singular Value Decomposition*) [31], sarà possibile risalire a posizione ed orientamento della punta del pennello.



Figura 6-26: *Pennello utilizzato per la decorazione manuale a cui sono stati aggiunti quattro marker per permettere le acquisizioni.*

Per quanto riguarda la procedura di calibrazione, si è deciso di utilizzare ancora una volta la tecnica implementata nelle librerie HALCON.

Il robot a nostra disposizione è stato fornito da ABB e in particolare il modello è IRB-120, un antropomorfo piccolo e leggero, il più piccolo prodotto da ABB. È particolarmente indicato per applicazioni dove i carichi in gioco non sono eccessivi e le aree di lavoro ridotte. Può manipolare un carico massimo di 3 kg (4 kg nella configurazione con polso verso il basso). Grazie agli ingombri minimi e all'elevata flessibilità è capace di sostituirsi ai vecchi metodi di automazione standard rigidi e poco adattabili. La sua struttura, le rotazioni massime associate a ciascun asse e gli ingombri vengono riportati in figura 6.28.

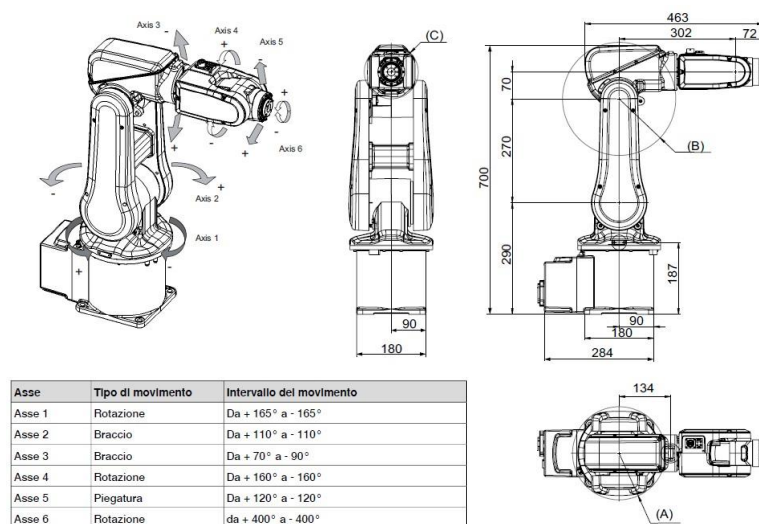


Figura 6-27 Struttura del robot IRB120, ampiezza di movimento dei singoli assi e ingombri.

I robot ABB possono essere controllati manualmente attraverso il flexpendant. Esso è sostanzialmente un'estensione del controllore. Il linguaggio ABB per scrivere programmi interpretabili dal controllore è RAPID. I programmi RAPID possono essere scritti direttamente da flexpendant oppure, alternativamente, possono essere sviluppati su un computer e successivamente trasferiti (attraverso una connessione TCP-IP o USB) sul controllore robot.

Il robot IRB-120 è stato installato in una cella di dimensioni 1200x1400 (figura 6.29). Sul piano di lavoro, per permettere al robot di eseguire i movimenti di

decorazione il più agilmente possibile, è stato fissato un supporto (figura 6.29) dove sono bloccate le piastrelle ed alloggiate le vasche del colore (figura 6.30).

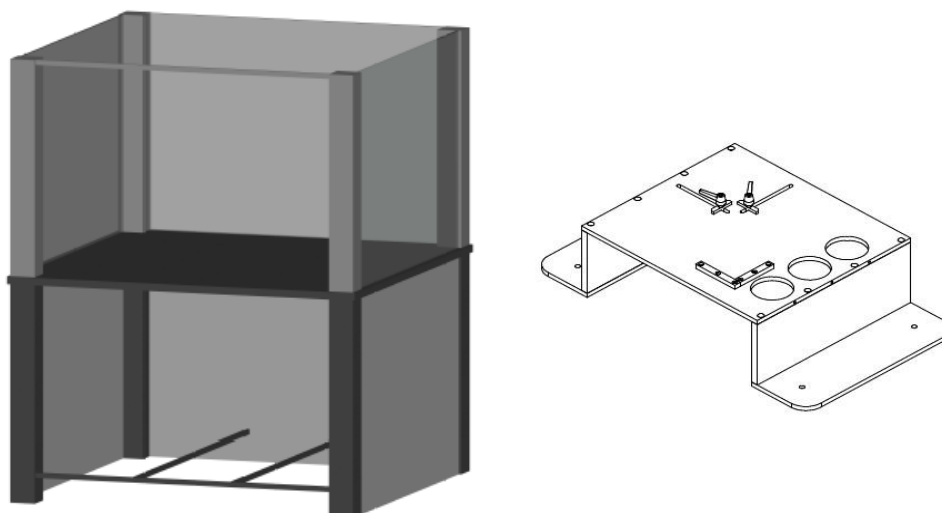


Figura 6-28: *Disegno CAD della cella e del supporto per le piastrelle.*

Come si può notare il supporto per piastrelle, disegnato apposta per questa applicazione, presenta tre fori di diametro 82 mm nel quale è possibile alloggiare le vasche per i colori già accennate. L'angolare permette di inserire una piastrella in maniera da poter eseguire l'operazione di decorazione in maniera riferita e ripetibile. Nelle due asole scorrono delle squadrette necessarie per bloccare il pezzo da decorare. Sono prese in considerazione piastrelle di forma sia quadrata che rettangolare le cui misure siano comprese tra 200x200 e 300x300. Le vasche per il colore attualmente utilizzate dai decoratori professionisti sono vasche in ceramica, come quella vista in figura 6.30:



Figura 6-29: *Struttura della vasca per il colore attualmente utilizzata nel processo di decorazione a mano.*

Esse presentano un canale attorno al corpo principale per poter raccogliere il colore rilasciato dal pennello. Una lieve pendenza del canale di raccolta del colore e due fori tra permettono al colore di defluire nuovamente all'interno. Tuttavia, questi contenitori devono essere modificati ed adattati all'ambiente industriale in cui devono essere utilizzati. Il problema principale delle vasche utilizzate per la decorazione a mano è quello del controllo di livello: hanno bisogno di essere periodicamente rabboccate. Questa è una condizione inaccettabile nell'ottica di una produzione automatizzata. Per questo motivo è stato progettato un diverso contenitore per il colore, in figura 6.31. Le vasche del colore possiedono una forma particolare: sono dotate di una canula di scarico e di una di aspirazione. Per la loro particolare costituzione i colori utilizzati per la decorazione di ceramiche necessitano di essere mantenuti costantemente in circolo per evitare il deposito di sedimenti. La canula di scarico rilascia il colore in un serbatoio di dimensioni superiori al contenitore per il colore utilizzato dal decoratore. Una pompa, collegata alla canula di aspirazione, ha il compito di prelevare colore dal serbatoio ed iniettarlo all'interno del contenitore. La canula di scarico, oltre ad avere il compito di far scaricare il liquido all'interno del serbatoio, possiede anche la funzione di controllo di livello. Se il colore all'interno del contenitore è troppo e supera il livello desiderato, defluirà dalla canula di scarico, se il colore è troppo poco, non potrà essere scaricato fino a quando il contenitore non si sarà riempito fino al livello voluto.



Figura 6-30: *Disegno CAD di una vasca per il colore.*

Alla flangia del robot è possibile fissare diversi pennelli a secondo del tipo di tratto desiderato: è possibile utilizzare pennelli a punta piatta e tonda ed ognuna delle due varianti è presente in diverse dimensioni. Attualmente il cambio pennelli è un'operazione che va effettuata a mano. Con lo scopo di industrializzare il processo si sta già studiando un setup che consenta di cambiare automaticamente pennello con una pinza e di utilizzare un maggior numero di possibili colori, attualmente limitato a tre. La figura 6.33 mostra l'isola di acquisizione e l'isola robotizzata per la decorazione.

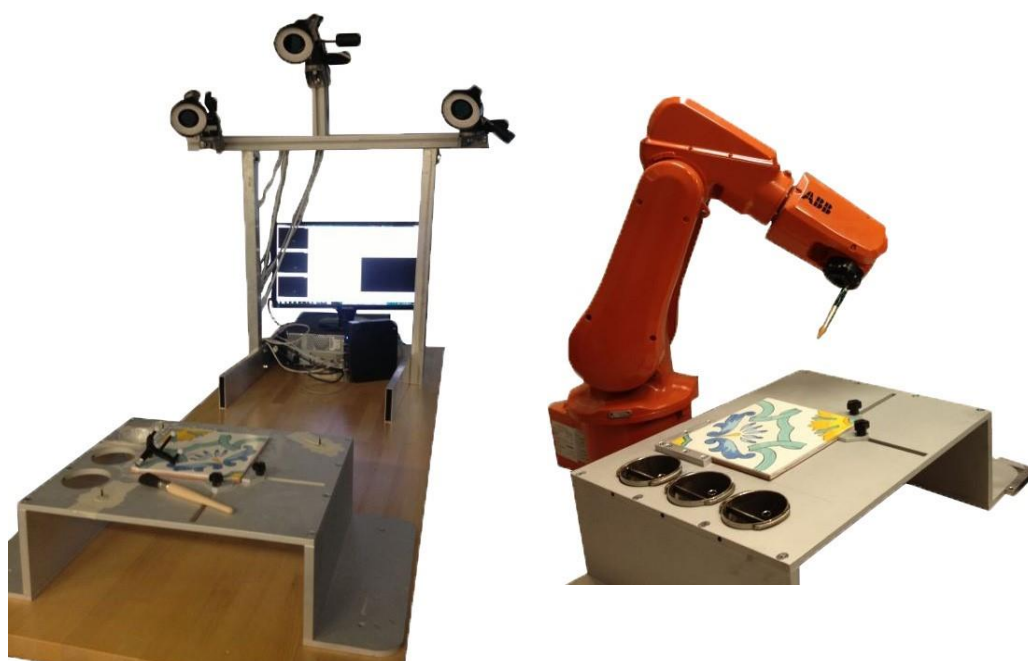


Figura 6-31: A sinistra la cella di decorazione manuale, sulla destra quella robotizzata.

I dati acquisiti dal sistema stereoscopico risultano affetti da un errore di misura che si ripercuote poi sulla fluidità dei movimenti del manipolatore. Infatti dalle prime prove si è notato che il robot vibrava in maniera evidente durante la riproduzione del movimento acquisito. Per risolvere questa problematica i dati acquisiti dal sistema stereoscopico vengono filtrati mediante un filtro passa basso.

6.2.2 Analisi dell'incertezza di misura

La verifica sperimentale dell'incertezza di misura avviene secondo i seguenti passi:

- si effettua un'analisi dell'incertezza generata dal processo di ricostruzione tridimensionale. Una prova di ricostruzione di un pattern di punti di cui sono note con grande accuratezza le distanze relative permette di verificare la bontà del processo di ricostruzione tridimensionale. È effettuata una seconda prova in cui viene ricostruito il singolo marker permettendo di stimare l'incertezza di misura della posizione di un marker sferico di 10 mm di diametro;
- si effettua un'analisi dell'incertezza generata dalle calibrazioni del TCP e del Work Object lato robot e dalla calibrazione del sistema di riferimento lato sistema di visione. Gli errori introdotti dalle diverse calibrazioni generano uno scostamento in traslazione e rotazione tra i sistemi di riferimento lato telecamera e lato robot;
- si conclude la qualificazione metrologica fornendo un'analisi della stima dell'incertezza della posizione della punta del pennello. Viene quindi verificato non solo il processo di ricostruzione tridimensionale, ma anche quello di registrazione rigida. Non è più utilizzato un singolo marker sferico, ma l'intera disposizione di marker sul pennello, in modo da essere in grado di ricostruire la posizione della punta. Inoltre vengono effettuate alcune prove per testare la robustezza del sistema di ricostruzione 3D e registrazione rigida a fronte di perdita di visibilità su un marker e di rotazione del pennello attorno al proprio asse.

6.2.3 Incertezza della ricostruzione 3D

Prima di potersi interfacciare con il sistema robot è utile verificare la bontà degli algoritmi di calibrazione e ricostruzione tridimensionale. Si sfrutta per questo tipo di verifica un pattern di punti di cui sono note con grande accuratezza le distanze relative. Una volta ricavate le coordinate immagine del centro dei 49 blob

è possibile verificare la bontà della procedura di ricostruzione tridimensionale. Sono state acquisite $k=30$ pose del calibratore e, per ciascuna di esse, è stata misurata col sistema stereoscopico la posizione 3D. Ancora una volta viene utilizzato il vettore delle distanze tra i blob per verificare l'accuratezza di misura. Conoscendo il vettore delle distanze reali, note con buona accuratezza, e misurando le distanze tra i blob col sistema di misura stereoscopico, è possibile ricavare un valore di scostamento medio delle distanze per ogni posa acquisita. Il calibratore è stato posto a distanze comprese tra 150 cm e 200 cm dalle telecamere. Si può notare dal grafico 6.34 come lo scostamento non oltrepassi mai il valore di 0.230 mm e sia in media pari a 0.124 mm.

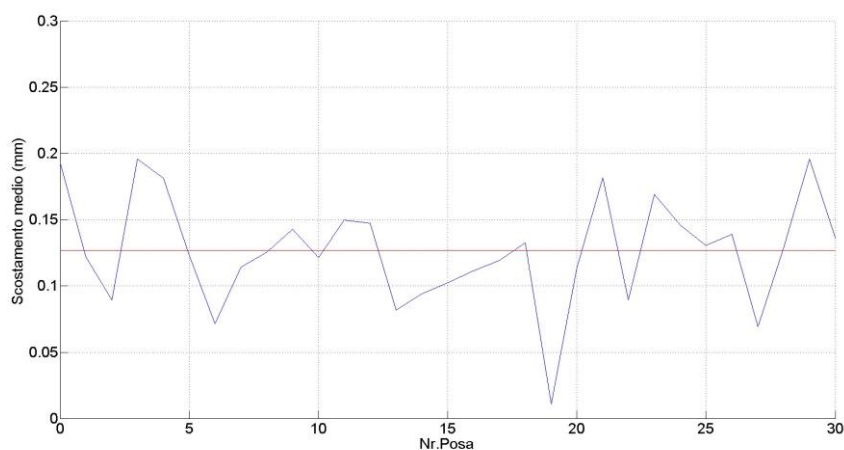


Figura 6-32: *Scostamento medio dalle distanze reali per ogni posa acquisita del calibratore.*

Per valutare l'incertezza del sistema di visione è stata inoltre ricostruita la posizione tridimensionale di un marker di 10 mm di diametro fissato alla flangia del robot. Per questa verifica non è necessario che manipolatore e sistema di visione siano riferiti al medesimo SdR, essendo interessati alla variabilità delle misure e non al loro valore assoluto. Ciò che si vuole osservare attraverso questo esperimento è la variazione della posizione del centro del marker lungo le tre direzioni, mantenendolo fisso nella scena. Per ricavare la stima dell'incertezza di misura si utilizza un approccio statistico. Vengono acquisite 200 immagini della scena e successivamente elaborate al fine di ricostruire la posizione del marker nello spazio. Una volta ottenute le posizioni 3D, da ogni serie verrà estratto un indicatore di

variabilità. Le misure sono distribuite come le gaussiane mostrate in figura 6.35. Le linee rosse (di riferimento) sono poste in corrispondenza dei valori $+3\sigma$ e -3σ , valori all'interno dei quali cade il 99.7% delle misure, ipotizzando una distribuzione gaussiana.

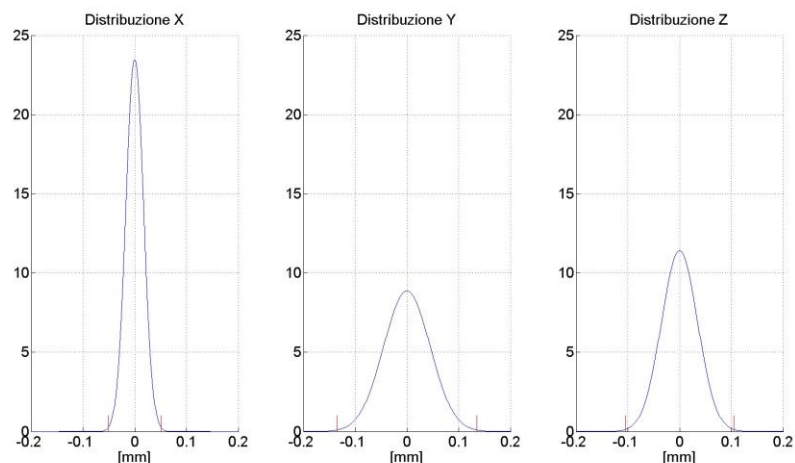


Figura 6-33: Distribuzione normalizzata delle misure di X, Y e Z del centro di un marker sferico di 10 mm di diametro sottoposto ad acquisizione statica.

Per ciascuna coordinata X, Y e Z viene calcolata la deviazione standard della serie di misure, definita anche *scarto quadratico medio* o *scarto tipo*, σ . Attraverso il valore di σ è possibile avere un'informazione di dispersione delle misure sperimentali: esso fornisce una stima della variabilità delle misure acquisite attorno al valore atteso. Il vettore degli scarti tipo ottenuto è (in millimetri) il seguente

$$\bar{\sigma} = \begin{bmatrix} \sigma_x & \sigma_y & \sigma_z \end{bmatrix} = \begin{bmatrix} 0.017 & 0.045 & 0.035 \end{bmatrix}$$

da cui, attraverso la formula

$$\|\bar{\sigma}\| = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$$

otteniamo uno scarto tipo complessivo pari a 0.059 mm, valore accettabile per l'applicazione considerata. È possibile giustificare i risultati ottenuti considerando la disposizione delle tre telecamere nello spazio.

6.2.4 Incertezza del sistema lato Robot

Per poter operare correttamente, il controllore robot necessita della definizione di un TCP e di un Work Object. Come abbiamo già detto precedentemente, la definizione del TCP è necessaria per poter identificare la terna a cui vogliamo che i movimenti richiesti al controllore siano riferiti. La definizione del Work Object, invece, è necessaria per poter riferire tutti i punti che verranno identificati durante la programmazione ad una terna di riferimento fissa nell'ambiente di lavoro. Entrambe le definizioni vengono effettuate tipicamente da un operatore umano. Di seguito vengono brevemente presentate le procedure necessarie per le rispettive calibrazioni e le incertezze da esse introdotte.

La procedura ABB di definizione del TCP sfrutta uno strumento di calibrazione, tipicamente una punta. La procedura richiede che vengano salvate almeno quattro configurazioni differenti del robot per cui il tool sia a contatto con la punta dell'attrezzo di calibrazione. Si veda figura 6.36.

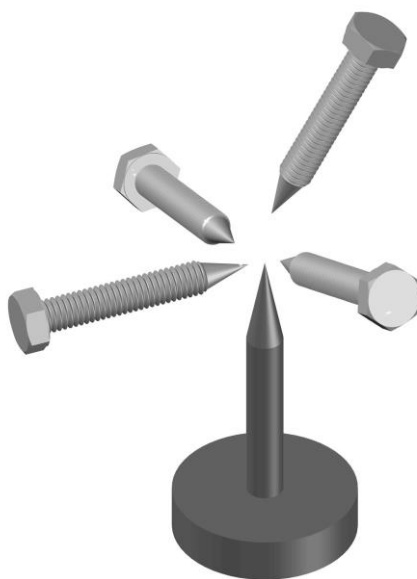


Figura 6-34: Procedura di calibrazione TCP sfruttando quattro pose del tool.

