



UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
TESI DI LAUREA MAGISTRALE

RICOSTRUZIONE SHAPE FROM
SILHOUETTE: AUTOCALIBRAZIONE
PER APPLICAZIONI A RAGGI X

RELATORE: Prof. Luca Poletto

CORRELATORE: Dott. Emanuele Simioni

LAUREANDO: *Marcomattia Mocellin*

Padova, aprile 2012

*"So che la Terra è rotonda perché ho visto l'ombra sulla Luna,
e mi fido delle ombre."*

Magellano

Indice

1	Introduzione	1
2	Radiazione X	5
2.1	Storia	5
2.2	Sorgenti	8
2.2.1	Tubo di Crookes	8
2.2.2	Tubo di Coolidge	9
2.2.3	Sorgenti Moderne	10
2.3	Applicazioni	13
3	Premessa Teorica	15
3.1	Shape from Silhouette: Visual Hull	17
3.1.1	Euristica Generale	23
3.1.2	Euristica Applicata	23
3.2	Geometria Proiettiva	25
3.2.1	Pinhole Camera	25
3.2.2	Matrice Fondamentale	30
3.2.3	Matrici Proiettive	31
3.3	Autocalibrazione	34
3.3.1	Fast Chord-to-Point Distance Accumulation Corner Detector	34
3.3.2	Cernita delle features	40
3.3.3	Normalized 8-point algorithm	44
4	Fase Sperimentale	45
4.1	Strumenti	45
4.2	Target	50
4.2.1	Modello Matematico: <i>Calibration Box</i>	51

4.2.2	Modello Matematico: Procedura	53
4.3	Applicazioni Precedenti	59
5	Ricostruzione	62
5.1	Galeone	62
6	Conclusioni	67
A	Implementazione MATLAB	70
	Bibliografia	72

Capitolo 1

Introduzione

La ricostruzione 3D è uno dei principali argomenti di indagine della *computer vision*, la branca della scienza il cui scopo è costruire descrizioni esplicite e significative di oggetti a partire da immagini [1]. Fin dalla nascita dei primi computer furono infatti evidenti i vantaggi di macchine capaci di "vedere", o meglio di trarre informazioni da immagini, per esempio nella cartografia o in ambito industriale. In quest'ultimo, per esempio, la ricostruzione 3D di oggetti trova applicazione attraverso software CAD o per la diagnostica. Con il progredire di tecnologia e strumenti sono state sviluppate molteplici tecniche per ottenere il modello 3D di un oggetto, come, per citarne qualcuna, la stereografia, la tomografia o lo Shape-from-Silhouette. Quest'ultima, oggetto di questa tesi, ricava l'informazione relativa all'oggetto da una serie di immagini acquisite intorno ad esso o, più precisamente, dalla silhouette ricavabile da esse.

L'applicazione dello Shape-from-Silhouette è variata negli anni a causa dell'evoluzione della tecnologia disponibile: inizialmente molto diffusa, è diventata obsoleta con la nascita degli scanner 3D (avvenuta intorno agli anni '80), che hanno reso possibile un sistema di misura diretta più preciso e con meno limiti rispetto allo Shape-from-Silhouette, in grado di eseguire un'effettiva ricostruzione solo di oggetti convessi.

Nonostante tale restrizione, questa tecnica presenta ancora notevoli vantaggi in alcuni campi, quali ad esempio l'archeologia: potendo essere facilmente adattato all'utilizzo con radiazione X e non richiedendo elaborazioni particolarmente complesse o infrastrutture ingombranti, lo Shape-from-Silhouette si pone come alternativa pratica e relativamente leggera (in termini computazionali) a tecni-

che di ricostruzione come la tomografia, di certo più performanti ma che d'altro canto necessitano di risorse non indifferenti. Inoltre, considerando di utilizzare la radiazione visibile, lo Shape-from-Silhouette è ancora oggi una tecnica valida per elaborare immagini, ovvero fare *imaging*, in tempi brevi e a basso costo. Queste caratteristiche rendono lo Shape-from-Silhouette degno di indagine, in special modo in Italia, dove l'archeologia trova uno dei più vasti e ricchi territori di applicazione.

Obiettivo

Lo Shape-from-Silhouette è stata una delle prime tecniche utilizzate nella ricostruzione 3D e quindi le sue caratteristiche sono ben note: obiettivo di questa tesi è quello di sviluppare una nuova metodologia per l'autocalibrazione del sistema di imaging, cioè l'estrazione dei parametri geometrici del sistema (indispensabili ai fini della ricostruzione) direttamente dalle immagini 2D disponibili.

Ci si è posto tale obiettivo in un'ottica di lavoro sul campo per siti archeologici: maggiore è la quantità di dati ottenibili direttamente dalle acquisizioni, minori sono il tempo e lo spazio necessari per ottenere la ricostruzione 3D desiderata. In particolar modo si è voluto evitare di dover aggiungere all'oggetto considerato un qualche sistema di riferimento (come può essere una griglia di dimensioni note), in modo da ridurre al minimo le operazioni di supporto all'acquisizione. Tuttavia in questa sede non è stato possibile eliminare totalmente tali operazioni, non potendo misurare con sufficiente precisione tutti i parametri essenziali (come distanza focale e posizione del punto principale) del sistema prospettico.

Implementazione

Per ottenere la ricostruzione 3D con un grado di dettaglio accettabile, lo Shape-from-Silhouette necessita di molteplici acquisizioni, ognuna da un punto di vista differente. A tal scopo, in questa tesi, si è scelto di far ruotare l'oggetto e acquisirne di volta in volta la silhouette, in modo da ottenere una sequenza di immagini lungo tutta la sua circonferenza, con il vincolo che l'oggetto debba essere ruotato di un angolo costante per tutte le acquisizioni.

Una volta ottenute tali silhouette, l'informazione necessaria allo Shape-from-Silhouette è disponibile sotto forma di sequenze di punti (detti *features* o *ereditari*) che mettono in relazione un'immagine con la precedente e la successiva in

base alla rotazione dell'oggetto. Le features scelte a tale scopo sono gli spigoli, in quanto presentano alcuni vantaggi: innanzitutto è relativamente semplice rivellarli attraverso algoritmi di elaborazione (o segmentazione) delle immagini detti *corner detector*; inoltre mantengono caratteristiche simili tra un'immagine e la successiva, ragion per cui è possibile associarli (e quindi mettere in correlazione immagini successive) senza incorrere in grandi ambiguità nella rivelazione.

Le sequenze così trovate permettono di ricostruire l'angolo che separa due immagini direttamente dalle immagini stesse. Tale angolo è necessario allo Shape-from-Silhouette per ottenere una corretta ricostruzione 3D.

Riassumendo le fasi dell'algoritmo realizzato, esse si suddividono come segue:

1. Acquisizione delle silhouette tutto intorno all'oggetto;
2. Segmentazione delle suddette immagini;
3. Individuazione delle features;
4. Creazione delle sequenze di features;
5. Correlazione delle immagini tramite le sequenze;
6. Calcolo dell'angolo di rotazione da un'immagine alla seguente;
7. Generazione del modello 3D.

In letteratura non sono riportati altri studi su questa tematica con le metodologie in esame in questa tesi.

Panoramica della Tesi

Questa tesi si compone di cinque capitoli, divisi come segue: nel primo capitolo è presente un'introduzione generale degli argomenti trattati; nel secondo viene esposta una breve storia della radiazione X, utilizzata in questo esperimento; nel terzo si trova la trattazione teorica generale necessaria a comprendere le varie fasi dell'elaborazione; nel quarto è descritto l'esperimento vero e proprio, sia nell'aspetto tangibile che in quello teorico specifico; infine nel quinto e ultimo capitolo vengono presentati i risultati sperimentali su un galeone in argento e le conclusioni tratte dallo studio.

Codice Sviluppato

A causa della mole dello stesso, si è optato per allegare in questa sede non l'intero codice Matlab utilizzato, ma il solo elenco di script realizzati con annessa descrizione, reperibile nell'Appendice A.

Capitolo 2

Radiazione X

Dal momento della loro scoperta, i raggi X hanno profondamente caratterizzato la tecnologia umana, rivoluzionando alcuni campi della scienza e fornendo inestimabile supporto in altri. In questo capitolo si intende offrire una panoramica storica e applicativa di tale radiazione.