La calibrazione è tanto più affidabile quanto più le configurazioni salvate differiscono tra loro per coordinate di giunto. È importante, inoltre, che passando

da una configurazione all'altra, vengano movimentati tutti e sei gli assi del robot. Una volta fornite al controllore le configurazioni, vengono elaborate, definendo il TCP e restituendo tre valori di errore. Uno è l'errore massimo e rappresenta, in millimetri, la distanza massima tra il TCP appena definito e le posizioni della punta del tool durante l'acquisizione delle configurazioni. Il secondo è l'errore medio ovvero la distanza media tra il TCP appena definito e le posizioni della punta del tool durante l'acquisizione delle configurazioni. Il terzo, invece, riporta l'errore minimo di definizione ed è simmetrico all'errore massimo. Una buona calibrazione fornisce un errore medio inferiore ai due decimi di millimetro. Sono state effettuate 15 prove ripetute di calibrazione TCP i cui risultati sono riportati in figura 6.37. Si può notare come l'errore medio di calibrazione sia di poco superiore ai due decimi di millimetro e come gli errori si scostino dal valore medio al più di un decimo di millimetro.

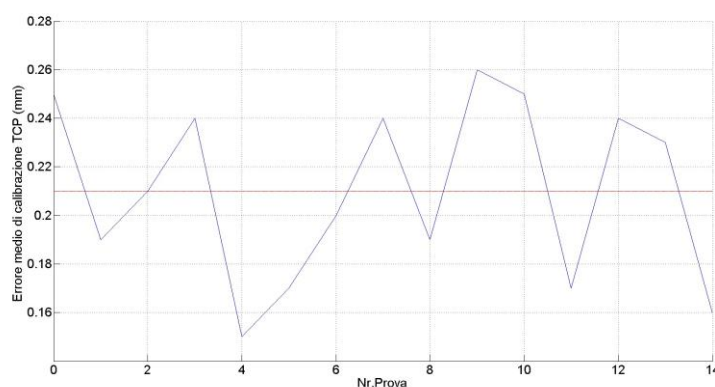


Figura 6-35: Scostamento medio dalle distanze reali per ogni posa acquisita del calibratore.

La procedura di definizione del Work Object, invece, richiede che venga fornita la posizione di due punti giacenti sull'asse x desiderato, in cui il secondo indica il verso dell'asse, e un punto sull'asse y . L'algoritmo di definizione opera in questo modo: attraverso i punti x_1 e x_2 vengono ricavati direzione e verso dell'asse x mentre attraverso il punto y_1 genera il piano ortogonale alla direzione x appena ricavata e passante per y_1 . Dall'intersezione tra piano e retta è possibile individuare la posizione dell'origine della terna Work Object. L'asse z viene identificato

completando la terna in maniera che risulti levogira. Per maggior chiarezza, si veda figura 6.38.

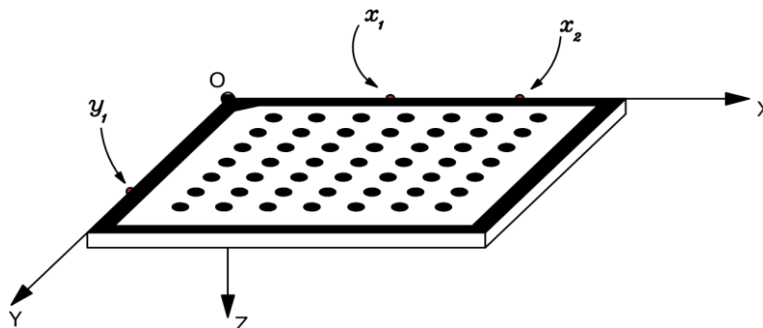


Figura 6-36: Procedura di definizione della terna work object.

La definizione del Work Object porterà con sé un errore di calibrazione che deve essere stimato. Ammettendo che l'operatore sia molto preciso nell'operazione di puntamento di x_1 , x_2 e y_1 , il tool è affetto da un errore di posizionamento medio di tre decimi di millimetro, al massimo. Se i punti vengono scelti in modo tale che l'influenza dell'incertezza del TCP sul Work Object venga minimizzata e, quindi, il punto x_1 non si discosti molto dall'origine della terna Work Object, la posizione di quest'ultima sarà affetta da un errore medio leggermente maggiore di quello ottenuto precedentemente durante la definizione del TCP. Stimare l'incertezza derivante dal processo di definizione del Work Object non è banale, verrà quindi considerato un cautelativo 0.5 millimetri come valore di incertezza della procedura di definizione del Work Object lato robot.

6.2.5 Misurazioni Camera Vs Robot

Sia la telecamera che il robot, come si è visto, sono affetti da errori generati durante la fase di calibrazione dell'una e dell'altro. Sono stati effettuati dei test per definire gli effetti di questi errori.

Prima di tutto viene misurato lo scostamento in traslazione e rotazione tra i sistemi di riferimento lato telecamera e lato robot. Per questa analisi viene utilizzata la medesima configurazione utilizzata in precedenza per la deduzione dell'incertezza del processo di ricostruzione tridimensionale del singolo marker.

Viene quindi utilizzato il medesimo marker vincolato alla flangia del robot. Il manipolatore antropomorfo ed il sistema di visione sono stati riferiti al medesimo SdR in modo tale da permettere il confronto tra il valore del riferimento e della ricostruzione. Il posizionamento del marker montato sulla flangia del robot viene considerato esente da errore e quindi perfettamente noto.

Viene programmato il controllore del robot affinché il marker, posizionato inizialmente nell'origine del sistema di riferimento, esegua tre movimenti di traslazione, lungo l'asse x , lungo l'asse y e lungo z . Per ogni traslazione il marker effettua un movimento di "andata", una pausa di due secondi nel punto di fine della traslazione ed un movimento di "ritorno" per rientrare in posizione di riposo (nel caso in oggetto, coincidente con l'origine degli assi). Traendo le dovute conclusioni, il marker riproduce la terna del Work Object.

Le posizioni dove il robot staziona sono le seguenti:

- l'origine degli assi $\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$;
- il punto di coordinate $\begin{pmatrix} 200 & 0 & 0 \end{pmatrix}$, lungo l'asse x ;
- il punto di coordinate $\begin{pmatrix} 0 & 200 & 0 \end{pmatrix}$, lungo l'asse y ;
- il punto di coordinate $\begin{pmatrix} 0 & 0 & -200 \end{pmatrix}$, lungo l'asse z .

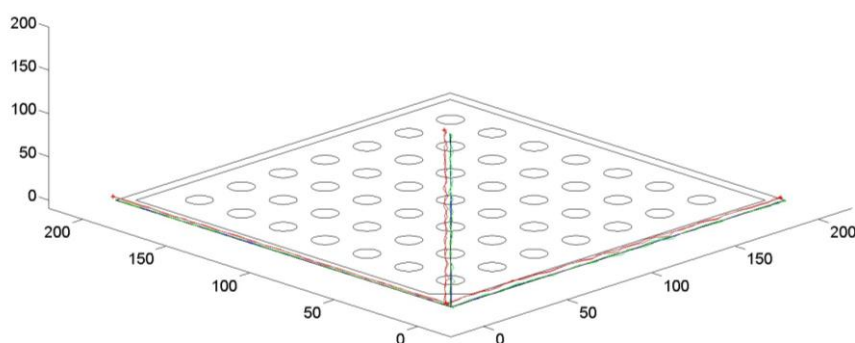


Figura 6-37: Ricostruzione degli assi della terna di riferimento acquisiti attraverso il sistema di visione.

Il movimento di riferimento e la ricostruzione sono riportati in figura 6.39. Con il colore rosso viene tracciata la ricostruzione 3D degli assi ottenuta sfruttando il

sistema di visione. Non essendo stato applicato alcuno stadio di filtraggio, l'acquisizione appare sporcata dal rumore di acquisizione. Il sistema di riferimento ricostruito appare, inoltre, traslato e ruotato rispetto al riferimento. Lo scostamento presente tra il SdR ricostruito e quello di riferimento è indicativo dell'accuratezza del processo di ricostruzione tridimensionale. Dato che l'offset in posizione ed orientamento sono di entità non trascurabile è possibile applicare l'algoritmo di registrazione rigida ai quattro punti dello spazio cartesiano evidenziati in precedenza. Si valuta un vettore traslazione così composto

$$\bar{v} = \begin{bmatrix} \varepsilon_x & \varepsilon_y & \varepsilon_z \end{bmatrix}^T$$

dove $\varepsilon_x = 1.16$ mm, $\varepsilon_y = 0.45$ mm e $\varepsilon_z = -0.28$ mm. Quindi

$$\bar{v} = \begin{bmatrix} 1.16 & 0.45 & -0.28 \end{bmatrix}^T$$

Gli errori sull'orientamento assumono invece questa forma

$$\bar{\alpha} = \begin{bmatrix} \varepsilon_\phi & \varepsilon_\theta & \varepsilon_\psi \end{bmatrix}^T$$

dove $\varepsilon_\phi = -0.01^\circ$, $\varepsilon_\theta = 0.11^\circ$ e $\varepsilon_\psi = -0.24^\circ$. Quindi

$$\bar{\alpha} = \begin{bmatrix} -0.01 & 0.11 & -0.24 \end{bmatrix}^T$$

L'errore di offset di posizione complessivo, per l'origine, può quindi essere valutato con la formula

$$\|\bar{v}\| = \sqrt{\varepsilon_x^2 + \varepsilon_y^2 + \varepsilon_z^2}$$

Che genera come risultato un valore di $\varepsilon = 1.27$ mm. Una volta applicata la traslazione v alle misure, si ottiene la traiettoria tracciata in verde in figura 6.39 che, filtrata attraverso un passa-basso di frequenza di taglio 5 Hz, genera la traiettoria blu. Questa ricostruzione, si può notare, riproduce in modo fedele gli assi di riferimento. È necessario effettuare questa correzione ogni volta che viene ridefinito il WorkObject del manipolatore o la terna di riferimento del sistema di visione. Sono definizioni che generalmente vengono effettuate solamente al momento del setup della cella.

6.2.6 Incertezza del sistema di acquisizione finale

Fino ad ora sono state testate puramente le prestazioni del sistema trascurando l'applicazione finale. Sono state affrontate verifiche delle prestazioni statiche e dinamiche del sistema telecamere-robot preso singolarmente per poterlo caratterizzare al meglio, escludendo le fonti di incertezza dovute all'algoritmo di ricostruzione della posizione e dell'orientamento della punta del pennello. Finora è stato infatti ricostruito un unico marker e sugli andamenti delle sue coordinate sono state effettuate le valutazioni di accuratezza ed incertezza del sistema. Ora è necessario però complicare il problema introducendo, nella valutazione dell'incertezza di misura, il processo di estrazione delle coordinate della punta del pennello. Per fare ciò, il procedimento è molto simile a quello seguito in sezione 6.2.3. Il pennello dotato di marker viene fissato sulla flangia del robot e movimentato in diverse posizioni dell'area di lavoro. È importante che venga mantenuto frontale al sistema di acquisizione, in maniera che tutti i marker siano visibili contemporaneamente. È una prova statica quindi il robot rimarrà fermo durante l'acquisizione delle immagini. Per ogni differente posizione del pennello, vengono acquisite 200 immagini dalle quali verranno estratte le coordinate u, v dei blob per poter permettere la successiva ricostruzione tridimensionale della posizione dei marker. Chiaramente, la posizione di ogni singolo marker presenterà l'incertezza calcolata in sezione 6.2.3. Dopo aver riordinato i marker, viene applicato l'algoritmo di registrazione rigida. L'algoritmo, essendo un algoritmo di minimizzazione, introdurrà ulteriore incertezza di cui vogliamo valutare l'entità. Il test è necessario per apprendere di quanto peggiori il valore dell'incertezza, indicato dal valore dello scarto tipo, dopo aver applicato l'algoritmo. Dalle traiettorie ottenute della posizione del centro viene estratto il vettore scarto tipo σ_i , dove i rappresenta l'indice della posa del pennello. Viene calcolato uno scarto quadratico medio complessivo σ_{av} come media dei σ_i . In figura 6.40 vengono rappresentate le distribuzioni normalizzate delle misure delle coordinate X, Y e Z della punta, aventi media nulla e scarto quadratico medio pari a σ_{av} .

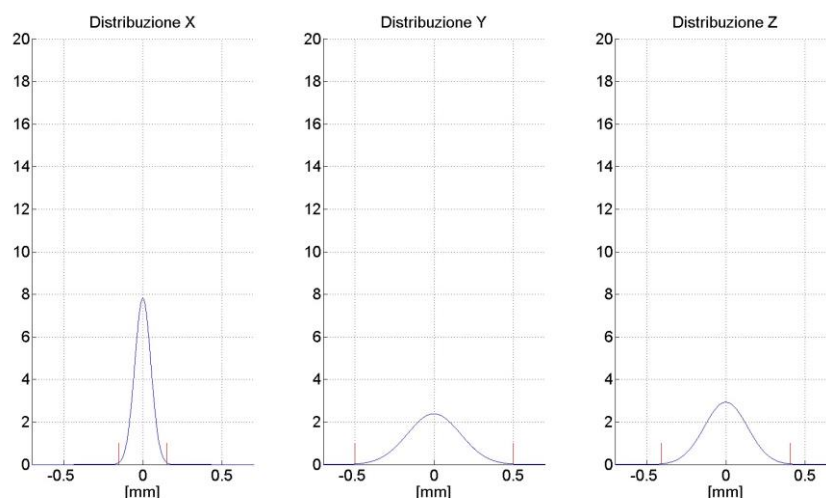


Figura 6-38: Distribuzione normalizzata delle misure di posizione della punta del pennello, non filtrate.

Attraverso il valore di σ_{av} è possibile avere un'informazione di dispersione delle misure sperimentali: esso fornisce una stima della variabilità delle misure acquisite attorno al valore atteso. Il vettore σ_{av} è così composto

$$\sigma_{av} = \begin{bmatrix} \sigma_x & \sigma_y & \sigma_z \end{bmatrix} = \begin{bmatrix} 0.051 & 0.167 & 0.136 \end{bmatrix}$$

da cui, attraverso la formula

$$\|\sigma_{av}\| = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$$

si ottiene un valore dello scarto tipo complessivo pari a 0.221 mm. Il valore di scarto tipo appena ottenuto implica che il 99.7% delle misure di posizione della punta del pennello potrà cadere in un intervallo di circa 1.3 millimetri, centrato sul valor medio. Un valore di variabilità così elevato non può essere accettato per l'applicazione studiata. Filtrando in frequenza le misure originali, con frequenza di taglio pari a 5 Hz, è stato possibile ridurre il rumore presente sulle misure di X, Y e Z della punta del pennello. Dopo l'applicazione del filtro, la distribuzione delle misure è cambiata come si vede in figura 6.41).

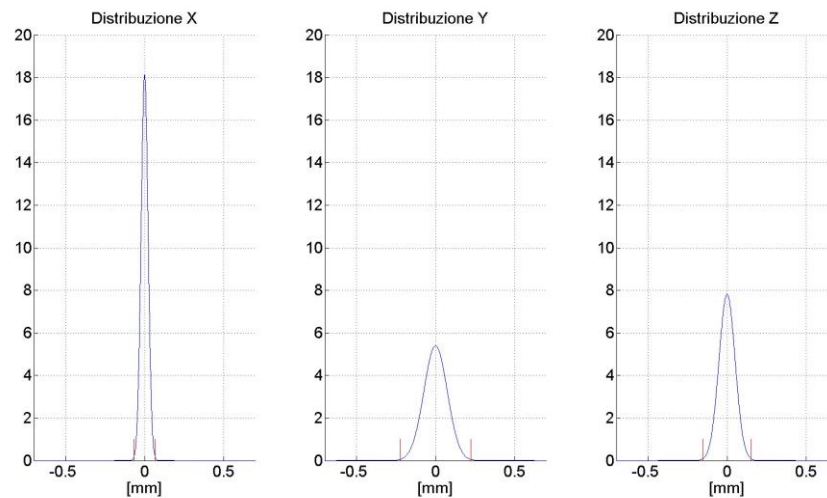


Figura 6-39: Distribuzione normalizzata delle misure di posizione della punta del pennello, filtrate attraverso un filtro di frequenza di taglio pari a 5 Hz.

Il vettore σ_{av} è così composto

$$\bar{\sigma} = \begin{bmatrix} \sigma_x & \sigma_y & \sigma_z \end{bmatrix} = \begin{bmatrix} 0.022 & 0.074 & 0.051 \end{bmatrix}$$

da cui, attraverso la formula

$$\|\sigma_{av}\| = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$$

otteniamo un valore dello scarto tipo complessivo pari a 0.093 mm, valore ridotto rispetto a quello precedentemente ottenuto. Si può notare, inoltre, come il valore di ripetibilità di misura lungo l'asse x sia notevolmente minore rispetto ad y e z. Considerando quindi che il 99.7% delle misure cade all'interno dell'intervallo $[-3\sigma, +3\sigma]$ si osserva che la procedura di calcolo delle coordinate della punta del pennello avrà una variabilità di $0.022 \cdot 6 = 0.132$ mm lungo la direzione x, di $0.074 \cdot 6 = 0.444$ mm lungo la direzione y e di $0.051 \cdot 6 = 0.306$ mm lungo la direzione z.

Una situazione che potrebbe potenzialmente presentarsi, durante l'acquisizione delle immagini, è quella in cui uno dei marker venga a trovarsi nascosto o dallo stelo del pennello o da un altro marker. Dato che l'algoritmo sfrutta una minimizzazione ai minimi quadrati il risultato fornito in caso di tre marker visibili

anziché quattro, potrebbe non essere il medesimo. È quindi interessante osservare come si modifica il valore dell'incertezza sul processo di ricostruzione della posizione della punta qualora si verificasse questo fenomeno. La prova effettuata è, ancora una volta, una prova statica. Il pennello viene mostrato in modo che l'insieme dei marker appaia frontalmente al sistema di misura. Uno per volta vengono oscurati tutti i marker e ne vengono osservati i risultati prodotti sulla ricostruzione della posizione della punta, a livello di incertezza di misura. Il pennello dotato di marker possiede la struttura in figura 6.42.

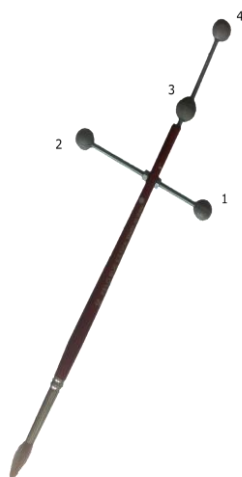


Figura 6-40: Pennello dotato di marker utilizzato per la prova e la decorazione, in cui i marker sono stati numerati da uno a quattro.

Nr Marker Mancante	σ_x (mm)	σ_y (mm)	σ_z (mm)	σ_{tot} (mm)
nessuno	0.022	0.074	0.051	0.093
1	0.012	0.195	0.062	0.204
2	0.030	0.117	0.082	0.145
3	0.018	0.097	0.109	0.147
4	0.023	0.121	0.094	0.154

Tabella 6.2: Valori, ottenuti sperimentalmente, dello scarto tipo delle misure di posizione della punta del pennello lungo le tre direzioni x , y e z in caso di oscuramento di un marker.

I risultati ottenuti dalle prove sono riportati in tabella 6.2. In tutti i casi di perdita di un marker durante l'acquisizione si verifica un netto peggioramento dell'incertezza di misura della procedura di estrazione delle coordinate della punta.

Fino ad ora è stato verificato il comportamento della procedura di ricostruzione tridimensionale della posizione della punta in caso di marker mostrati frontalmente alle telecamere e in caso di perdita di visibilità su un marker. Non è ancora stato affrontato, invece, il caso in cui lo stelo del pennello fosse parallelo all'asse z del riferimento del sistema di visione, e ruotasse attorno al suo asse, avvicinando tra loro le proiezioni dei marker nell'immagine e creando oscuramenti parziali sui singoli marker. Il caso è mostrato in figura 6.43. Sono state effettuate delle prove statiche, fissando il pennello con i marker sulla flangia del robot e sono state effettuate acquisizioni, ruotando il pennello di 30° alla volta, attorno all'asse del pennello, in senso orario. Si vuole capire quale sia l'effetto dell'orientamento del pennello attorno al suo asse sul valore dell'incertezza della ricostruzione tridimensionale della posizione della punta. Per ciascun orientamento vengono acquisite 200 immagini e, quindi, 200 posizioni della punta del pennello. Da ogni serie viene estratto il vettore scarto quadratico medio σ^2 , scomponibile lungo le tre direzioni in σ_x , σ_y e σ_z .



Figura 6-41: Rotazione in senso orario del pennello.

I risultati sono riportati nella tabella seguente:

Rotazione (gradi)	σ_x (mm)	σ_y (mm)	σ_z (mm)	σ_{tot} (mm)
0	0.022	0.074	0.051	0.093
30	0.038	0.078	0.059	0.104
60	0.015	0.046	0.053	0.072
90	0.055	0.072	0.063	0.110
120	0.018	0.062	0.055	0.085
150	0.029	0.067	0.065	0.098
180	0.010	0.052	0.035	0.063

Tabella 6.3: Valori, ottenuti sperimentalmente, dello scarto delle misure di posizione della punta del pennello tipo lungo le tre direzioni x , y e z al variare dell'orientamento del pennello.

In tabella sono riportati solamente i risultati di rotazioni tra 0° e 180° poiché tra 180° e 360° la situazione è la medesima, data la conformazione del pennello. Osservando la tabella si può notare che non si presentano sostanziali variazioni per quanto riguarda l'incertezza di misura, come ci si sarebbe potuto invece aspettare, al variare dell'angolo. Oscuramenti parziali dei marker non provocano sostanziali variazioni nel risultato della procedura di ricostruzione tridimensionale al contrario del caso in cui si verifica la perdita di visibilità di uno dei quattro marker. È stata effettuata l'analisi solamente per rotazioni attorno all'asse z in quanto ritenute le più critiche perché più probabili. Durante il processo di decorazione manuale, infatti, è spesso presente il caso in cui il pennello sia fatto ruotare attorno al suo asse. Si presenta invece molto più raramente il caso in cui il pennello venga fatto ruotare significativamente attorno agli assi x ed y del sistema di riferimento, posizionandolo quasi orizzontale.

6.2.7 Conclusioni

È stato qui presentato un metodo innovativo per la programmazione offline di un robot per la decorazione di supporti ceramici. Il metodo proposto è interamente automatizzato, tuttavia l'accuratezza risultante dipende fortemente dalla procedura di calibrazione. Il metodo in oggetto ha il vantaggio di essere flessibile e di non richiedere che le diverse componenti del sistema, quali robot e telecamere, siano posti in posizioni prestabilite. Esso è inoltre scalabile, permettendo di effettuare misure su aree di lavoro differenti.

Sono stati effettuati test per definire il grado di accuratezza del metodo proposto. I principali contributi all'accuratezza vengono forniti dalle procedure di registrazione dei sistemi di acquisizione e di attuazione robot. Questi tuttavia possono essere compensati, anche se non interamente, in fase di elaborazione dati eliminando così gli errori sistematici. Il grado di accuratezza raggiunto permette di produrre risultati conformi agli standard delle decorazioni manuali. Il metodo fornito permette di raggiungere uno scarto tipo del risultato del processo di estrazione della posizione della punta del pennello pari a 0.1 mm. Questo valore è quello raggiunto in condizioni standard, ovvero in tutte quelle condizioni per cui non si verifica la perdita di un marker. Per ottenere queste prestazioni è stato necessario filtrare in frequenza, con frequenza di taglio pari a 5 Hz, le misure di posizione della punta del pennello. Il valore ottenuto, senza l'intervento del filtro sarebbe stato quasi tre volte maggiore. Inoltre, si osserva come l'incertezza del processo di estrazione delle coordinate della punta del pennello sia un valore notevolmente minore lungo x , rispetto ad y e z . È stato ottenuto un vettore scarto tipo pari a $\bar{\sigma}$ così composto.

$$\bar{\sigma} = \begin{bmatrix} \sigma_x & \sigma_y & \sigma_z \end{bmatrix} = \begin{bmatrix} 0.022 & 0.074 & 0.051 \end{bmatrix}$$

Questo sbilanciamento nelle componenti del vettore ripetibilità è dovuto alla posizione relativa delle telecamere utilizzate per l'acquisizione. Esse infatti sono posizionate frontalmente all'asse x del piano di lavoro e inclinate di circa 60 gradi rispetto all'orizzontale. In questo modo la misura di x risulta quella meno affetta da

incertezza, mentre la restante incertezza viene spartita quasi equamente tra la misura di y e di z . Il risultato finale è un tratto fluido, paragonabile a quello di un decoratore.

Il sistema è stato verificato mediante una campagna sperimentale effettuata con ABB e un tecnico esperto in decorazioni di ceramiche. Alcuni risultati sono riportati nelle immagini seguenti.

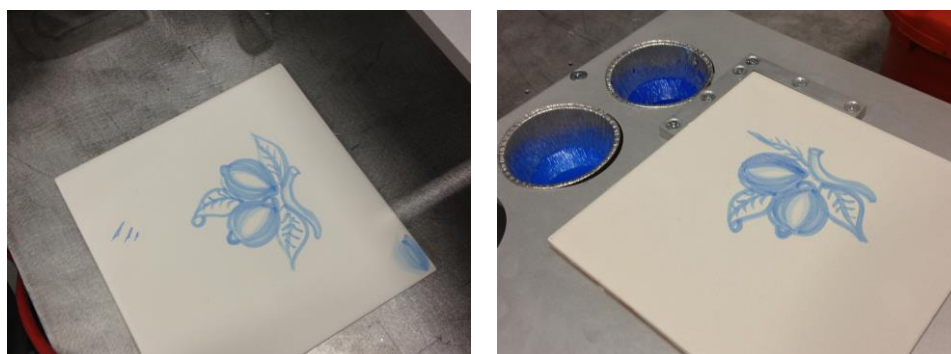


Figura 6-42: Confronto tra una piastrella decorata a mano ed una decorata dal manipolatore IRB 120: particolare effettuato a mano libera.



Figura 6-43: Confronto tra una piastrella decorata a mano ed una decorata dal manipolatore IRB 120: decorazione all'interno di contorni.

6.3 Bin picking

Lo studio di sistemi che siano in grado di individuare e afferrare oggetti ammassati o disposti casualmente all'interno di cassette o cassoni è attualmente di grande interesse nel settore dell'automazione industriale. Queste applicazioni vengono indicate con il nome di *bin picking*. Il contesto produttivo attuale, essendo molto spesso caratterizzato da lotti di produzione piccoli e differenziati tra loro, richiede una sempre maggiore flessibilità. E' necessario, quindi, garantire la possibilità di riconfigurare velocemente e semplicemente la cella di lavoro al cambio del prodotto. L'introduzione dei robot nelle linee di produzione, permettendo una manipolazione programmabile dei prodotti, consente di superare parte di questi limiti aumentando la flessibilità del processo. Tuttavia i robot, se non vengono dotati di sensori avanzati per la percezione dell'ambiente, vedono pesantemente limitata la propria flessibilità nelle fasi di asservimento. Ciò determina una diminuzione importante della flessibilità globale del sistema.

I sensori normalmente presenti in una cella di lavoro, infatti, sono nella maggior parte dei casi specializzati per un singolo compito, come la verifica della presenza e del passaggio di un pezzo o controlli di spessori e lunghezze. I manipolatori sono quindi "ciechi" e la classica programmazione a punti permette la replica di movimenti pre-programmati che obbliga i manipolatori a operare in un ambiente rigidamente strutturato con un asservimento in posizioni fisse. Sistemi basati su visione artificiale possono essere la chiave per superare tali limitazioni. La loro integrazione nel processo produttivo ha quindi come obiettivo l'incremento della flessibilità nelle operazioni di asservimento macchina o nelle lavorazioni con ripresa, rimuovendo i limiti dell'impostazione classica legati alla necessità di ordinare i pezzi da manipolare.

La realizzazione di un sistema di bin picking richiede di affrontare in modo sinergico una serie di problematiche differenti tra loro. In primo luogo sono necessarie procedure in grado di riconoscere in maniera robusta i pezzi da afferrare all'interno della scena. In alcuni casi di presa da nastro o da piano, in cui i pezzi sono correttamente singolarizzati (non sovrapposti e distanziati l'uno dall'altro) e

non si necessita di informazioni sulla profondità per il loro riconoscimento, algoritmi di identificazione 2D sono adeguati allo scopo. In tutte le altre situazioni, soprattutto quando il volume di lavoro si sviluppa in profondità e/o i pezzi sono disposti casualmente uno sull'altro, sono necessarie procedure che permettano di ricavare la posa tridimensionale dell'oggetto nella scena. Si rende necessario, quindi, lo sviluppo di metodologie di riconoscimento basate su algoritmi di pattern matching 3D che permettano di riconoscere diverse tipologie di pezzi modificando unicamente alcuni parametri di controllo. Infatti un'ipotetica necessità di implementare procedure di riconoscimento specifiche per il singolo pezzo lederebbe la possibilità di ottenere soluzioni flessibili. Queste metodologie devono essere inoltre caratterizzate da una robustezza al rumore di misura.

Un sistema di misura adeguato per le operazioni di bin picking deve essere in grado di ottenere una ricostruzione 3D della scena che sia sufficientemente accurata e con risoluzione adeguata per far sì che la procedura di riconoscimento individui correttamente i pezzi. La scelta del sistema di misura più opportuno è dovuta al contesto produttivo e viene guidata da una serie di fattori specifici come la forma, le dimensioni, le caratteristiche superficiali degli oggetti e i tempi di ciclo richiesti.

Tra le tecniche di ricostruzione tridimensionale descritte nel paragrafo 2.1, le principali alternative che possono essere utilizzate in ambito industriale sono sistemi basati su stereoscopia e triangolazione attiva. All'interno del VBLab è già stata realizzata una soluzione al problema del bin-picking basata sulla triangolazione attiva. Questa è stata poi sviluppata e industrializzata dalla SpinOff del Politecnico di Milano chiamata ISS (riferimento sito ISS) ed è attualmente presente sul mercato in diverse parti del mondo. Partendo dalle conoscenze già maturate all'interno del VBLab si è voluto realizzare una soluzione alternativa basata sulla stereoscopia. La scelta di sviluppare una seconda alternativa di misura ha permesso di confrontare pregi e difetti delle varie soluzioni e comprendere in quali contesti risultino maggiormente adeguate. Il problema che si intende risolvere in questo lavoro è solamente quello della ricostruzione 3D della nuvola di punti che rappresenta il cassone, poiché gli ulteriori passaggi quali l'algoritmo di pattern matching 3D, la generazione della traiettoria migliore per prelevare il pezzo,

algoritmi di *collision avoidance* e progettazione della pinza, sono già stati affrontati e risolti in passato durante lo sviluppo della soluzione con triangolazione laser.

Sono state poi implementate procedure di riconoscimento basate su algoritmi di pattern matching. I modelli di ricerca possono provenire o da un modello CAD 3D o, tramite reverse engineering, dalla misura del pezzo stesso.

E' infine presentata la qualificazione metrologica degli strumenti di misura sviluppati. Essa si compone di due parti: l'analisi teorica dell'incertezza di misura e la verifica sperimentale di tali incertezze, in modo da verificarne l'adeguatezza nel caso di studio (capitolo 5). E' stata inoltre effettuata un'analisi di sensitività per investigare come un'errata stima dei parametri di calibrazione possa influire sulla misura.

In ultima istanza, sono state effettuate due campagne sperimentali (una per il sistema stereoscopico, una per quello a triangolazione attiva) in cui diversi pezzi disposti alla rinfusa sono stati misurati e riconosciuti. Le relative pose sono state comunicate a un robot antropomorfo che ha correttamente proceduto alla loro presa. Questa è la migliore verifica per dimostrare l'affidabilità del sistema realizzato.

6.3.1 Setup e taratura del sistema

Il sistema stereoscopico per il bin picking sviluppato in questo lavoro deve essere in grado di misurare il 3D di oggetti disposti alla rinfusa in un cassone. Dalla nuvola di punti ottenuta, ovvero la discretizzazione dell'oggetto di misura con punti di cui si conoscono le coordinate nello spazio, un algoritmo di pattern matching individua le pose dei singoli pezzi che in seguito possono essere prelevati da un robot antropomorfo. La cella di lavoro progettata è composta da:

- Sistema di visione, costituito fisicamente da una telecamera e un illuminatore.
- Sistema di movimentazione, costituito da un braccio robotico antropomorfo.
- Una stazione di cambio utensile nella quale sono alloggiata la pinza e la telecamera.
- Un PC per l'elaborazione dei dati.

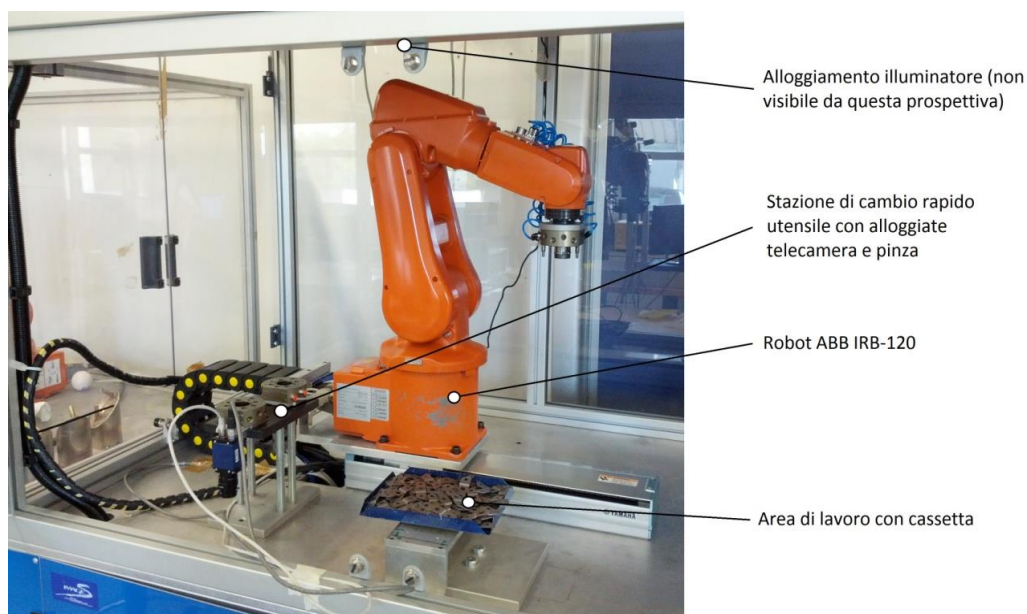


Figura 6-44: *Visione d'insieme della cella di lavoro.*



Figura 6-45: *Particolari della stazione di cambio utensile (sinistra) e dell'illuminatore (destra).*

In particolare il sistema stereoscopico sviluppato in questo lavoro si basa sul principio dell'hand eye camera, ovvero per acquisire le diverse immagine necessarie per la misurazione 3D si utilizza solamente una telecamera movimentata opportunamente da un robot. Questa tipologia di sistema garantisce una notevole flessibilità poiché consente di provare diverse configurazioni di posizionamento e distanza delle telecamere semplicemente agendo sul software robot senza dover fisicamente modificare la struttura di supporto. Inoltre l'utilizzo di una sola telecamera riduce i costi del sistema.

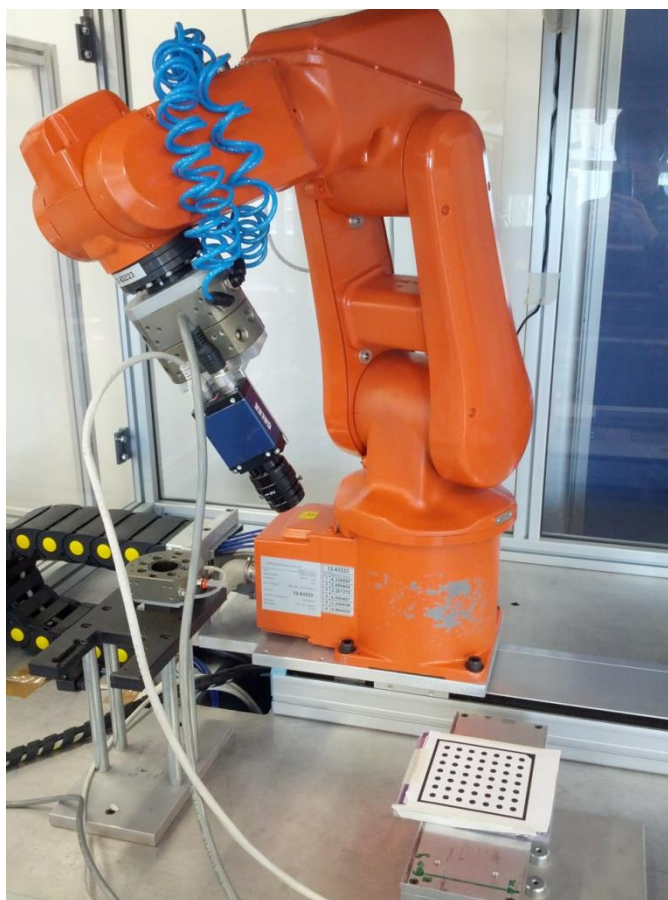


Figura 6-46: *Telecamera movimentata da robot in configurazione hand-eye.*

La scena deve però presentare caratteristiche statiche poiché, dato che le immagini acquisite devono necessariamente inquadrare la scena nella stessa posizione, è necessario tenere conto dei tempi di movimentazione necessari a spostare il braccio robotico nelle posizioni stabilite. Nelle operazioni di bin picking questa condizione è sempre verificata, quindi una configurazione di questo tipo può essere utilizzata senza problemi.

Con lo scopo di approfondire e studiare un sistema basato su stereoscopia, un sistema composto da una sola telecamera permette di provare diverse configurazioni di posizionamento e di illuminazione per valutare le prestazioni e l'incertezza sulle misure. In ambito industriale è possibile, una volta indentificata la configurazione che restituisca risultati conformi alle specifiche volute, procedere

all'allestimento di un sistema definitivo in cui la movimentazione tramite robot è sostituita da un sistema multi-camera.