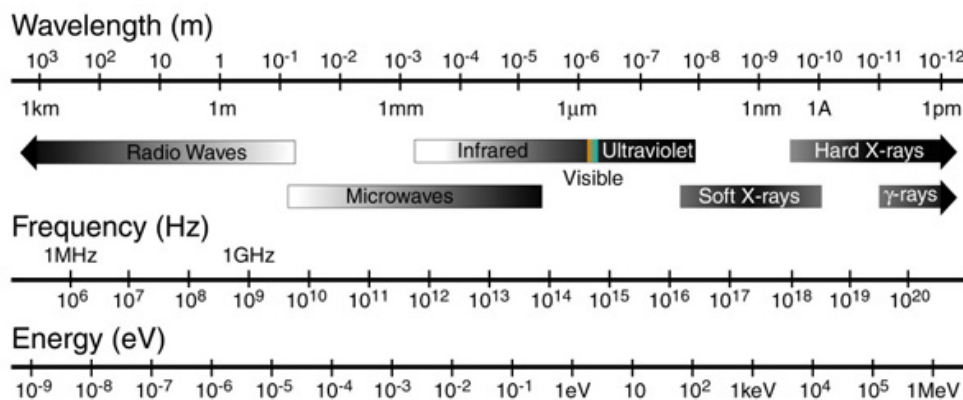


Figura 2.1: Spettro Elettromagnetico

2.1 Storia

La radiazione X deve il suo nome al proprio scopritore Wilhelm Conrad Röntgen. Nel 1895, anno della prima pubblicazione di un articolo su tali raggi, non si era ancora compresa la natura di queste radiazioni e quindi Röntgen li battezzò raggi-X [3]. Pur essendo conosciuti anche come raggi di Röntgen, la prima denominazione fu quella che prese piede e divenne ufficialmente il nome della regione dello spettro elettromagnetico compreso tra 0.01 e 10 nm.

Röntgen scoprì la radiazione quasi per caso, durante un esperimento coinvolgente i raggi catodici emessi da un tubo di Crookes e uno schermo fluorescente. Egli notò come lo schermo emanasse un bagliore pur essendo il tubo coperto da cartone, e quindi ipotizzò l'esistenza di una radiazione ancora sconosciuta capace di attraversare corpi e oggetti. Sempre a Röntgen si deve anche la nascita della radiografia e la prima lastra radiografica della storia, l'immagine delle ossa della mano di sua moglie [4].

Con quest'ultima scoperta, si aprì un nuovo orizzonte di applicazioni per i raggi X, dalla medicina alla spettroscopia, anche grazie alla spinta per lo sviluppo di nuove sorgenti. Infatti i tubi di Crookes erano inaffidabili e scostanti nell'emissione, perché necessitavano, per funzionare, di una certa quantità di gas al proprio interno, di solito aria, che permettesse lo scorrere di corrente; l'aria, con il tempo, veniva assorbita all'interno del tubo stesso e quindi alterava l'emissione della radiazione. Vennero ideate molteplici soluzioni a questo problema, ma fu solo nel 1904 che si trovò una valida alternativa ai tubi di Crookes: in questo anno John Ambrose Fleming realizzò il tubo a vuoto¹ che permetteva lo scorrere di elettroni attraverso l'emissione termoionica. Il tubo a vuoto è l'elemento fondamentale del tubo di Coolidge, strumento tutt'oggi alla base di molte sorgenti [2].

Una volta avuta a disposizione una sorgente affidabile per i raggi X, l'interesse attorno a questa² radiazione crebbe a dismisura, spaziando dalla medicina all'industria. Nel 1912 William Henry Bragg e suo figlio William Lawrence completarono i primi esperimenti nella scienza della spettroscopia X, formulando anche la celebre condizione di Bragg, $n\lambda = 2d \sin \theta$. A loro fecero seguito molti altri, ma tutti sfruttando le nuove possibilità introdotte dai raggi x.