In questo lavoro si è deciso di utilizzare una telecamera Dalsa Genie HM1400 le cui specifiche tecniche sono mostrate in figura :

Feature	Specification
Part Number	CR-GM0x-H140x
Resolution	1400 x 1024
Total Data Rate	92 MB/s
Max. Line Rate	75 fps
Pixel Size	7.4 μ m
Number of Camera Taps	1
Output Format	GigE Vision
Size	44 mm x 29 mm x 67 mm



Figura 6-47: Datasheet della telecamera Genie HM1400.

Tale scelta è motivata in primo luogo dalla necessità di acquisire immagini ad alta risoluzione. Un maggior numero di pixel nell'immagine permette una ricostruzione stereoscopica maggiormente dettagliata. La telecamera scelta ha un sensore con risoluzione di 1400x1200 pixel, elevata per gli standard delle telecamere industriali. Inoltre la Genie HM1400 ha la possibilità di essere comandata da un segnale di trigger esterno. Questa caratteristica è fondamentale in quanto permette di sincronizzare l'acquisizione delle immagini con il movimento del manipolatore, in particolare si deve essere in grado di acquisire l'immagine quando il robot raggiunge le posizioni stabilite. Il robot presenta una serie di output digitali che possono essere comandati via software e utilizzati come segnali di trigger.

Alla telecamera è stata accoppiata un'ottica caratterizzata da una lunghezza focale di 12 mm. Quest'ottica è stata scelta perché consente di inquadrare, all'interno della cella di lavoro in esame, l'intero volume di lavoro.

Il robot utilizzato in questo lavoro è l'ABB IRB-120, un manipolatore antropomorfo a 6 assi. Questo robot manipola un carico massimo di 3 kg (4 kg con polso verso il basso) e ha un'estensione del volume di lavoro di 580 mm.

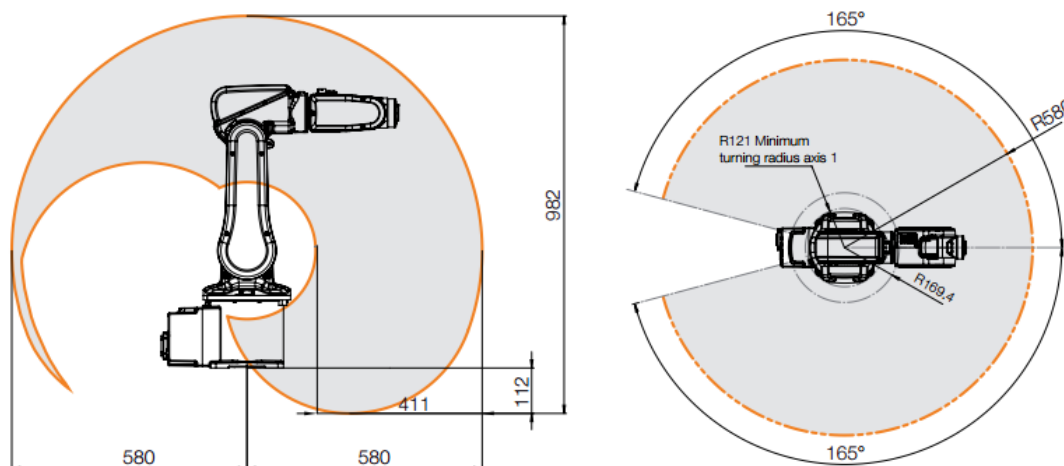


Figura 6-48: Volume di lavoro del manipolatore ABB IRB-120.

La sua funzione all'interno del sistema è duplice: movimentata la telecamera per l'acquisizione delle immagini e successivamente afferra con una pinza il pezzo da prelevare. Dato che l'obiettivo del sistema è quello del bin picking, la pinza utilizzata dipende dalla geometria del pezzo da prendere. Tuttavia il sistema sviluppato non pone limiti sulla scelta dei componenti da afferrare. Dimensionando opportunamente la pinza e il robot è possibile adattare il sistema a qualsiasi tipologia di pezzo. Sia la telecamera che la pinza sono fissate a dei moduli di cambio rapido utensile, che permettono al robot di agganciare l'utensile necessario tramite un azionamento pneumatico.

La programmazione dei movimenti e del ciclo di lavoro è codificata attraverso il linguaggio di programmazione Rapid e compilato con il software RobotStudio. Il raggiungimento della posizione e la dinamica del movimento sono gestiti da un controllore dedicato.

Il robot possiede inoltre un banco di 25 input e output digitali che consentono il controllo di elementi esterni, ad esempio l'elettrovalvola che comanda l'apertura e chiusura della pinza o il trigger della telecamera.

Per l'applicazione in studio è necessario che in ogni ciclo di acquisizione il robot si muova nella stesse posizioni con un'incertezza inferiore alla tolleranza di presa. Da datasheet forniti da ABB si ritrovano dei valori di quella che loro chiamano precisione statica, ovvero il più grande errore di posizione che si genera in seguito a un movimento del robot, e la ripetibilità, cioè la capacità del robot di posizionarsi nello stesso punto dello spazio effettuando, in istanti successivi, gli stessi movimenti. Per il manipolatore ABB IRB 120 il datasheet riporta un valore di precisione statica di 0.02 mm ed un valore di ripetibilità statica di 0.01 mm. Questo valore è di un ordine di grandezza inferiore rispetto all'incertezza sulla misura dello strumento, come sarà verificato in seguito, quindi si può considerare trascurabile.

La taratura del sistema di misura è stata fatta ancora una volta mediante la tecnica implementata nelle librerie di HALCON. In questo caso la calibrazione viene fatta facendo acquisire il calibratore fermo dalla telecamera in pose diverse tra loro. Così facendo è possibile calcolare i parametri intrinseci della telecamera. Acquisendo poi il calibratore ripreso solamente dalle pose utilizzate per la misura è possibile trovare gli estrinseci, vale a dire le pose relative.

L'ultimo passaggio necessario per terminare la calibrazione è la registrazione tra sistema di visione e robot. I due sistemi infatti fino a questo punto sono indipendenti tra loro e non è possibile utilizzare le posizioni ricavate direttamente dal pattern matching come riferimento di posizione per il robot. Questo perché robot e visione sono riferiti a due sistemi di coordinate diversi.

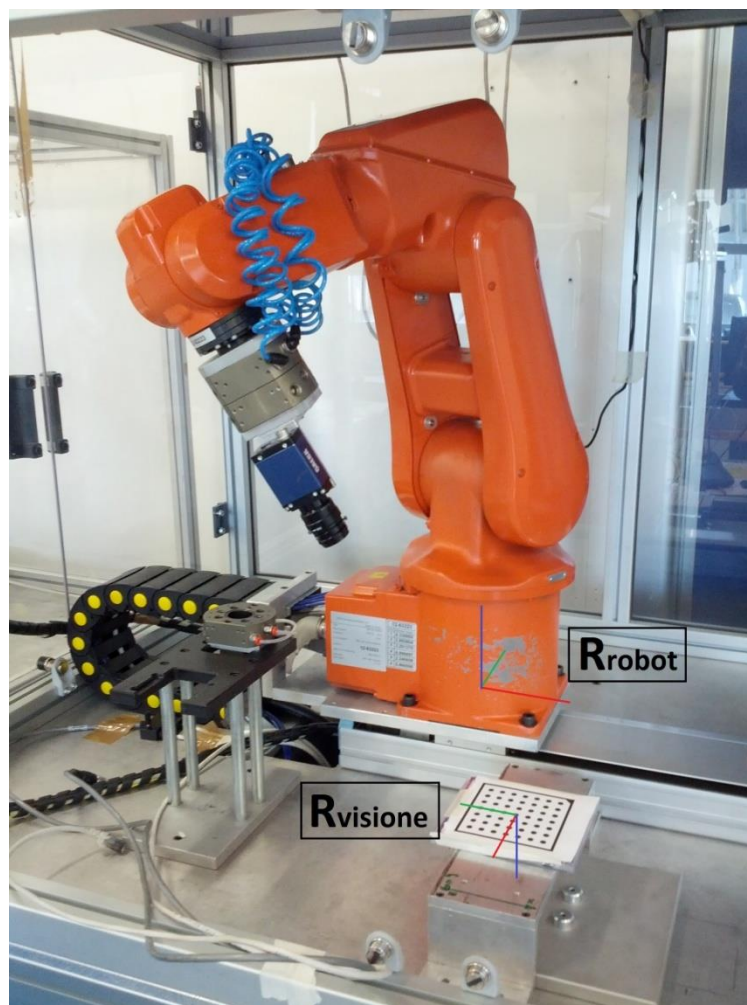


Figura 6-49: Panoramiche dei due sistemi di riferimento utilizzati in fase di registrazione.

In particolare per quanto riguarda il robot l'origine del sistema di riferimento (R_{robot}) è posizionata al centro della base. Il sistema di visione invece sfrutta il fatto che, come spiegato precedentemente, i parametri estrinseci sono ricavati relativamente a una posizione iniziale del calibratore. Per evitare cambi di coordinate e trasformazioni che potrebbero introdurre incertezze nelle misure si è deciso di utilizzare la posa iniziale del calibratore come sistema di riferimento principale ($R_{visione}$), sia per la ricostruzione stereoscopica, sia per il pattern matching. Tutte le coordinate dei punti ricavate tramite stereoscopia saranno quindi relative a questo sistema di riferimento. Successivamente, la posa stimata dal

pattern matching viene tralata nel sistema di riferimento robot mediante la trasformazione:

$$H_{pezzo_{robot}} = H_{calibratore_{robot}} * H_{pezzo_{visione}}$$

Dove $H_{pezzo_{robot}}$ è la posa del punto di presa in coordinate robot, $H_{calibratore_{robot}}$ è la posa del calibratore sempre in coordinate robot e $H_{pezzo_{visione}}$ è la posa del punto di presa stimata dal pattern matching, ovvero riferita al sistema di riferimento solidale al calibratore. Il vantaggio principale che si ottiene in questo modo è che la posizione $H_{pezzo_{robot}}$ ottenuta dopo il cambio di sistema di riferimento è direttamente il set point di posizione da inviare al robot per la presa.

L'informazione che manca per poter effettuare il calcolo è la posizione del calibratore rispetto al sistema di riferimento R_{robot} . Il modo più semplice ma allo stesso tempo preciso di ricavarlo è puntare con l'end effector il calibratore e utilizzare queste posizioni per ricavarne posizione e orientamento rispetto alla terna R_{robot} . Per individuare la posizione di una terna ed il suo orientamento nello spazio sono necessari minimo tre punti: il primo punto definisce l'origine del sistema di coordinate, il secondo permette di individuare la direzione di uno degli assi, nel caso specifico l'asse x, definita come la direzione del segmento che unisce il primo punto con il secondo; il terzo punto deve essere preso perpendicolare alla retta che unisce i primi due. In questo modo si definisce l'orientamento dell'asse y. In ultimo da tre punti passa uno ed un solo piano, calcolando la normale relativa a questo piano e passante per il primo punto si ottiene l'asse z.

Un modo semplice per poter ottenere punti che soddisfino queste caratteristiche con elevata accuratezza è quello di puntare con il robot i marker stampati sul calibratore.

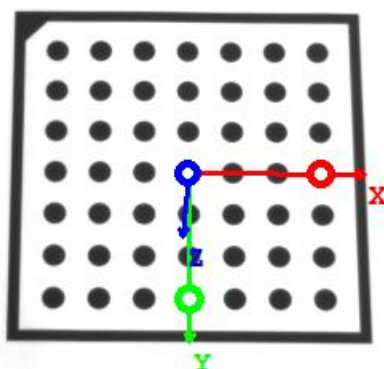


Figura 6-50: Marker del calibratore scelti come punti notevoli per la procedura di registrazione.

Per rendere il procedimento più veloce si potrebbe semplificare introducendo l'ipotesi che il calibratore sia parallelo al piano individuato dagli assi x e y della terna di riferimento R_{robot} . In questo modo sarebbero sufficienti due punti solamente per il calcolo dell'orientamento dell'asse x .

Dato che lo scopo del prototipo realizzato durante il lavoro è quello di testare diverse soluzioni per l'approccio al problema del bin picking tramite stereoscopia, è stato necessario mettere a punto un'interfaccia software che permetta di gestire le varie fasi del processo, la comunicazione con il robot e un'interfaccia con l'utente. Questo software è stato realizzato in C#, tramite Visual Studio 2010. L'interfaccia permette di scegliere prima di ogni acquisizione la configurazione voluta tra le quattro precedentemente trattate. Dopodiché esegue l'acquisizione delle immagini e la successiva ricostruzione tridimensionale della scena. Infine permette di visualizzare graficamente i risultati ottenuti dal pattern matching e di scegliere il pezzo da prendere. Alla fine di ogni ciclo salva su disco le immagini, la nuvola di punti e le pose ricavate per consentire un'eventuale lavoro di verifica offline.

Il robot è stato programmato tramite il software Robotstudio e la comunicazione robot-PC è basata sul protocollo TCP/IP. In questo modo, attraverso l'invio di opportune stringhe, si sincronizza l'esecuzione del codice HALCON con i movimenti del robot.

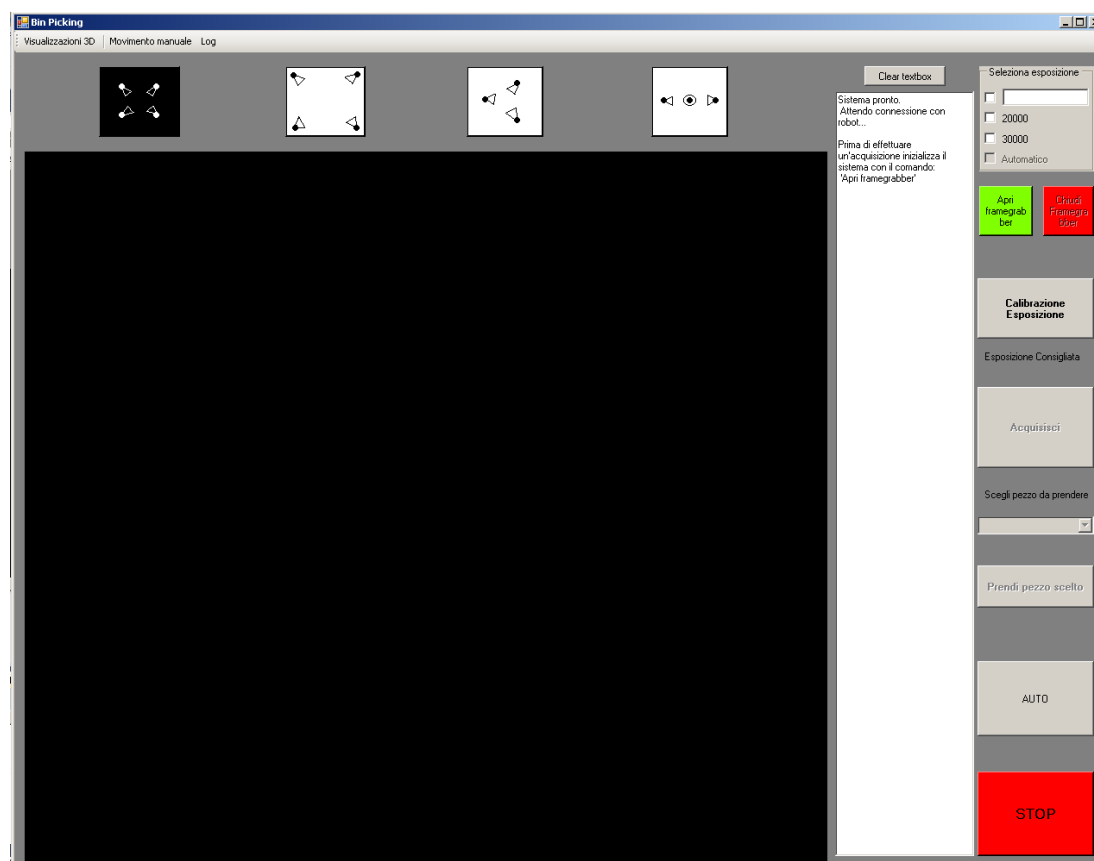


Figura 6-51: *Interfaccia utente.*

6.3.2 Illuminazione

Inizialmente il problema della determinazione della nuvola di punti è stato affrontato utilizzando la luce ambiente e regolando il tempo di esposizione della telecamera per avere delle immagini correttamente illuminate. Questa soluzione è risultata molto svantaggiosa nel caso in cui i pezzi che il robot deve prelevare, dunque i pezzi da trovare con l'algoritmo di pattern matching, risultino molto uniformi e non presentino numerosissime caratteristiche particolari (chiamate feature) o texture. L'algoritmo di pattern-matching 3D infatti necessita di un numero elevato di punti da riconoscere nella nuvola di punti, presenti omogeneamente su tutta la superficie del pezzo. Il processo di ricerca dei punti notevoli nelle immagini fa uso principalmente di tecniche di edge detection, ovvero identifica solamente quelle zone all'interno delle immagini che presentano pixel

adiacenti con valori di luminosità molto diversi, identificate come bordi. Nella nuvola di punti ottenuta quindi vengono ricostruiti solamente i bordi degli oggetti presenti sulla scena, oppure le zone al loro interno che presentano texture come scritte, spigoli, linee di confine tra colorazioni differenti. Una nuvola di punti in cui sono ricostruiti solo i bordi degli oggetti non è sufficiente e non restituisce alcun risultato.

Per risolvere il problema si è considerato quindi di mettere a punto un sistema di illuminazione che possa facilitare il riconoscimento dei pezzi. Si passa quindi da una tipologia di ricostruzione stereoscopica passiva ad una attiva. Il sistema di illuminazione è stato sviluppato in modo da consentire allo stesso tempo di mantenere le condizioni di luce che permettano la visibilità dei bordi e nel frattempo cercare di aumentare artificialmente le texture visibili sui pezzi, in modo da renderli riconoscibili al pattern matching. In questo lavoro si sono messe alla prova diverse soluzioni di illuminazione per cercare di raggiungere questo obiettivo, confrontando i risultati ottenuti con diversi pattern proiettati e diverse tipologie sorgenti di illuminazione, a luce visibile e infrarossa.

Il primo apparecchio verificato è l'MPro110 della 3M, un proiettore dalle dimensioni molto compatte dotato di unità di illuminazione LED. Di seguito le principali specifiche tecniche:

Description		Specification
Imager	Technology	LCOS
	Size	11.9 mm (0.47")
	Image Resolution	640 x 480 Pixels
	Image Aspect Ratio	4:3
Projection Distance		305–1800 mm (12"–70")
Image Size		163–975 mm (6.4"–38.4")
Throw Ratio (Width/Distance)		0.53
Lens Offset		1:1 (none)
Inputs		VGA, Composite Video, DC Power
Compatible Resolution		VGA, SVGA, XGA
Refresh Rates		60 Hz

Figura 6-52: Specifiche tecniche del proiettore Mpro110.

Il principale vantaggio di questo apparecchio sta nel fatto che, una volta collegato al PC, è possibile proiettare qualsiasi immagine avendo così la possibilità di mettere alla prova un gran numero di trame differenti.

Le principali tipologie di texture sperimentate sono:

- Marker neri di varie dimensioni su sfondo bianco.
- Linee nere di vario spessore e direzione su sfondo bianco.
- Rumore bianco a gradazione continua.
- Rumore bianco solo in tonalità bianco e nero.

Di seguito sono presentate alcune immagini delle trame utilizzate:

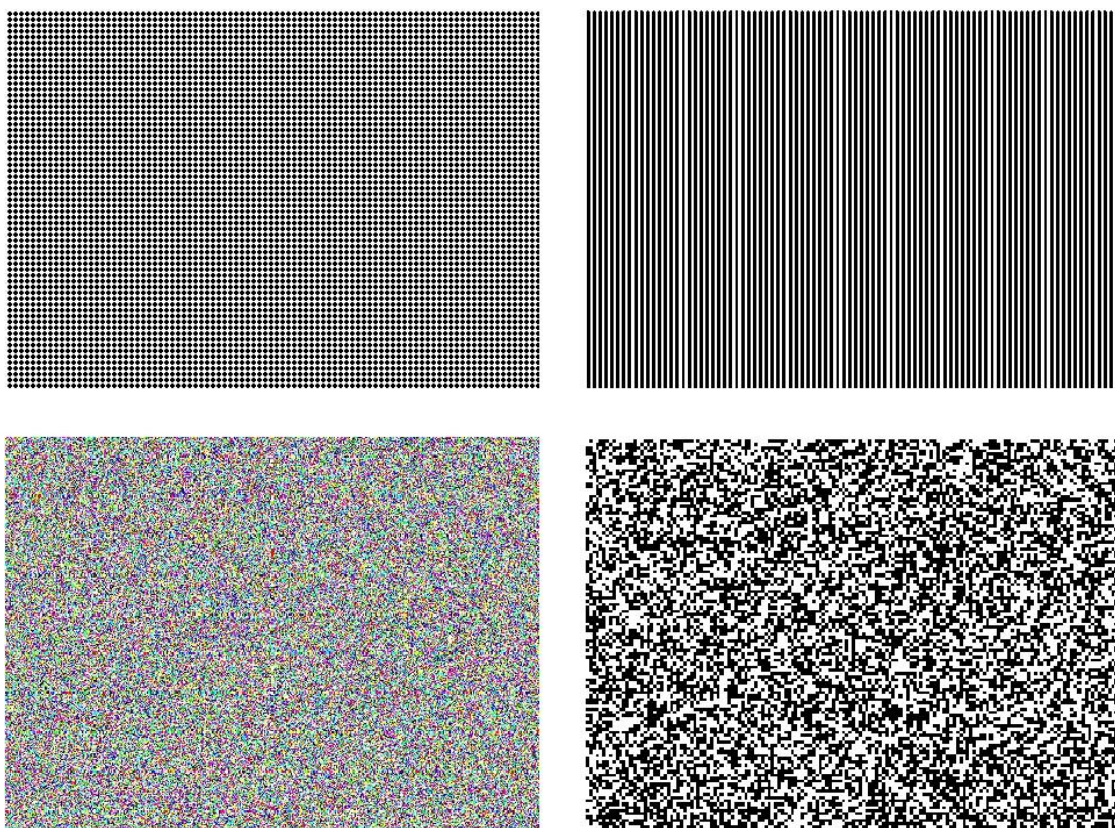


Figura 6-53: Esempi delle differenti trame provate. In ordine dall'alto a sinistra e in senso orario: punti, linee verticali, rumore bianco, rumore bianco in tonalità bianco e nero.

La texturizzazione mediante rumore bianco ha una duplice funzione. Infatti, oltre a rendere diversificata la superficie dei pezzi, permette di ridurre le ambiguità in fase di accoppiamento dei punti omologhi. Se le geometrie proiettate sono tutte uguali, troppo piccole e vicine tra loro può succedere che nelle varie immagini acquisite si creino delle ambiguità nel riconoscimento dei punti omologhi. Per

definizione in un immagine a rumore bianco ogni pixel ha un'intensità luminosa indipendente dagli altri quindi, in teoria, è possibile utilizzare trame più fitte, e quindi più dettagliate, che permettono di generare una nuvola di punti più omogenea.

Tutte le diverse trame proiettate con l'MPro110 sono state testate ma è subito emerso il principale limite di questa soluzione: la potenza del proiettore. Utilizzando il proiettore con la scena illuminata normalmente le immagini proiettate sono sovrastate dalla luminosità ambientale e diventano praticamente invisibili sia ad occhio nudo, sia alla telecamera.

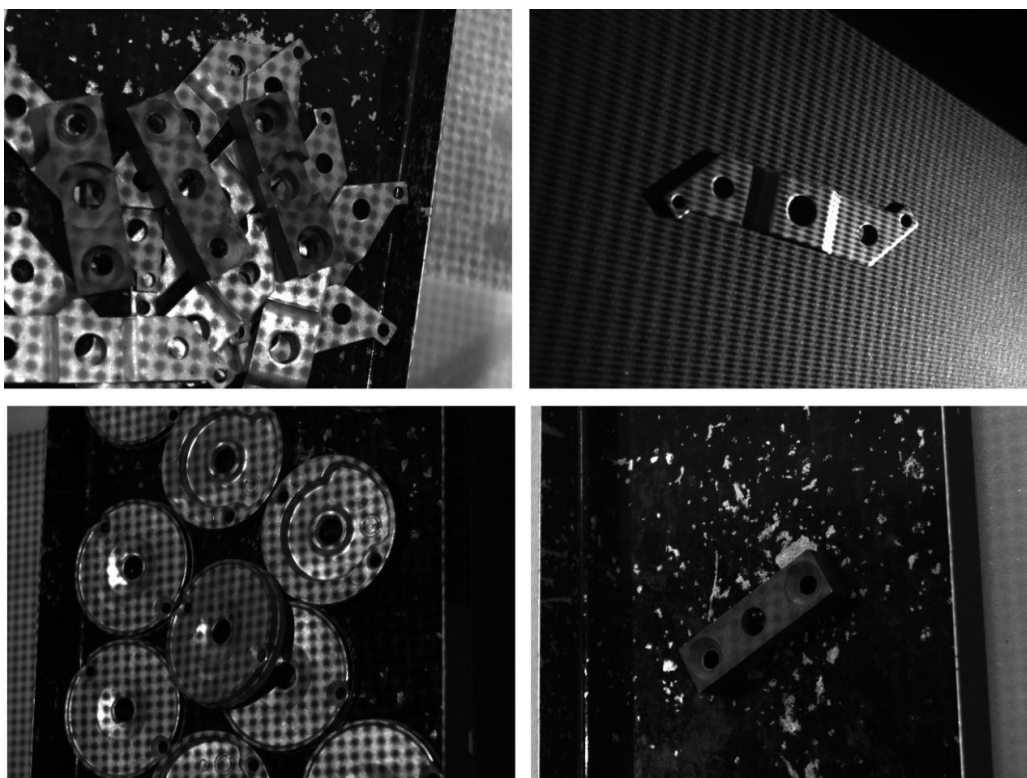


Figura 6-54: Esempi di diverse tipologie di pezzi illuminate con il proiettore MPro110. Si nota che la scarsa illuminazione ambientale necessaria a rendere le trame visibili compromette la qualità generale delle immagini.

L'unico modo per rendere visibili le trame è lavorare in ambiente completamente scuro, ed utilizzare il proiettore come fonte di illuminazione principale. Tuttavia la scarsa potenza fa sì che i vantaggi che si ottengono dall'aggiungere texture sui pezzi siano compensati dalla perdita di dettaglio dovuta

alla scarsa illuminazione. E' stato necessario quindi cercare un metodo che permettesse di rendere visibili i marker proiettati senza andare a modificare l'illuminazione esterna.

La soluzione adottata prevede dunque di utilizzare per l'illuminazione della scena e la proiezione delle texture due sorgenti di luce a lunghezze d'onda differenti, in modo che non interferiscano tra loro. La fonte di illuminazione della scena rimane la luce visibile (illuminazione ambientale) mentre per la proiezione delle texture si utilizzano gli infrarossi. In questo modo si può modificare l'illuminazione ambientale senza perdere definizione sul pattern proiettato.

Le caratteristiche che la sorgente di luce infrarossa deve soddisfare sono la capacità di proiettare un pattern luminoso che ad una distanza di circa un metro sia fitto ma non confuso e sufficientemente potente da essere visibile dalla telecamera in ogni condizione di luce ambientale. Un dispositivo che soddisfa adeguatamente i requisiti è il proiettore presente all'interno del Microsoft Kinect. Questo dispositivo, ideato come controller di gioco per la console Xbox360, permette di eseguire ricostruzioni tridimensionali tramite luce strutturata. In questo lavoro però non utilizziamo tutte le sue potenzialità ma ci limitiamo a utilizzarne l'illuminatore infrarossi. Tale scelta è dettata dal fatto che il dispositivo ha dimensioni molto ridotte ed ha un costo molto basso. Tuttavia il proiettore integrato in questa soluzione non è programmabile, quindi non è possibile modificare la griglia diffusa.

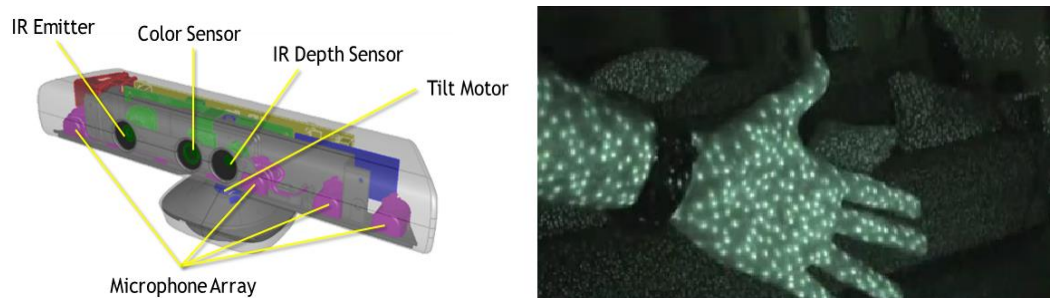


Figura 6-55: Schema del Microsoft Kinect e immagine a infrarossi della griglia proiettata.

In seguito ad alcune prove effettuate la soluzione si è rivelata adatta, infatti i punti risultano fitti, ben definiti su tutta la superficie e soprattutto sempre visibili dalla telecamera, con qualsiasi condizione di illuminazione. La ricostruzione ottenuta dai pezzi illuminati in questo modo risulta dettagliata a sufficienza per poter essere utilizzata per il pattern matching con buoni risultati.

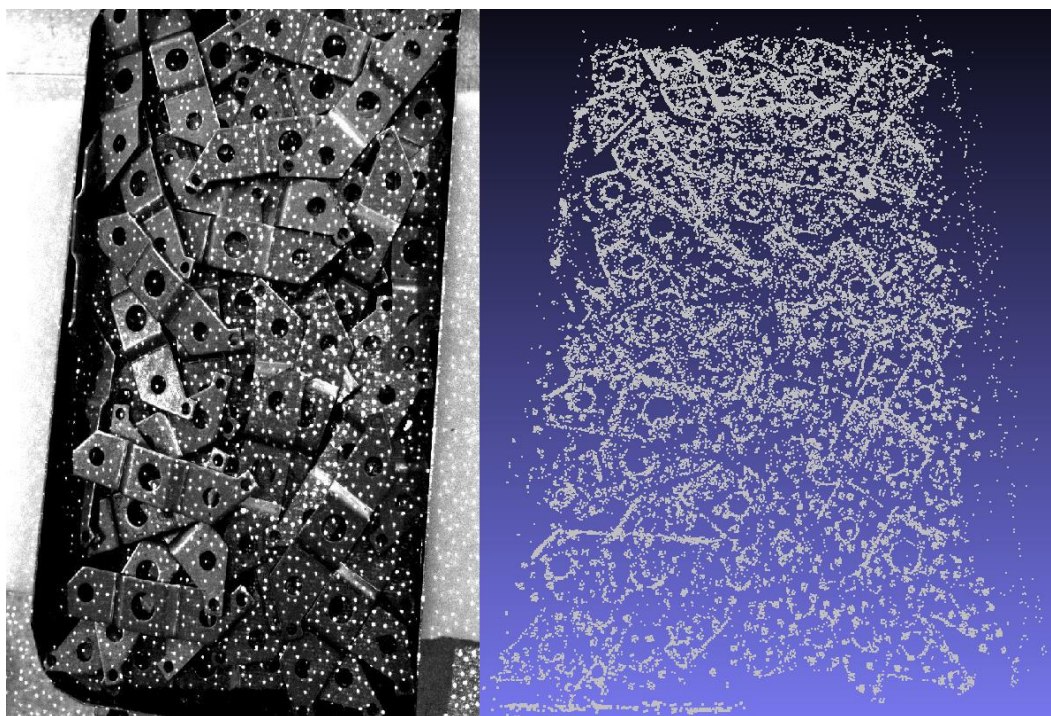


Figura 6-56: Ricostruzione tridimensionale della cassetta di pezzi illuminata con proiettore a infrarossi. Si notano nella prima immagine i punti della trama proiettata.

6.3.3 Scelta delle posa della telecamera

Il primo passo effettuato è stato uno studio sull'influenza del numero di immagini e il posizionamento relativo delle telecamere sull'incertezza di misura, in modo da definire la configurazione che garantisca la ricostruzione 3D migliore.

Come spiegato nel paragrafo 3.1 la ricostruzione stereoscopica si basa sulla ricerca all'interno delle immagini dei punti omologhi. Questi sono i punti che nelle varie immagini rappresentano lo stesso punto fisico. Maggiore è il numero di queste corrispondenze che si possono trovare tanto più dettagliata è la ricostruzione finale.

La condizione ottimale che si cerca di ottenere è posizionare le telecamere in modo da riprendere la scena da punti di vista e orientamenti diversi tra loro ma allo stesso tempo osservare gli stessi punti fisici in più immagini. Ad esempio una baseline più larga permette di osservare la scena da punti di vista più angolati, riducendo l'errore di triangolazione, ma allungando i tempi di ricostruzione dato che la ricerca dei punti da accoppiare deve essere eseguita su una porzione maggiore dell'immagine.

Durante questo lavoro sono state considerate quattro diverse configurazioni di posizionamento in cui variano sia il numero di pose telecamera sia la loro posizione relativa. Quattro configurazioni permettono di provare il sistema incrementando il numero di telecamere dal minimo possibile per ottenere la stereoscopia (due), arrivando fino a quattro. Il limite di quattro è stato scelto poiché in ambito industriale raramente si supera quel numero di telecamere su un solo sistema. Dato che la configurazione a quattro posizioni è quella che restituisce i risultati migliori si è deciso di analizzare più a fondo quel tipo di setup modificando la distanza relativa tra le telecamere.

Le posizioni all'interno di ogni configurazione sono state scelte tenendo conto dei vincoli relativi alle dimensioni del robot e della cella di lavoro. In particolare la lunghezza dei giunti limita il volume nel quale si può movimentare la telecamera. La telecamera è sempre posizionata alla massima altezza possibile in modo da poter inquadrare la massima estensione di area di lavoro possibile. Tutte le posizioni riportate sono riferite alla terna posizionata nella base del robot.

Per comodità, dato che nei paragrafi successivi si fa riferimento a queste posizioni, si specificano le quattro configurazioni con un codice identificativo.

- **Configurazione A:**

La prima e anche più semplice delle configurazioni è basata su due sole posizioni di ripresa. Le telecamere sono disposte al di sopra della scena, con una baseline (distanza tra i centri ottici) di 300 mm.

	Camera 1	Camera 2
X [mm]	-40	267
Y [mm]	200	-200
Z [mm]	475	472

Tabella 6.4: Posizionamento delle telecamere in configurazione A.

- **Configurazione B**

Nella seconda configurazione testata si hanno tre posizioni di ripresa, disposte a triangolo equilatero. Ogni lato di questo triangolo misura 350 mm.

	Camera 1	Camera 2	Camera 3
X [mm]	-40	234	235
Y [mm]	200	-100	-327
Z [mm]	475	471	470

Tabella 6.5: Posizionamento delle telecamere in configurazione B.

- **Configurazione C e configurazione D**

Le ultime due configurazioni sono composte da 4 posizioni di ripresa, disposte a quadrato al di sopra della scena. La differenza sta nella distanza relativa tra due punti di ripresa differenti: in un caso è di 200 mm nell'altro di 400 mm.

	Camera 1	Camera 2	Camera 3	Camera 4
X [mm]	0	0	200	200
Y [mm]	-150	-250	-250	-150
Z [mm]	478	476	476	478

Tabella 6.6: Posizionamento delle telecamere in configurazione C.

	Camera 1	Camera 2	Camera 3	Camera 4
X [mm]	-110	-140	305	301
Y [mm]	-313	-100	-302	-100
Z [mm]	410	410	410	403

Tabella 6.7: *Posizionamento delle telecamere in configurazione D.*

Stabilite le configurazioni precedentemente descritte è possibile verificarne le prestazioni di ricostruzione tridimensionale mediante alcune prove. Per mostrare quanto descritto nel paragrafo 6.3.2 sulla necessità di illuminare la scena con un pattern di punti, due prove principali sono riportate nel seguito: la prima utilizzando solamente la luce ambiente, la seconda con l'illuminazione attiva a infrarossi.

Ogni ricostruzione restituisce una nuvola di punti più o meno fitta a partire dalla quale l'algoritmo di pattern matching ricava la posizione e l'orientamento dei singoli pezzi. Per questa verifica si considera sia una scena ideale in cui sia presente un solo pezzo da riconoscere, sia una condizione reale di utilizzo in cui i pezzi sono disposti casualmente nella cassetta. In entrambe le condizioni si valutano i risultati del pattern matching. La bontà delle pose utilizzate può essere valutata attraverso la percentuale di punti del modello del pezzo che sono stati trovati nella scena. Questo valore è restituito dal parametro score in output dalla funzione di pattern matching. Un valore di score pari a 1 indica che il 100% dei punti del modello è stato individuato all'interno della nuvola di punti.

- **ILLUMINAZIONE AMBIENTALE:**

- Configurazione A:

	Pezzo singolo	Cassetta completa
Numero punti	649	12737
Tempo ricostruzione [s]	0.181	0.6
Pezzo riconosciuto da pattern matching	no	no
score	/	/

Tabella 6.8: Risultati della ricostruzione con telecamere in configurazione A e illuminatori spenti.

Avendo solamente due immagini la configurazione A permette la ricostruzione più rapida rispetto alle altre. Tuttavia il numero di punti ricostruiti è basso rispetto alle altre configurazioni e si perdono gran parte delle informazioni sullo spessore dei pezzi, in particolare al crescere dello spessore stesso. Infatti è possibile che i bordi laterali del pezzo siano visibili solamente in una delle due immagini, mentre nell'altra risultano ostruiti dal pezzo stesso. Il pattern matching non restituisce alcun risultato, né nella situazione ideale con un solo pezzo, né con la cassetta.

- Configurazione B:

	Pezzo singolo	Cassetta completa
Numero punti	1730	20318
Tempo ricostruzione [s]	0.83	3.5
Pezzo riconosciuto da pattern matching	si	no
score	0.8	/

Tabella 6.9: Risultati della ricostruzione con telecamere in configurazione B e illuminatori spenti.