Nel 1921 i fratelli de Broglie, Maurice e Louis, arrivarono a proporre la loro ipotesi di dualità onda/particella proprio grazie al risultato delle analisi fatte con e su i raggi X. Negli anni seguenti si videro progressi soprattutto nel campo della diagnostica medica, con lo sviluppo della TAC³, basata sul metodo di rappresentazione di una singola fetta del corpo umano (la *stratigrafia*) proposto dal radiologo italiano Alessandro Vallebona nel 1930. Questo portò allo sviluppo di nuove e più sicure tecniche radiografiche e del microscopio a raggi X, basato

¹Noto anche come valvola termoionica.

²E molte altre: di questi anni infatti sono le scoperte dei vari raggi α , β e γ .

³Tomografia Assiale Computerizzata; il primo tomografo computerizzato fu installato nel 1971.

anche sui lavori di focalizzazione ottica di Paul Kirkpatrick e Albert Baez. [3]

2.2 Sorgenti

La storia dei raggi X non può essere scissa dallo sviluppo delle sue sorgenti. In poco più di un secolo si passò infatti dall'inaffidabile seppur portatile tubo di Crookes all'attualmente unica sorgente a spettro continuo: il sincrotrone.

2.2.1 Tubo di Crookes

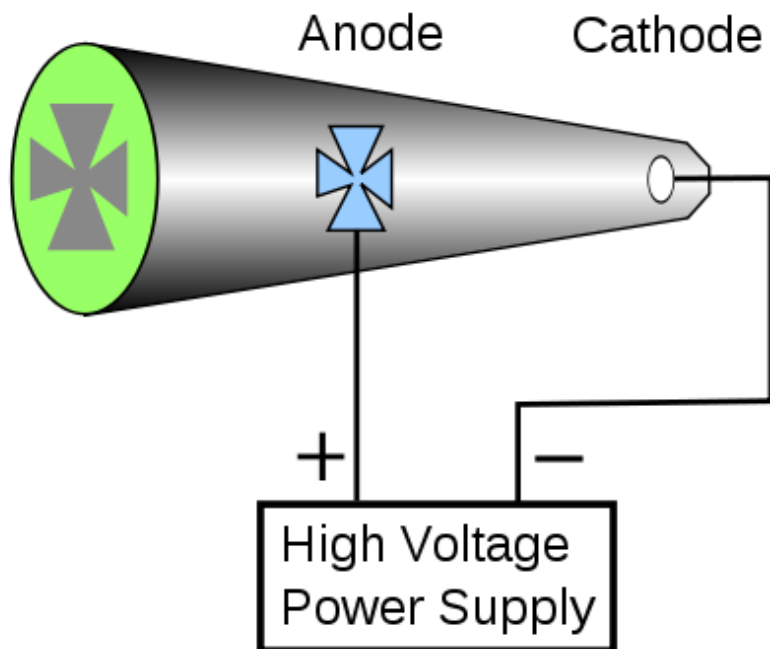


Figura 2.2: Schema di un Tubo di Crookes

Il tubo di Crookes è stata la prima sorgente di raggi X, pur non essendo stato ideato per questo scopo. Esso infatti era utilizzato come sorgente di raggi catodici, così chiamati perché emessi dal catodo. Tali raggi venivano emessi ionizzando l'aria presente all'interno del tubo stesso, creando così elettroni liberi. Da qui l'impossibilità, per tale tecnologia, di creare e mantenere il vuoto all'interno del tubo, risultando conseguentemente in scarsa affidabilità e consistenza nell'emissione.

Nonostante questi difetti, i tubi di Crookes raggiungevano potenze del tutto rispettabili (da qualche kW fino ad anche 100 kW) tuttavia, a causa dell'umidi-

tà presente al loro interno, non potevano garantire a lungo tali prestazioni. Per ovviare a questo intrinseco problema si pensò di modificarne la struttura, aggiungendo dei “softeners” che regolassero la pressione all’interno del tubo. Tale modifica è mostrata in figura 2.3, dove si nota un’altra modifica effettuata all’originale per ottenere raggi X: venne aggiunta una placca di platino (scelto per la sua alta temperatura di fusione) su cui i raggi catodici impattavano, generando così raggi X attraverso Bremsstrahlung oppure fluorescenza X¹.