Questa si è rivelata essere la configurazione più lenta nella ricostruzione. Tuttavia la maggior parte dei problemi della configurazione a due camere vengono risolti, in particolare quello riguardante lo spessore del pezzo. L'algoritmo di pattern matching individua il pezzo solamente nel caso ideale di pezzo singolo.

➤ Configurazione C e D:

	Pezzo singolo	Cassetta completa
Numero punti	1041	19161
Tempo ricostruzione [s]	0.93	2.9
Pezzo riconosciuto da pattern matching	si	no
score	0.8	/

Tabella 6.10: Risultati della ricostruzione con telecamere in configurazione C e illuminatori spenti.

	Pezzo singolo	Cassetta completa
Numero punti	1380	26316
Tempo ricostruzione [s]	0.35	1.6
Pezzo riconosciuto da pattern matching	si	no
score	0.84	/

Tabella 6.11: Risultati della ricostruzione con telecamere in configurazione D e illuminatori spenti.

Le configurazioni a quattro telecamere sono un compromesso tra velocità di ricostruzione e dettaglio della stessa. La principale differenza tra una configurazione stretta (C) e larga (D) sta nel tempo di elaborazione e nella capacità di ricostruire lo spessore dei pezzi. La configurazione con baseline più larga (400 mm) permette di osservare meglio i lati verticali degli oggetti dato che l'angolo

di incidenza tra la telecamera e la verticale è maggiore. Questo consente di ricostruire più facilmente i bordi esterni di oggetti particolarmente alti. Questo vantaggio si paga con un aumento dei tempi di ricostruzione. L'aumento è determinato dal fatto che, avendo baseline più ampia lo stesso punto fisico è rappresentato nelle diverse immagini in zone distanti tra loro. L'algoritmo di ricerca dei punti quindi impiega più tempo a cercare le corrispondenze in un'area più estesa all'interno delle immagini.

Comunque in entrambi i casi il pattern matching restituisce sempre la posa corretta del pezzo nel caso ideale mentre non si ottiene alcun risultato con la cassetta completa.

Come risulta dalle prove effettuate, in nessuno dei casi testati l'algoritmo di pattern matching riesce a individuare i pezzi disposti casualmente nella cassetta, mentre nelle configurazioni a più di due camere la ricostruzione di un singolo pezzo è sufficientemente dettagliata per poterlo riconoscere. Queste prove hanno portato all'utilizzo di un'illuminazione attiva come descritto nel paragrafo 6.3.2.

- **ILLUMINAZIONE ATTIVA**

Vengono qui ripetute le prove descritte sopra utilizzando in aggiunta alla componente di luce ambientale anche un sistema di illuminazione attiva.

Per ogni configurazione sono stati verificati i tempi di ricostruzione, il numero di punti ricostruiti, la capacità del pattern matching di riconoscere i pezzi, sia isolati che nella cassetta, e lo score risultante. Inoltre vengono presentate le differenze in percentuale del numero di punti ricostruiti e del tempo di ricostruzione rispetto alle senza illuminatore infrarossi.

	Configurazione A		Configurazione B	
	Pezzo singolo	Cassetta	Pezzo singolo	Cassetta
Numero punti	859	14365	2347	31747
Variazione numero di punti %	+32.3%	+12.78%	+35.6%	+56.2%
Tempo ricostruzione [s]	0.3	1.15	1.52	6.67
Variazione tempo ricostruzione %	+65%	+43.7%	+83.1%	+62.7%
Pezzo riconosciuto da pattern matching	si	si	si	si
Score	0.98	0.88	0.99	0.97

Tabella 6.12: Risultati della prova di ricostruzione con illuminatore acceso riferiti alle configurazioni A e B

	Configurazione C		Configurazione D	
	Pezzo singolo	Cassetta	Pezzo singolo	Cassetta
Numero punti	2387	25250	2028	36419
Variazione numero di punti %	+33.2%	+15.9%	+46.9%	+38.3%
Tempo ricostruzione [s]	1.91	4.5	0.64	2.79
Variazione tempo ricostruzione %	+105.3%	+40.6%	+82.8%	+46.8%
Pezzo riconosciuto da pattern matching	si	si	si	si
Score	0.99	0.94	0.98	0.99

Tabella 6.13: Risultati della prova di ricostruzione con illuminatore acceso riferiti alle configurazioni C e D

Dai risultati illustrati in Tabella 6. e Tabella 6.1 si nota come, in seguito all'aggiunta di texture sulla superficie tramite l'illuminazione precedentemente progettata, i pezzi vengono riconosciuti anche quando sparsi casualmente nella cassetta. Questo è dovuto alla maggiore densità di punti ricostruiti sulla scena. Rispetto alle prove con illuminatore spento si registra un aumento medio del numero di punti pari al 30%. Il maggior numero di punti permette alla procedura di pattern matching di individuare correttamente i pezzi nella cassetta.

Questo aumento dei punti ricostruiti implica però una crescita significativa dei tempi di ricostruzione; nelle quattro configurazioni si registra un aumento medio del 65%.

6.3.4 Validazione del sistema stereoscopico

Una volta realizzato il sistema di ricostruzione tridimensionale si è passati alla sua qualificazione metrologica. In primo luogo si esegue un'analisi sull'incertezza attesa che si ottiene variando il posizionamento relativo delle telecamere. In questo modo si vuole verificare che le configurazioni di posizionamento scelte nel paragrafo 6.3.3 siano caratterizzate da un'incertezza compatibile con le specifiche del sistema.

Nel processo di triangolazione si vanno a commettere degli errori sistematici dovuti al fatto che i raggi che collegano i centri delle telecamere con il punto da misurare sono sghembi e il punto che minimizza la loro distanza viene calcolato matematicamente.

Questa incertezza dipende generalmente da tre fattori:

- La distanza relativa tra i diversi punti di vista, o baseline
- L'angolo di inclinazione delle telecamere, espresso rispetto alla retta che congiunge le due camere.
- La lunghezza focale.

Baseline [mm]	d
Orientamento telecamera rispetto alla baseline [°]	β
Lunghezza focale [mm]	f
Dimensione pixel [μm]	p

Tabella 6.14: sigle dei parametri utilizzati per il calcolo.

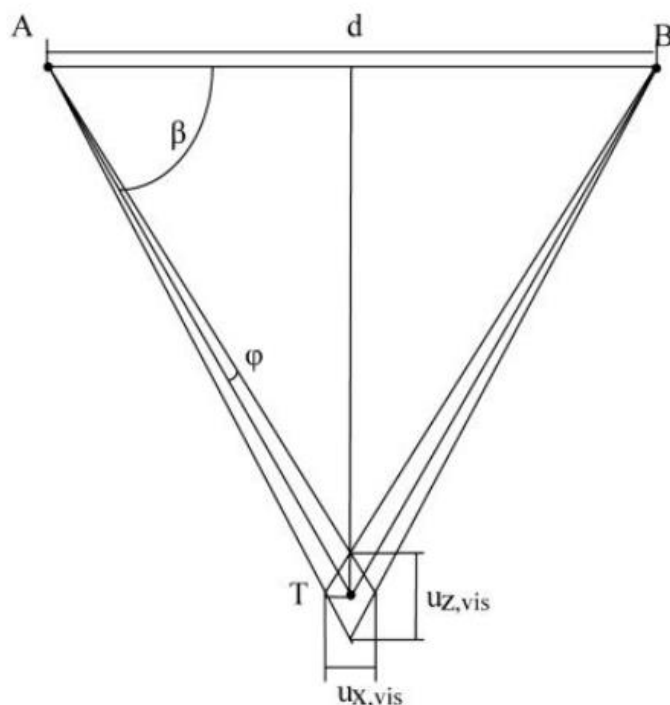


Figura 6-57: Rappresentazione dell'incertezza di triangolazione

Le incertezze lungo le tre direzioni principali u_x , u_y e u_z si possono calcolare tramite le relazioni:

$$u_x = u_y = \frac{d}{2\sqrt{3}} \frac{(\text{tg}(\beta) - \text{tg}(\beta - \varphi))}{\text{tg}(\beta - \varphi)}$$

$$u_z = \frac{d}{2\sqrt{3}} (tg(\beta) - tg(\beta - \varphi))$$

Dove:

$$\varphi = arctg\left(\frac{0.1 \cdot p}{f}\right)$$

Si vuole valutare l'incertezza attesa ottenuta variando questi parametri. I valori inseriti sono stati scelti per essere congruenti con quelli effettivamente utilizzati nelle prove sperimentali.

Per quanto riguarda la baseline si valuta un range a partire dai 200 mm fino a 600 mm, che comprende tutte e quattro le configurazioni sviluppate. L'angolo di inclinazione varia da un minimo di 60° per le configurazioni con baseline maggiore a uno di 80° per quelle con distanze tra telecamere più ridotte. La focale si considera costante, dato che, a partire dalle specifiche dimensioni della cella di lavoro, è stata scelta per poter inquadrare al meglio la scena e quindi non ha senso variarla.

	$\beta = 60^\circ$	$\beta = 70^\circ$	$\beta = 80^\circ$
d= 600 mm	0.008 mm	0.01 mm	0.02 mm
d= 400 mm	0.01 mm	0.02 mm	0.04 mm
d= 200 mm	0.02 mm	0.03 mm	0.06 mm

Tabella 6.25: Incertezza attesa di triangolazione lungo le direzioni x e y al variare della baseline e dell'angolo di incidenza.

	$\beta = 60^\circ$	$\beta = 70^\circ$	$\beta = 80^\circ$
d= 600 mm	0.01 mm	0.03 mm	0.02 mm
d= 400 mm	0.02 mm	0.06 mm	0.04 mm
d= 200 mm	0.11 mm	0.2 mm	0.35 mm

Tabella 6.16: Incertezza attesa di triangolazione lungo la direzione z al variare della baseline e dell'angolo di incidenza.

I risultati mostrano come in tutte le configurazioni analizzate l'incertezza teorica lungo x e y rimane inferiore al decimo di millimetro, mentre quella lungo z cresce considerevolmente al diminuire della baseline.

Dopo aver discusso l'incertezza teorica di triangolazione si effettua una campagna di prove sperimentali con l'obiettivo di analizzare le prestazioni dello strumento di misura sviluppato. L'analisi è divisa in tre parti nelle quali si vanno a esaminare i risultati ottenuti nelle tre sottoinsiemi fondamentali del sistema: la taratura del sistema ottico, la ricostruzione stereoscopica e la registrazione visione-robot.

- **VALIDAZIONE DEL PROCESSO DI TARATURA:**

Sono state eseguite due tipologie di prove per valutare la correttezza del processo di calibrazione e ricostruzione.

Il primo set di prove eseguite consiste nella ricostruzione di un insieme di punti di cui si conosce con elevata accuratezza la distanza relativa. Così facendo è possibile verificare lo scostamento tra le distanze dei punti reali e le distanze calcolate con la ricostruzione. Si è scelto di utilizzare come oggetto di test un calibratore, composto da 49 blob disposti a una distanza relativa di 12.5 mm uno dall'altro, posizionato in diverse pose. Il software possiede alcune funzioni che permettono di riconoscere in un'immagine questo tipo di calibratore e di restituire direttamente la posizione dei blob nello spazio. Sono state quindi acquisite le immagini del calibratore nelle 10 pose differenti e per ognuna sono state ricavate le coordinate tridimensionali di tutti i blob. Si sono quindi calcolate le distanze di ogni blob dai rimanenti e si ricava, per ogni posa del calibratore, un vettore contenente tutte le distanze misurate. Dalla differenza in modulo tra il vettore delle distanze reali e quello delle distanze misurate si trova un vettore degli scostamenti tra le distanze per ogni posa. Viene poi trovata la media tra gli elementi di ognuno dei vettori degli scostamenti. Così facendo si ha un valore di scostamento medio per ogni posa.

Questa prova è stata eseguita per tutte le 4 configurazioni di posizionamento telecamere precedentemente descritte.

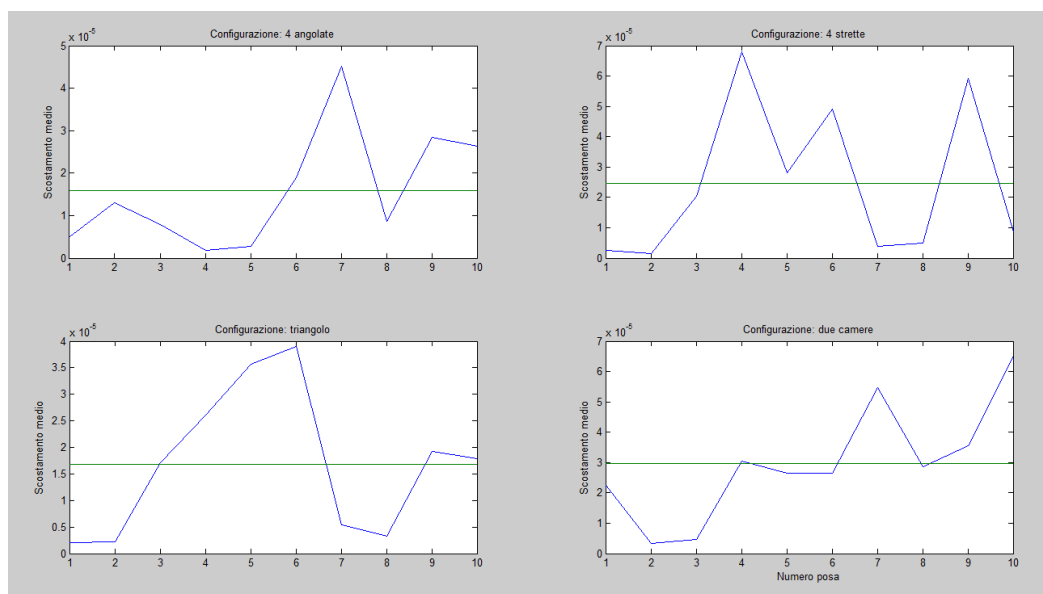


Figura 6-58: Andamento dello scostamento medio nelle quattro configurazioni. In blu: scostamento medio in ogni posa, verde: media dello scostamento nelle 10 prove.

Come si può vedere dai grafici in figura 6.60, lo scostamento medio è simile in tutte le quattro configurazioni e si attesta intorno ai 0.02 mm. Dato che in questo caso la posizione del centro dei blob viene calcolata tramite algoritmi di blob detection come media di tutti i pixel che compongono il blob stesso, la posizione restituita in coordinate immagine è più accurata rispetto a quella che si ricava da edge detection. E' lecito quindi aspettarsi un'incertezza inferiore a quella ottenibile da prove di ricostruzioni dei pezzi in condizioni reali in cui le coordinate immagine sono ricavate tramite edge detection, che non effettua alcuna media tra i valori. In particolare i risultati ottenuti in questa prova sono comparabili e validano quelli ricavati nell'analisi dell'incertezza attesa, dato che quel calcolo non tiene conto di eventuali errori nelle coordinate immagine dei punti.

La seconda prova eseguita per la validazione della calibrazione prevede la stima dell'incertezza restituita dallo strumento di misura nella ricostruzione di punti complanari. Si effettua la ricostruzione stereoscopica di una superficie perfettamente planare, nel caso specifico un calibratore e a partire dai punti ricavati si calcola tramite un algoritmo di *Principal Components Analysis* [30] l'equazione

del piano interpolante. E' possibile quindi misurare la distanza di ogni punto della ricostruzione dal piano così ricavato e ricavarne uno scostamento medio.

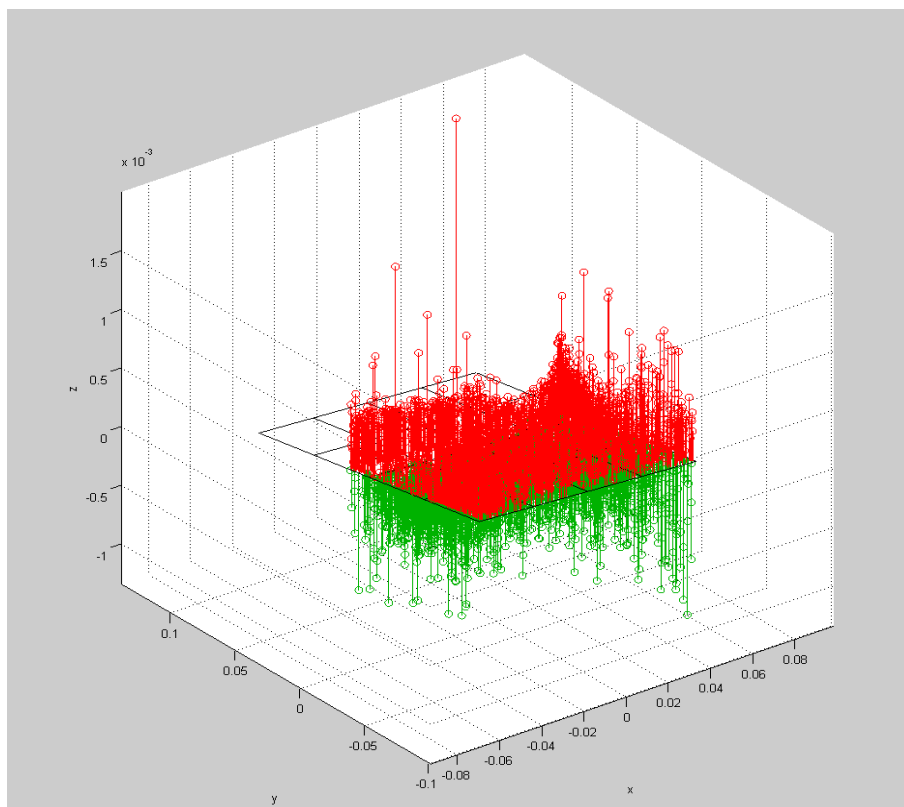


Figura 6-59: rappresentazione grafica delle distanze di ogni punto ricostruito dal piano interpolante

La stessa prova è stata effettuata per tutte le quattro configurazioni precedentemente definite. I risultati ottenuti sono mostrati in tabella 6.17:

	Scostamento medio [mm]
Configurazione A	0.2
Configurazione B	0.23
Configurazione C	0.16
Configurazione D	0.18

Tabella 6.17: scostamento medio dei punti ricostruiti rispetto al piano interpolante.

Lo scostamento medio si attesta intorno a 0.2 mm; quindi tutte e quattro le configurazioni si dimostrano ugualmente accurate nel ricostruire superfici planari.

- VALIDAZIONE DEL PROCESSO DI RICOSTRUZIONE

Si passa ora a verificare l'incertezza introdotta dalla ricostruzione stereoscopica. Le prove consistono nel ricostruire la nuvola di punti tridimensionale di oggetti di dimensioni note e misurarne le dimensioni in termini di distanza punto-punto. Dato che le nuvole di punti sulle quali si devono effettuare le misure sono molto fitte e la misura che si ottiene dipende fortemente dai punti che si scelgono per effettuare la misura stessa, ogni valore presentato è la media di più misurazioni effettuate sullo stesso misurando.