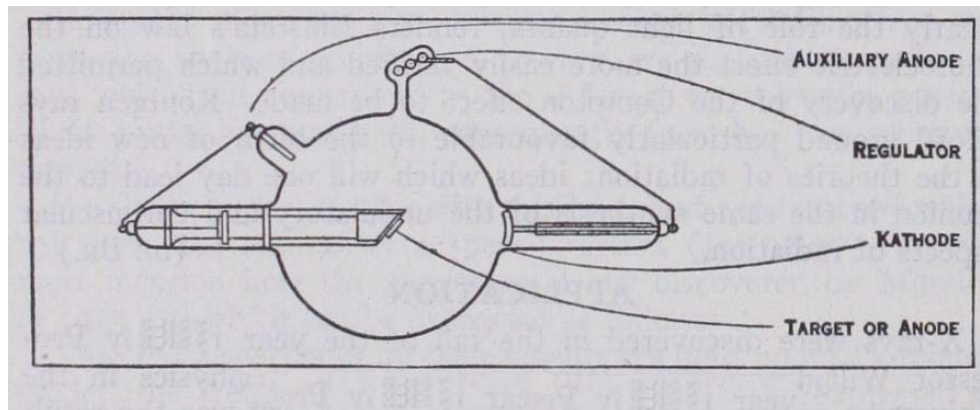


Figura 2.3: Tubo di Crookes con sistema “softener”, consistente in un cristallo di mica o un agente chimico che rilascia del gas proporzionalmente alla temperatura.

2.2.2 Tubo di Coolidge

I tubi di Crookes rimasero l’unica sorgente di radiazione X fino all’invenzione della valvola termoionica, che permette l’emissione di elettroni in vuoto. L’idea alla base del tubo di Coolidge è quella di sfruttare “l’evaporazione” di elettroni da un filamento incandescente per poi farli collidere con un bersaglio adeguato creando così i raggi X.

Unica attenzione da porre è quella di mantenere raffreddato il bersaglio, requisito facilmente soddisfatto grazie al design di base del tubo, come visibile in figura 2.4. Infatti Coolidge prevede la possibilità di raffreddare a liquido l’anodo del proprio dispositivo, possibilità che, unita al funzionamento a vuoto del tubo, garantiva così quella continuità di prestazioni che nei tubi di Crookes non era ot-

¹Per Bremsstrahlung si intende la radiazione emessa da particelle cariche quando queste vengono accelerate o subiscono urti, mentre la fluorescenza X è l’emissione di radiazione dovuta al decadimento di elettroni da livelli energetici più esterni verso livelli più interni.

tenibile. Questa maggior affidabilità però venne raggiunta con un compromesso sulla potenza massima: mediamente ci si assesta su un massimo di 20kW, contro i 100 raggiunti dai tubi di Crookes.

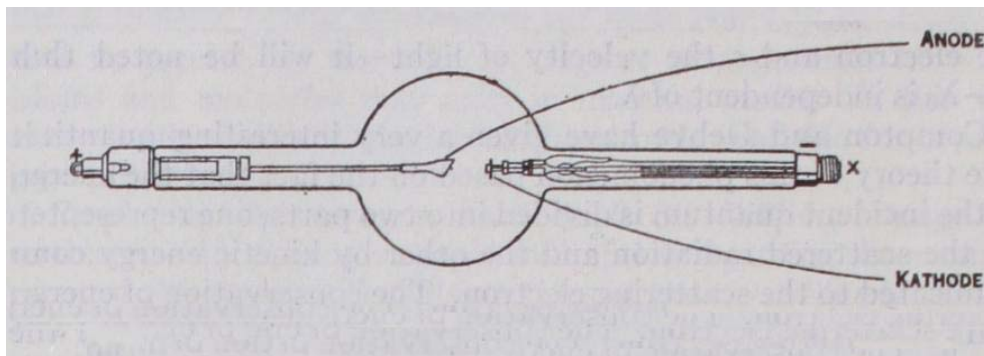


Figura 2.4: Tubo di Coolidge

Quello visto in figura 2.4 è una delle due tipologie base del tubo di Coolidge, quella a finestra laterale, l'altra è quella a finestra in coda. La differenza si trova essenzialmente nel tipo di bersaglio e nella direzione di emissione dei raggi X. Nei tubi a finestra laterale il bersaglio è relativamente spesso e l'emissione avviene, ovviamente, lateralmente. Nel tipo a finestra in coda, invece, il bersaglio solitamente dev'essere sottile, in quanto la generazione non avviene per riflessione ma per trasmissione.

2.2.3 Sorgenti Moderne

Ad oggi i tubi di Coolidge vengono largamente utilizzati, seppur con varie modifiche. Tra queste vi sono le sorgenti Microfocus, come la Hamamatsu L8601-01 utilizzata in questa tesi. Le sorgenti Microfocus vengono così dette perché il fascio di elettroni viene focalizzato sul bersaglio ottenendo grande coerenza spaziale.¹ Questa modifica progettuale impone una serie di modifiche visibili nello schema in figura 2.5, rendendo le sorgenti microfocus più complicate dal punto di vista costruttivo, ma allo stesso tempo stabili e compatte.

¹Questo a patto di limitare la potenza disponibile, in quanto grandi potenze danno luogo a densità di carica notevoli, che per effetti elettrostatici tendono ad allargare lo spot.

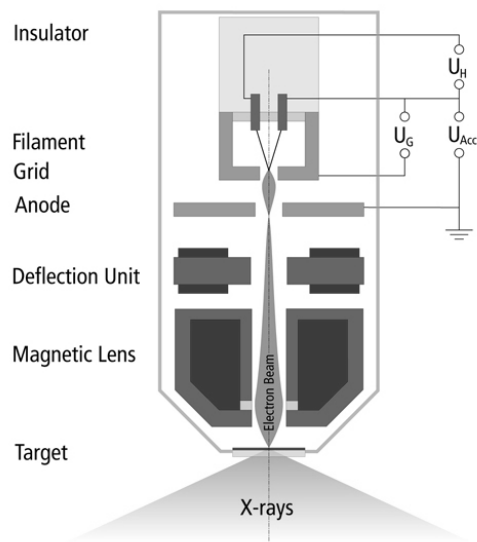


Figura 2.5: Schema di funzionamento di una sorgente Microfocus

Un'altra tipologia di tubo radiogeno è composta dai tubi ad anodo rotante, usati in medicina per le radiografie, che essenzialmente utilizzano un bersaglio più grande e messo in rotazione, per ottenere maggiore area di indagine a parità di radiazione emessa. Contemporaneamente si garantisce una maggiore vita operativa in quanto l'anodo viene irradiato solo in parte, riducendo lo stress cui viene sottoposto. Le sorgenti finora trattate sono accomunate dal fatto che i fenomeni alla base dell'emissione dei raggi X siano o emissione per Brehmsstrahlung oppure per fluorescenza X (altresì nota come emissione caratteristica). Entrambi i fenomeni si devono all'interazione di elettroni fortemente energizzati con gli atomi dell'anodo, nello specifico si parla di fluorescenza qualora l'interazione avvenga con elettroni appartenenti ad orbitali interni all'atomo¹; si parla di Brehmsstrahlung quando invece gli elettroni incidenti sfiorano il nucleo dell'atomo senza esserne catturati ma solo deflessi: questa deflessione solitamente porta ad emissione X.

Tuttavia esistono altre sorgenti X, sia naturali, come le Supernovae, i Quasar e i buchi neri, sia artificiali, come la radiazione di sincrotrone. Quest'ultima è ad oggi l'unica sorgente capace di fornire uno spettro X continuo, in quanto sfrutta l'emissione di radiazione da parte di una particella carica accelerata a velocità

¹Quando l'elettrone incidente ne eccita un altro appartenente ad un orbitale interno ne determina anche l'espulsione e quindi la liberazione di uno stato energetico profondo. Per questo motivo altri elettroni decadono dagli orbitali più esterni, decadimento accompagnato dall'emissione di fotoni X.