Si è scelto di utilizzare due tipi di riferimento per le misure: inizialmente si sono effettuate le prove misurando dei blocchetti di riscontro (o blocchetto Johnson) di quattro diverse lunghezze, dopodiché si sono misurati gli oggetti reali utilizzati nel bin picking. I blocchetti di riscontro sono dei parallelepipedi in acciaio lavorati in modo da ottenere due facce contrapposte perfettamente parallele e distanziate tra loro di una quota precisa, detta spessore nominale. A seconda degli errori massimi di spessore nominale i blocchetti vengono classificati in diverse classi di precisione, che definiscono i limiti di accuratezza garantiti. Per le prove effettuate sono stati utilizzati i blocchetti di spessore nominale pari a 100, 10, 2 e 1 millimetro. In questo modo si verifica l'accuratezza della ricostruzione in un range di grandezze che è compatibile con le dimensioni dei pezzi utilizzati. I risultati delle prove effettuate con i blocchetti di riscontro sono presentati in tabella 6.18:

	Configurazione D	Configurazione C	Configurazione B	Configurazione A
Lunghezza [mm]	Lunghezza [mm]	Lunghezza [mm]	Lunghezza [mm]	Lunghezza [mm]
100	100.18	100.09	100.15	100.11
10	10.21	10.11	10.14	10.26
2	2.09	2.12	2.25	2.1
1	1.15	1.2	1.28	1.24

Tabella 6.18: Risultati delle misure effettuate sulle ricostruzioni dei blocchetti di riscontro

Successivamente si sono ripetute le misure utilizzando come oggetto di riferimento uno dei pezzi utilizzati nel corso del lavoro per il binpicking, così da verificare l'accuratezza in condizioni di utilizzo reali. Le grandezze misurate in queste prove sono rappresentate in figura 6.62:

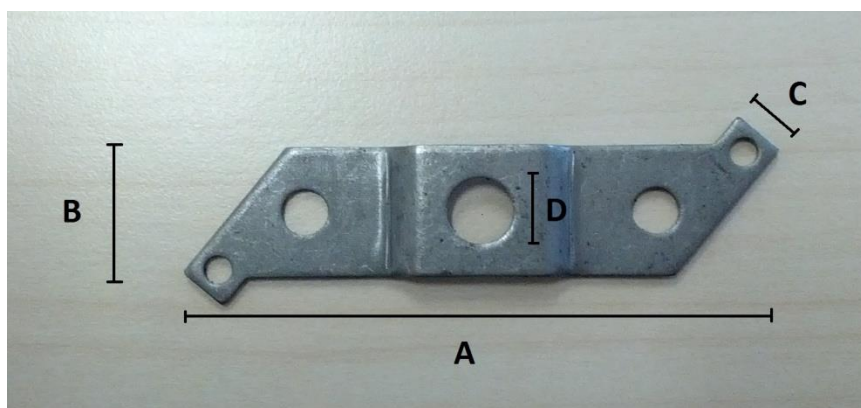


Figura 6-60: Dimensioni misurate.

	Sigla	Dimensioni [mm]
Lunghezza	A	79
Larghezza corpo principale	B	17
Larghezza estremità	C	7
Diametro foro principale	D	8.3
Spessore	E	1.4

Tabella 6.19: Dimensioni delle parti misurate.

Le prove sono effettuate sia con illuminazione accesa, sia spenta, per verificare se e quanto il pattern proiettato sui pezzi influisce sulla qualità della ricostruzione.

	Configurazione D	Configurazione C	Configurazione B	Configurazione A
A [mm]	79.14	78.92	79.13	79.3
B [mm]	17.07	17.21	16.85	17.2
C [mm]	7.1	6.87	7.32	7.2
D [mm]	8.42	8.14	8.23	8.42
E [mm]	1.57	1.52	1.57	1.57

Tabella 6.20: misure ottenute dalla misura del pezzo con illuminatori accesi.

	Configurazione D	Configurazione C	Configurazione B	Configurazione A
A [mm]	79.0	79.21	79.28	79.19
B [mm]	17.14	17.19	17.23	17.18
C [mm]	7.14	7.27	7.27	7.16
D [mm]	8.39	8.27	8.34	8.22
E [mm]	1.53	1.49	1.39	1.61

Tabella 6.21: misure ottenute dalla misura del pezzo con illuminatori spenti.

In conclusione si nota come su tutte le prove effettuate l'incertezza si attesta intorno a 0.2 mm. Questo è un buon risultato perché dimostra che l'accuratezza delle dimensioni degli oggetti ricostruiti nelle nuvole di punti non dipende significativamente dalla configurazione di posizionamento delle telecamere, né dalle dimensioni dell'oggetto stesso. Si può concludere che, come previsto, l'incapacità dell'algoritmo di pattern matching di individuare i pezzi senza pattern di luce proiettato è dovuto solamente al minor numero di punti ricostruiti e non da problemi nella ricostruzione delle dimensioni dei pezzi, dato che anche senza illuminazione non viene introdotta nessuna incertezza addizionale sulle misure.

- VALIDAZIONE DELLA REGISTRAZIONE TRA SISTEMA DI VISIONE E ROBOT:

Come ultima prova si vuole verificare la corretta registrazione del sistema di visione rispetto al robot. Si fissa un pezzo alla pinza robot e lo si deposita in modalità di movimento manuale in un punto qualsiasi dello spazio di lavoro. La posizione del pezzo è nota con accuratezza, corrisponde infatti alla posizione della pinza robot. Si ricostruisce quindi la scena tridimensionale e vi si applica l'algoritmo di pattern matching. Avendo imposto al robot e al sistema di visione lo stesso sistema di riferimento, la posizione restituita dal pattern matching deve corrispondere con quella della pinza in fase di posizionamento. L'obiettivo è misurare lo scostamento tra la posa reale del pezzo e quella ricavata da pattern matching.

Tuttavia misurando in questo modo la posizione del pezzo si introduce un errore sistematico sull'asse Z, dovuto al particolare posizionamento degli assi nel modello

CAD del pezzo. Il pattern matching restituisce la posizione dell'origine del sistema di riferimento che è stato posizionato in fase di creazione del CAD.

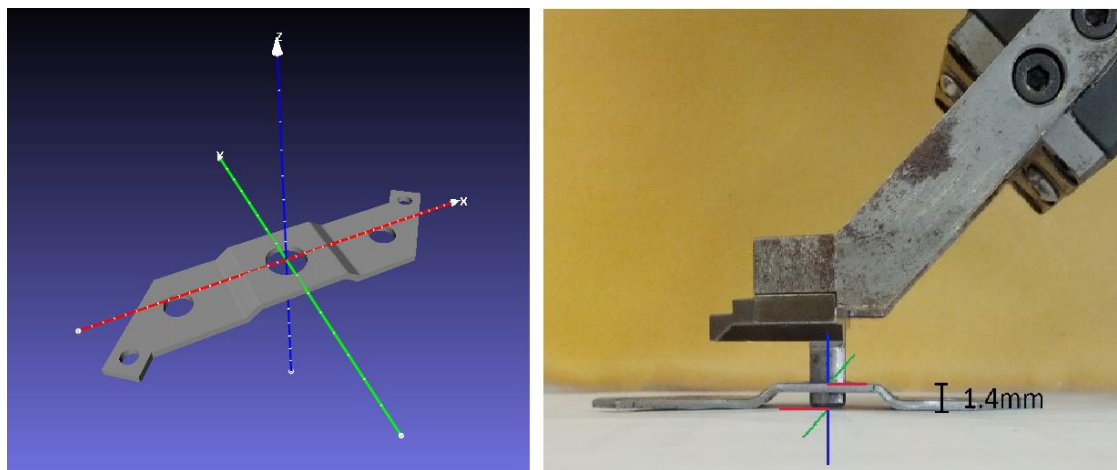


Figura 6-61: *Errore sistematico dovuto al posizionamento della pinza.*

Come mostrato in figura 6.63 la terna nel modello è posizionata al di sopra del pezzo, mentre per quanto riguarda la pinza la sua posizione è misurata usando come riferimento la punta della pinza stessa. Per poter afferrare correttamente il pezzo, la pinza deve scendere nel foro di una quota pari almeno allo spessore del pezzo. Ci si aspetta quindi uno scostamento lungo Z compatibile con lo spessore del pezzo, pari a 1.5 mm. Come per le altre prove effettuate con pattern matching, viene riportato anche il valore di score, ovvero la percentuale di punti del modello che sono stati ritrovati nella scena.

La prima prova è stata effettuata con illuminazione a infrarosso accesa e con condizioni di illuminazione esterna ottimali. Lo stesso pezzo è stato ricostruito a partire dalle immagini riprese nelle quattro configurazioni precedentemente illustrate.

	Posizione misurata	Posizione stimata dall'algoritmo di pattern matching			
		Configurazione D	Configurazione C	Configurazione B	Configurazione A
X [mm]	102.05	101.79	101.94	101.65	102.15
Y [mm]	-269.1	-268.93	-268.97	-269.25	-268.69
Z [mm]	-9.73	-8.41	-8.42	-8.43	-8.36
Rx [gradi]	0	359.5	359.88	359.7	359.46
Ry [gradi]	0	359.77	359.85	359.74	359.65
Rz [gradi]	180	181.67	181.35	0.93	0.91

Tabella 6.22: coordinate spaziali e orientamento della posa stimata da pattern matching, prova effettuata con illuminatore acceso.

	Configurazione D	Configurazione C	Configurazione B	Configurazione A
Scostamento lungo X [mm]	-0.26	-0.11	-0.4	0.1
Scostamento lungo Y [mm]	0.17	0.13	-0.15	0.41
Scostamento corretto lungo Z [mm]	-0.18	-0.19	-0.2	-0.13
Score	0.98	0.99	0.97	0.96

Tabella 6.23: scostamenti tra la posa misurata tramite robot e quella ricavata da pattern matching, nella prova con illuminatore acceso.

I risultati ottenuti mostrano che in caso di illuminazione ottimale l'algoritmo di pattern matching individua i pezzi con un'incertezza comparabile in tutte e 4 le configurazioni. Con illuminatore acceso e con un pezzo singolo presente sulla scena, si conferma che il pezzo viene riconosciuto con un valore di score comparabile in tutte le configurazioni. Si può inoltre confermare quanto osservato precedentemente riguardo lo scostamento lungo l'asse Z, infatti tutte le misure presentano un offset di circa 1.5 mm.

Si ripete ora la stessa prova senza l'illuminazione a infrarossi. In questo modo i pezzi sono privi di texture e verranno ricostruiti solamente i bordi. Tuttavia essendo presente sulla scena un solo pezzo e senza ulteriori elementi di disturbo, ci si aspetta che l'algoritmo ritrovi comunque la posa correttamente, ma con un valore di score molto inferiore.

	Posizione misurata	Posizione stimata dall'algoritmo di pattern matching			
		Configurazione D	Configurazione C	Configurazione B	Configurazione A
X [mm]	85.41	85.45	85.5	85.12	/
Y [mm]	-236.77	-236.17	-236.06	-235.96	/
Z [mm]	-0.4	1.6	1.23	0.99	/
Rx [gradi]	0	359.3	359.43	359.8	/
Ry [gradi]	0	359.23	359.5	359.62	/
Rz [gradi]	180	1.32	181.2	2.04	/

Tabella 6.24: coordinate spaziali e orientamento della posa stimata da pattern matching, prova effettuata con illuminatore spento.

	Configurazione D	Configurazione C	Configurazione B	Configurazione A
Scostamento lungo X [mm]	0.04	0.09	-0.29	/
Scostamento lungo Y [mm]	-0.6	-0.71	-0.81	/
Scostamento lungo Z [mm]	2	1.63	1.39	/
Scostamento lungo Z corretto [mm]	0.5	0.13	-0.11	/
Score	0.76	0.78	0.7	/

Tabella 6.25: scostamenti tra la posa misurata tramite robot e quella ricavata da pattern matching, nella prova con illuminatore spento.

Dalla tabella 6.24 si nota in primo luogo che l'algoritmo di pattern matching non ha trovato il pezzo nella configurazione a due telecamere mentre nelle altre configurazioni il valore di score è sceso sotto al 80%. Inoltre lo score peggiora di quasi il 10% passando da una configurazione a quattro telecamere a quelle da 3 e 2. Tuttavia gli scostamenti sono nello stesso ordine di grandezza di quelli ricavati nella prova con illuminazione accesa. Questo significa che il pattern matching utilizzato in questo lavoro consente, in presenza di pochi pezzi sulla scena, di restituire una posizione corretta del pezzo anche con valori di score bassi, intorno al 70%. Nella ricostruzione con sole due camere il numero di punti è troppo basso per poter riconoscere l'oggetto nella scena.

Le pose così ricavate, se passate al robot, consentono a quest'ultimo di afferrare correttamente i pezzi, come mostrato nella figura 6.64.



Figura 6-62: Presa del pezzo

In conclusione si può affermare che l'incertezza totale che si accumula sul sistema, data dall'unione dell'incertezza del sistema ottico, del pattern matching e della registrazione visione-robot risulta minore della tolleranza necessaria alla corretta presa del pezzo da parte del robot. Il sistema sviluppato risulta quindi funzionante e permette di effettuare operazioni di bin picking.

6.3.5 Conclusioni e confronto con la tecnica a triangolazione laser

La prova di diverse tipologie di illuminatori ha dimostrato l'importanza e la necessità di una corretta progettazione di un sistema di illuminazione che possa incrementare la risoluzione delle ricostruzioni; in particolare la proiezione di un pattern adeguato migliora notevolmente l'accuratezza delle procedure di riconoscimento basate su algoritmi di pattern matching 3D.

La configurazione del sistema di misura ha diversi effetti sull'incertezza, valutati grazie al confronto di risultati sia teorici che pratici ottenuti da varie prove. Per il sistema stereoscopico si sono valutate le prestazioni ottenute con diverse configurazioni di posizionamento delle telecamere. Ne è risultato che un buon compromesso tra velocità di ricostruzione e risoluzione della nuvola di punti risultante si ottiene utilizzando quattro posizioni di ripresa e mantenendo la baseline ridotta. Dal punto di vista dell'accuratezza di misura non si sono rilevati particolari scostamenti variando il numero di telecamere da due a quattro. Bisogna tener presente però che nel prototipo realizzato il volume di misura era limitato e non è stato possibile provare a variare considerevolmente la distanza delle telecamere dalla scena, dato che le dimensioni del robot che movimentava la telecamera non lo permettevano. Ci si aspetta che aumentando significativamente la distanza e il volume di lavoro le differenze tra le ricostruzioni ottenute variando il numero di telecamere risultino molto più marcate.

Si è constatato che l'incertezza ottenuta dipende molto dalla qualità dei componenti utilizzati, in particolare della telecamera. Infatti disponendo di una telecamera con un sensore con risoluzione maggiore si otterrebbe una riduzione dell'incertezza di misura, poiché si avrebbe una diminuzione del rapporto pixel/mm, a parità di distanza del punto da misurare e di ottica utilizzata.

Il corretto funzionamento di un sistema di bin picking è il risultato di misurazioni sufficientemente accurate, ma anche di una corretta procedura di registrazione dello strumento di visione con il manipolatore robotico. È stata posta particolare attenzione alla fase di registrazione, che è risultata adeguata.

Sia le misure ottenute, sia i risultati della registrazione visione-robot sono stati valutati mediante prove sperimentali specifiche, che hanno permesso di verificare che i sistemi realizzati garantiscano un'incertezza sufficientemente ridotta da permettere la presa dei pezzi tramite il robot.

L'incertezza totale, infatti, è risultata dell'ordine del decimo di millimetro, valore inferiore alla tolleranza massima di presa, questo permette di soddisfare gli obiettivi preposti, quindi una corretta presa dei pezzi ammassati nella cassetta.

Grazie allo sviluppo all'interno del VBLab di due strumenti che utilizzano tecniche di misura differenti per risolvere il problema di bin picking, è stato possibile svolgere un confronto su pregi e difetti di ognuno. Si ricorda che oltre alla tecnica basata su stereoscopia qui descritta, il problema era già stato affrontato dallo staff del VBLab mediante una tecnica a triangolazione laser. I principali fattori che differenziano i due sistemi sono: il tempo richiesto per effettuare la misura della scena e la risoluzione ottenibile in termini di numero di punti.

La stereoscopia è una tecnica di misura densa; le informazioni necessarie alla ricostruzione vengono ottenute istantaneamente (a meno del tempo di esposizione necessario per l'acquisizione di immagini in base alle condizioni di luminosità della scena) mediante l'acquisizione di più immagini della scena. La velocità nella ricostruzione è limitata solamente dai tempi computazionali degli algoritmi numerici. Utilizzando algoritmi ottimizzati da questo punto di vista, a parità di potenza computazionale del pc utilizzato, è possibile ottenere tempi di ricostruzione ridotti. Nel sistema sviluppato in questo lavoro la ricostruzione dell'intera scena richiede 1.5 s.

Tuttavia la ricostruzione stereoscopica si è rivelata essere fortemente dipendente dalle condizioni di illuminazione della scena. E' necessario mantenere quanto più possibile costante la quantità di luce che incide sulla scena; anche piccoli cambiamenti possono influire sulla qualità della ricostruzione. E' stata proposta una soluzione che risolve in parte queste problematiche ma nello stesso tempo rallenta notevolmente le operazioni di ricostruzione.

Il numero di punti ricostruiti dipende fortemente dalla quantità di dettagli che la scena presenta. Se gli oggetti non presentano texture notevoli sulla loro superficie, il numero di punti ricostruito si rivela insufficiente per poter garantire il riconoscimento dei pezzi. E' stato quindi necessario utilizzare un proiettore che permettesse di aumentare artificialmente la texture presente sulla scena. Attraverso una serie di prove si sono riconosciuti come elementi fondamentali nella scelta del proiettore la sua potenza, che deve essere sufficientemente elevata da permettere di rendere visibili i pattern proiettati con qualsiasi condizione di luce, e la forma e dimensione dei pattern stessi. Il pattern proiettato deve essere fitto ma non confuso e deve essere uniformemente proiettato su tutta la scena.

In conclusione un sistema di bin picking basato su stereoscopia si adatta in modo efficace a situazioni in cui sia possibile controllare in modo molto accurato l'illuminazione dell'area di lavoro e in cui i pezzi da afferrare presentino caratteristiche visibili molto accentuate come colori differenti tra loro, scritte o texture particolari.

Lo scanner laser che utilizza una tecnica di misura non densa richiede un certo tempo per la registrazione delle informazioni. Sono necessarie molte immagini nelle quali la lama laser viene movimentata in diversi punti della scena per poter avere una ricostruzione 3D che sia adatta ad un sistema di bin picking. In questo caso la velocità complessiva è limitata sia dal tempo richiesto per la scansione della scena sia dagli algoritmi di calcolo.

L'estrazione dei punti del profilo a partire dall'immagine a pieno sensore è la procedura più onerosa (in termini computazionali); i tempi necessari a questa operazione possono essere ridotti utilizzando una telecamera intelligente (*smart camera*) che grazie ad un'elettronica programmabile incorporata svolge l'operazione di estrazione delle coordinate pixel del profilo laser a valle di ogni singola acquisizione, così da non dover memorizzare le immagini ed analizzarle con il pc. D'altra parte, tale scelta incrementa il costo totale dello strumento di misura.

Per quanto riguarda l'acquisizione, quando la risoluzione desiderata è elevata, nel caso in cui si devono misurare oggetti di forme complesse, è necessario regolare opportunamente due parametri: velocità di scansione e frequenza di acquisizione. La velocità di scansione può essere ridotta, così da poter acquisire molte più immagini. La frequenza di acquisizione può essere invece aumentata, per poter disporre di un numero maggiore di immagini; in questo modo si dilatano i tempi della ricostruzione.