relativistiche [5]. Come visibile in figura 2.6, gli elementi di un sincrotrone sono composti dal generatore di elettroni accelerati, il LINAC¹, un anello di sincronizzazione, o linea di trasferimento, per iniettare gli elettroni dove possano essere utilizzati, cioè nello storage ring. Qui, oltre ai magneti necessari al mantenimento degli elettroni in traiettoria, si trovano le tre famiglie di sorgenti: i bending magnet, gli ondulatori e i wiggler. I primi sfruttano l'emissione di radiazione da parte di particelle cariche che si muovono in traiettoria circolare. Infatti queste particelle, nello specifico trattasi di elettroni, emettono radiazione a spettro estremamente largo (dagli infrarossi agli X duri) lungo un cono tangente alla loro traiettoria, che si restringe sempre più man mano che ci si avvicina a velocità relativistiche. Gli ondulatori permettono di selezionare una precisa lunghezza d'onda, in quanto consistono di una struttura magnetica periodica, ma con campo debole, che causa agli elettroni piccole oscillazioni angolari e quindi emissione in poche lunghezze d'onda.

I wiggler infine sono strutture magnetiche periodiche come gli ondulatori, ma a campo forte, per cui le oscillazioni e le loro armoniche degenerano fino a creare quasi uno spettro continuo come quello generato dai bending magnet, però centrato a energie più elevate. Come è possibile intuire, il sincrotrone permette di raggiungere energia, brillantezza e coerenza spaziale estremamente elevate, ma si tratta di uno strumento praticamente limitato alla sola ricerca, in quanto estremamente costoso, dispendioso in termini energetici, bisognoso di un vuoto estremamente spinto per funzionare correttamente e, non ultimo, di grandi dimensioni: i più compatti necessitano comunque di un'intera stanza dedicata.

¹LINear ACcelerator, in cui gli elettroni generati per effetto termoionico raggiungono velocità lineari prossime a quelle della luce.

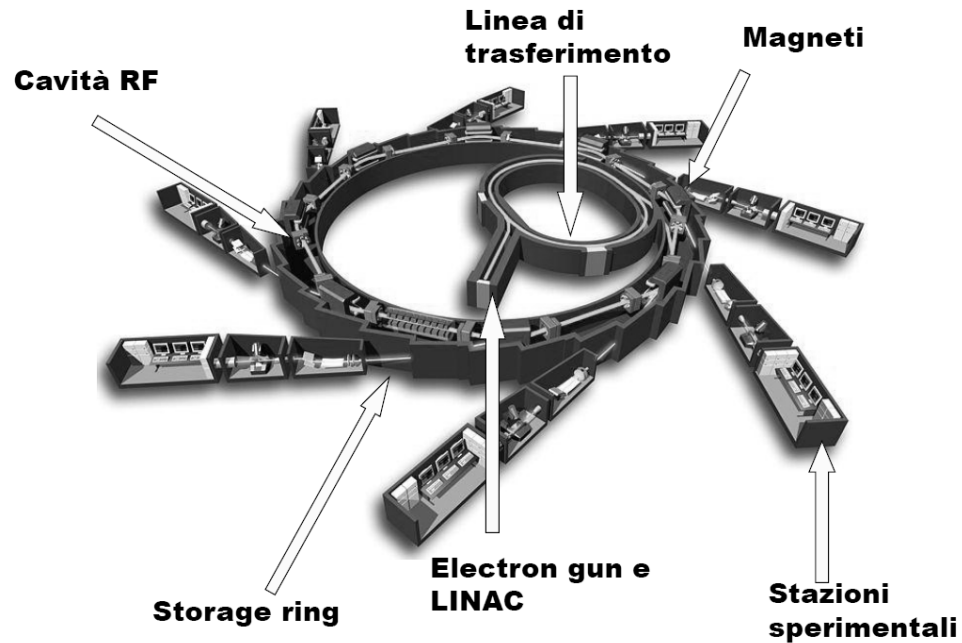


Figura 2.6: Elementi di un Sincrotrone

2.3 Applicazioni

Come già riportato, la prima applicazione dei raggi X è stata in ambito medico. Attraverso le radiografie divenne possibile esaminare la condizione delle ossa dei pazienti, migliorando di molto il trattamento e l'abilità diagnostica del tempo. La radiografia ricevette un'ulteriore seppur triste spinta grazie alla Grande Guerra, in cui la possibilità di vedere dove fosse fisicamente il danno (o il proiettile) migliorava di molto le probabilità di sopravvivenza dei soldati. Tuttavia l'utilizzo che se ne faceva allora è ben diverso da quello attuale, in quanto non si conoscevano ancora i pericoli delle radiazioni ionizzanti: la radiazione X viene usata ancora oggi per molteplici esami, ad esempio TAC o radiografie [6] [7], ma in condizioni rigidamente controllate e da parte di personale formato allo scopo.