In applicazioni in cui i tempi ciclo sono stringenti la tecnica più efficiente è la stereoscopia. Inoltre quando i pezzi presentano zone di occlusione e sottosquadri, l'utilizzo di uno scanner laser può dar luogo ad alcune criticità, perché la lama laser può non essere in grado di illuminare queste zone, dato che il proiettore laser è vincolato a ruotare attorno a un asse in posizione fissa.

Quando però sono necessari molti punti distribuiti uniformemente sulle superfici degli oggetti, la scelta di utilizzare lo scanner laser consente un riconoscimento robusto, anche in diverse condizioni di illuminazione ambiente. Per la ricostruzione di oggetti uniformi il laser consente di avere un numero elevato di punti lungo tutta la superficie.

7. Conclusioni

Lo scopo di questa tesi è la qualificazione metrologica di sistemi stereoscopici e la loro applicazione in tre diversi ambiti. L'evoluzione della tecnologia di ripresa e di elaborazione delle immagini consente di raggiungere prestazioni sempre migliori in termini di accuratezza delle ricostruzioni tridimensionali mediante stereoscopia. Sicuramente l'accuratezza di un sistema stereoscopico dipende dall'hardware utilizzato, ma a parità di hardware può dipendere da diversi altri fattori: illuminazione ambientale, posizione del sistema rispetto al misurando, ma soprattutto dalla qualità della procedura di taratura. Si è voluto dunque approfondire lo studio della procedura di taratura del sistema stereoscopico, che ha un effetto significativo sui risultati della misura ma può essere completamente controllata con opportuni accorgimenti. Essa è composta dalla calibrazione delle singole telecamere, processo nella quale vengono stimati i parametri intrinseci delle telecamere, e dalla registrazione delle stesse, vale a dire la stima delle pose relative. Il processo che più influisce sulla buona riuscita della taratura di un sistema stereoscopico è la calibrazione delle singole telecamere. Si è dunque voluto dedicare una parte del lavoro allo studio di alcuni dei principali algoritmi di calibrazione esistenti in letteratura e in particolare al confronto tra la tecnica di Zhang [2] ottimizzata mediante algoritmo di Levenberg-Marquardt (Zhang-LM) [8] e la tecnica di Heikkilä [1]. La prima è una tecnica molto diffusa e utilizzata in letteratura; essa è in uso presso il VBLab del Politecnico di Milano, laboratorio in cui è stato svolto l'intero lavoro, da diversi anni ed è stata implementata, validata e

consolidata con diversi lavori. La seconda, invece, è meno nota ma è stata implementata nelle librerie per la visione artificiale chiamate HALCON e prodotte da MVTec, che sono in uso da poco presso il laboratorio. Data la comodità e la facilità di utilizzo delle procedure di calibrazione implementate in HALCON, una parte del lavoro è stato dedicato al confronto metrologico tra la tecnica di Heikkilä e la tecnica di Zhang-LM. Dalle prove effettuate si è potuto riscontrare che la prima raggiunge prestazioni paragonabili alla seconda e può dunque essere utilizzata come valida alternativa nelle procedure di calibrazione di telecamere. La tecnica implementata nelle librerie HALCON presenta dei vincoli applicativi stringenti come ad esempio la necessità di conoscere con buona accuratezza la dimensione fisica dei pixel della telecamera. Questo vincolo però è compensato da molti vantaggi a livello di comodità e semplicità della procedura per le acquisizioni delle immagini di calibrazione e dalla possibilità di avere funzioni per la taratura di sistemi stereoscopici a più telecamere. Per questi motivi durante il lavoro è stata utilizzata la tecnica di Heikkilä per affrontare il problema della calibrazione delle telecamere che compongono il sistema stereoscopico.

Un sistema a tre telecamere è stato realizzato in modo da effettuare una serie di prove atte a qualificare metrologicamente la tecnica utilizzata. Per prima cosa si è studiato a fondo la procedura di taratura del sistema stereoscopico. In particolare la procedura di calibrazione utilizzata effettua una minimizzazione dell'errore di retroproiezione, vale a dire la distanza tra i punti di un calibratore individuati nelle immagini (i dati di partenza) e i punti dello stesso calibratore trasformati sulla base dei parametri di telecamera trovati, quindi retroproiettati. Questo parametro è comunemente utilizzato per definire la bontà della calibrazione e si è voluto studiare il legame presente tra esso e l'accuratezza di misura tridimensionale. Si vuole quindi capire la dipendenza dell'errore 3D effettuato in fase di misura dall'errore 2D restituito in fase di calibrazione. Per fare ciò sono stati studiati due diversi errori 3D. Il primo è dato dagli scostamenti delle distanze dei marker di un calibratore planare dal proprio valore reale. Il secondo invece è definito come lo scostamento tra distanze dei marker di un calibratore posto in una prima posa dai marker dello

stesso calibratore posto in una seconda posa e il loro valore reale. È stato possibile ottenere il valore reale delle distanze tra due pose differenti in quanto il calibratore inquadrato è stato montato in pinza a un robot antropomorfo in grado di generare spostamenti con accuratezza trascurabile rispetto a quella del sistema di misura.

Il risultati mostrano che l'errore di retroproiezione è una buona stima dell'accuratezza della calibrazione, in quanto entrambi gli errori 3D calcolati crescono al crescere dell'errore 2D. Esiste tuttavia un intervallo di valori dell'errore di retroproiezione (<0.22 pixel) in cui l'errore 3D rimane contenuto e sostanzialmente costante al variare dell'errore 2D. Questo significa che, per valori molto bassi di errore di retroproiezione, non è detto che un valore poco più alto di un altro implichi una peggior accuratezza di calibrazione.

È stato inoltre verificato che per effettuare correttamente la taratura del sistema stereoscopico è conveniente utilizzare almeno 15 pose differenti. Sopra questo valore è possibile affermare che l'errore di retroproiezione restituito dalla procedura di calibrazione definisce la bontà della calibrazione e garantisce una buona accuratezza in fase di misura. Utilizzando meno di 15 immagini invece si rischia di incorrere in false buone calibrazioni, vale a dire calibrazioni che restituiscono un errore 2D molto basso ma non garantiscono buone accuratezze al sistema di misura. Questo avviene in particolare nei casi in cui si utilizzano poche pose del calibratore tutte molto simili tra loro; la procedura di calibrazione riesce a minimizzare l'errore 2D con un set di parametri validi solo per quelle particolari pose, ma con i quali il sistema stereoscopico non è in grado di effettuare misure tridimensionali con buone accuratezze (errore 3D alto).

Inoltre è stato mostrata l'importanza di calibrare in tutto lo spazio in cui si prevede di dover effettuare misure tridimensionali col sistema stereoscopico. Infatti, da una serie di prove, si è mostrato come l'accuratezza di misura peggiori se vengono effettuate misure fuori dallo spazio di calibrazione.

Utilizzando il metodo di Monte Carlo poi si è proceduto all'analisi della propagazione dell'incertezza durante la procedura di taratura del sistema. Partendo da alcune prove necessarie a stimare le incertezze della fase di acquisizione dei punti del calibratore, è stato poi possibile studiare le incertezze dei parametri di

calibrazione stimati. I risultati mostrano che tutti i parametri presentano distribuzioni gaussiane con deviazione standard molto contenute. Per poter poi comprendere gli effetti di tale incertezze sulla ricostruzione tridimensionale, vengono eseguite alcune analisi di sensitività della procedura di ricostruzione, studiando come la modifica del valore di alcuni dei parametri principali (con valori scelti all'interno della propria distribuzione), avesse effetti sulle coordinate 3D ricostruite dal sistema. Dalle prove si sono potute osservare variazioni nell'ordine di pochi decimo di millimetro, che definiscono dunque l'accuratezza che il sistema in esame può raggiungere. Dall'analisi di sensitività è stato inoltre mostrato che la coordinata affetta da una maggior incertezza è quella perpendicolare al piano del sensore.

Una volta effettuata la validazione metrologica del sistema stereoscopico esso è stato configurato in modo ottimale per essere utilizzato in tre diversi casi applicativi. Questi sono rappresentati da tre problemi differenti fra di loro, tutti risolti in maniera innovativa mediante la realizzazione di un sistema stereoscopico dedicato. In particolare le tre applicazioni sono: misure in galleria del vento, programmazione automatica di robot industriali e guida robot per la presa da cassone di parti alla rinfusa.

La prima nasce dalla necessità della Galleria del Vento del Politecnico di Milano di effettuare misure di spostamento in tre dimensioni senza introdurre effetti di carico sul misurando, in modo da non disturbare la vena fluida del vento. Il sistema stereoscopico realizzato garantisce scalabilità e flessibilità, caratteristiche indispensabili per questa applicazione. Per aumentare la robustezza del sistema di misura sono stati aggiunti degli illuminatori infrarossi e filtri passa banda. Con questa configurazione il sistema è in grado di misurare le posizioni tridimensionali di particolari marcatori catarifrangenti posti sul misurando. Il sistema è stato validato con alcune prove di ricostruzione di oggetti noti e applicato a numerose campagne sperimentali. Tra queste si sottolinea la ricostruzione tridimensionale della forma di una vela di un modello di imbarcazione sottoposto all'azione del vento (detta *flying shape*). Grazie all'estrazione di alcuni parametri geometrici

comunemente utilizzati dai velai per la fabbricazione delle vele, è stato possibile confrontare la *flying shape* con la *design shape*, vale a dire la ricostruzione tridimensionale teorica che viene realizzata dal velaio mediante alcuni software di simulazione. Studiando poi l'andamento delle forze scambiate a terra dall'imbarcazione, misurate con un bilancia a sei assi, con la *flying shape* è stato possibile analizzare le forme realmente assunte dalla vela durante i punti di massima spinta. In alcune campagne sperimentali per la misura di spostamenti di strutture civili, i dati acquisiti dal sistema stereoscopico sono stati confrontati con alcuni dati acquisiti da accelerometri posti sulla struttura. Le misure dei due strumenti sono risultate compatibili, validando ulteriormente il sistema stereoscopico realizzato.

Il secondo ambito applicativo nasce da un progetto di ricerca effettuato con la società ABB Robotica Italia, nell'ambito del quale è stato realizzato un dimostratore robotizzato in grado di riprodurre fedelmente le traiettorie manuali eseguite da un operatore che afferra un opportuno tracciatore completo di marcatori sferici catarifrangenti. In particolare è stato realizzato un dimostratore in grado di effettuare la programmazione offline di un robot per la decorazione di supporti ceramici. Il metodo proposto è interamente automatizzato e ha il vantaggio di essere flessibile e scalabile, consentendo aree di misura differenti. Sono stati effettuate alcune prove per definire il grado di accuratezza del metodo proposto. I principali contributi all'accuratezza vengono introdotti dalle procedure di registrazione dei sistemi di acquisizione e di attuazione robot. Il grado di accuratezza raggiunto permette di produrre risultati conformi agli standard delle decorazioni manuali. In particolare mediante un opportuno trattamento offline dei dati di acquisizione è stato possibile raggiungere uno scarto tipo del risultato del processo di estrazione della posizione della punta del pennello pari a 0.1 mm. In ultimi, il sistema è stato interamente verificato mediante una campagna sperimentale effettuata con ABB e un tecnico esperto in decorazioni di ceramiche.

Infine, nell'ultima applicazione, è stato realizzato un sistema di presa di oggetti posizionati in maniera casuale in un cassone, mediante un robot antropomorfo. Per

tale applicazione si è notata fin da subito la necessità di una corretta progettazione di un sistema di illuminazione che potesse incrementare la risoluzione delle ricostruzioni. Dopo numerose prove il sistema è stato integrato con un dispositivo in grado di proiettare un pattern adeguato sulla cassetta contenente i pezzi da prelevare, in modo da aumentare i punti ricostruiti dal sistema e migliorare notevolmente l'accuratezza delle procedure di riconoscimento dei pezzi basate su algoritmi di pattern matching 3D. È stata inoltre studiata la miglior configurazione del sistema stereoscopico, vale a dire la miglior disposizione delle telecamere per aumentare l'accuratezza di misura. Per fare ciò si è scelto di utilizzare una sola telecamera movimentata mediante il braccio robotico, in modo da non avere limiti sul numero di punti di osservazione della scena utilizzabili e poter valutare le prestazioni ottenute con diverse configurazioni di posizionamento delle telecamere. Da queste prove è risultato che un buon compromesso tra velocità di ricostruzione e risoluzione della nuvola di punti risultante si ottiene utilizzando quattro posizioni di ripresa e mantenendo la baseline ridotta. Sia le misure ottenute, sia i risultati della registrazione visione-robot sono stati valutati mediante prove sperimentali, che hanno permesso di verificare che il sistema realizzato garantisca un'incertezza sufficientemente ridotta da permettere la presa dei pezzi tramite il robot. L'incertezza totale è risultata dell'ordine del decimo di millimetro, valore inferiore alla tolleranza massima di presa (dell'ordine del millimetro); questo permette di soddisfare gli obiettivi preposti, quindi una corretta presa dei pezzi ammucchiati nella cassetta. La problematica era già nota all'interno del VBLab del Politecnico di Milano in quanto era già stata affrontata utilizzando un triangolatore laser come sistema di scansione. Da un confronto con il sistema già realizzato si è concluso che un sistema di bin picking basato su stereoscopia risulta essere decisamente più veloce ma poco robusto alle variazioni di luce della scena, a differenza del sistema a triangolazione laser. Il sistema stereoscopico rappresenta dunque la soluzione migliore nel caso in cui il tempo ciclo richiesto debba essere molto basso e vi sia la possibilità di controllare le condizioni di luce della scena. La tecnica a lama laser invece risulta vincente dove non si hanno vincoli temporali troppo stringenti ma la scena è soggetta alle variazioni di luminosità tipiche degli ambienti industriali.

Bibliografia

- [1] J. Heikkilä. *Geometric Camera Calibration Using Circular Control Points*. IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 22, NO. 10, October 2000
- [2] Zhengyou Zhang. *A Flexible New Technique for Camera Calibration*. Microsoft Research, One Microsoft Way, Redmond, USA. December 1998.
- [3] Y. Ma, S. Soatto, J. Kosecka, S. S. Sastry. *An Innovation to 3-D Vision*. Springer, 2004.
- [4] E. Trucco, A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
- [5] Roger Y. Tsai. *A Versatile Camera Calibration Techniaue for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses*. IEEE, Journal of Robotics and Automation, Vol. RA-3, August 1987.
- [6] Wang Qi, Fu Li, Liu Zhenzhong. *Review on Camera Calibration*. Shenyang Institute of Aeronautical Engineering, Shenyang 110136, China.
- [7] Hou Zhenjie, Zhao Jing, Gu Liguu, Lv Guoling. *Automatic Calibration Method Based on Traditional Camera Calibration Approach*. Inner Mongolia Agriculture University, Hohhot, China.
- [8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Second edition. Cambridge University Press, 2003.
- [9] D.C. Brown. *Decentering Distorsion of Lenses*. Photometric Engineering, pages 444-462, Vol. 32, No. 3, 1966.
- [10] G. Mainetti, *Calibrazione di telecamere per sistemi di visione stereoscopica: confronto tra algoritmi genetici e tecniche tradizionali*, Politecnico di Milano, Italy, 2011.

- [11] Liu Wan-Yu, Xie Kai. *A Camera Calibration Method based on Neural Network optimized by Genetic Algorithm*. Heilongjiang Scientific Research Foundation and the Scientific Research Foundation for the Returned Overseas Chinese Scholars.
- [12] Andrea Fusiello. *Visione Computazionale*. Università degli Studi di Verona, Italy, 2008.
- [13] Matteo Pedraglio. *Analisi e qualificazione metrologica di calibrazioni di telecamere matriciali per applicazioni di Bin-Picking*. Politecnico di Milano, Italy, 2009.
- [14] R.C. Gonzales, R.E. Woods, S.L. Eddins. *Digital Image Processing Using Matlab*. Pearson Education, 2004.
- [15] C. Harris and M. Stephens. *A Combined Corner and Edge Detector*. Proceedings of the 4th Alvey Vision Conference. pp. 147–151. 1988.
- [16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery. *Numerical Recipes. Third edition*. Cambridge University Press, 2007.
- [17] E. Trucco, A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
- [18] Pradit Mittrapiyanuruk, *A Memo on How to Use the Levenberg-Marquardt Algorithm for Refining Camera Calibration Parameters*, Robot Vision Laboratory, Purdue University, West Lafayette, IN, USA.
- [19] Manolis I.A. Lourakis and Antonis A. Argyros. *Is Levenberg-Marquardt the Most Efficient Optimization Algorithm for Implementing Bundle Adjustment?*, Institute of Computer Science, Foundation for Research and Technology, Crete, GREECE.
- [20] JCGM 102:2011: Supplement 2 to the “Guide to the expression of uncertainty in measurement” – *Propagation of distributions using Monte Carlo method*
- [21] Nicholas Metropolis and S. Ulam. *The monte carlo method*. Journal of the American Statistical Association, 44(247):335–341, 1949.

-
- [22] Cuthbert C. Hurd. *A Note on Early Monte Carlo Computations and Scientific Meetings*. *Annals of the History of Computing*, 7(2):141–155, April-June 1985.
- [23] F. J. Massey Jr. , *The Kolmogorov-Smirnov Test for Goodness of Fit*, *Journal of the American Statistical Association*, Vol. 46, Issue 253, 1951
- [24] V. DePuy, V. W. Berger, Y. Zhou, *Wilcoxon-Mann-Whitney Test*, *Encyclopedia of Statistics in Behavioral Science*, Oct. 2005
- [25] F. Fossati, G. Mainetti, M. Malandra, R. Sala, P. Schito, A. Vandone, *Offwind Sail Flying Shapes Detection*, 5th High Performance Yacht Design Conference Auckland, 10-12 March, 2015
- [26] F. Fossati, S. Muggiasca, I. Bayati, C. Bertorello, *Lecco Innovation Hub Sailing Yacht Lab Project – A Sailing Research Infrastructure*, 3rd International Conference on Innovation in High Performance Sailing Yachts, Lorient, France, 2013.
- [27] F. Fossati, I. Bayati, F. Orlandini, S. Muggiasca, A. Vandone, G. Mainetti, R. Sala, C. Bertorello, E. Begovic, *A Novel Scale Laboratory For Yacht Engineering Research*, *Ocean Engineering*, 2015
- [28] Alessandro Perego. *Qualificazione metrologica di uno scanner laser a tempo di volo per la ricostruzione tridimensionale di forme in ambiente non strutturato*. Politecnico di Milano, Italy, 2008.
- [29] Patent application n° PD2014A000249
- [30] Jolliffe I.T., *Principal Component analysis, Series: Springer Series in Statistics*, 2nd ed., Springer, NY, 2002.
- [31] Arun, Huang, Blostein. *Least-Squares fitting of Two 3-D Point Sets*. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. Vol. PAMI-9, No. 5, September 1987.
- [32] www.abb.com
- [33] www.mvtec.com/halcon/
- [34] <http://www.windtunnel.polimi.it/cirive/cirive.htm>