Oggi l'utilizzo di raggi X non è limitato al solo campo medico, ma si riscontra in molteplici ambiti: la radiografia viene utilizzata comunemente come metodo di diagnostica industriale (per esempio per verificare saldature oppure che il contenuto rispetti una data specifica [8]); in ambito scientifico lo spettro X permette ulteriori analisi della materia (attraverso la cristallografia) o dello spazio (tramite progetti quali la missione CHANDRA) [9] o simili [10]; in ambito artistico-archeologico, infine, permette di indagare la storia di opere d'arte oppure

di verificare la presenza e la forma di manufatti nascosti alla radiazione visibile [11] [12].

Capitolo 3

Premessa Teorica

Tra le molteplici tecniche di ricostruzione 3D esistenti, lo Shape-from-Silhouette (d'ora in poi SfS) è stata una tra le prime ad essere investigata. Più precisamente, è una tecnica di ricostruzione tridimensionale a partire da immagini, solitamente fotografie, con un'unica premessa: le immagini devono essere acquisite in numero sufficiente attorno all'oggetto, ovvero ci deve essere un numero adeguato di punti di presa. Questa, come si vedrà, è la condizione necessaria per poter arrivare a ricostruire il target considerato, a prescindere dalla specifica tecnica con cui si ricostruirà il suo modello.

Nell'ambito di questa tesi, lo SfS viene applicato non su delle normali fotografie (cioè fatte con radiazione visibile) ma su delle immagini realizzate tramite raggi X: in questo caso lo schema ottico cambia leggermente e si perde l'informazione ottenibile da ombre e colori dell'oggetto, ma rimane l'informazione topologica. L'impiego della radiazione X permette l'estensione di questa tecnica a vari campi, quali medicina, controllo qualità e, come si vedrà nei prossimi capitoli, archeologia.

Si supponga, per esempio, di voler esaminare un qualche manufatto monolitico come una statuetta o un pendente non visibile a occhio nudo per via di ossidazione o residui non removibili. In questo caso la tecnica più affermata di ricostruzione 3D è la tomografia X. Questa tecnica consiste nel ricostruire oggetti tridimensionali sovrapponendo più "fette" ottenute con un'onda che attraversa completamente l'oggetto stesso. Poiché è la lunghezza d'onda¹ a determinare la risoluzione della ricostruzione, quella della tomografia X si trova nell'ordine di

¹Che nel caso dei raggi X è compresa tra 0.001 e 10 nanometri.

decine di nanometri lungo l'intera sezione dell'oggetto, sia esterno che interno. I requisiti necessari all'applicazione di tale tecnica sono tuttavia stringenti: è necessario un passo angolare di 1 o 0.5 gradi (cioè tipicamente è necessario acquisire 360 o 720 foto intorno all'oggetto) una sorgente tale da riuscire a trapassare l'oggetto¹ e un elaboratore dalle discrete capacità o, in alternativa, un quantitativo non trascurabile di tempo. Qui emergono i vantaggi di una tecnica quale lo SfS: il suo approccio infatti necessita di un passo angolare compreso tra 5 e 20 gradi, una sorgente a bassa energia² e un minor carico computazionale. Questo si traduce, rispetto alla tomografia, in una risoluzione decisamente inferiore, limitata a frazioni di millimetro, e a un effettivo limite nella ricostruzione: lo SfS infatti non può rilevare tutte le concavità dell'oggetto, ma solo quelle osservabili nel banco di immagini. D'altro canto un vantaggio non trascurabile è la maggior velocità di realizzazione (minuti contro ore) e una migliore portabilità e semplicità d'utilizzo. Queste caratteristiche (considerato anche che nel caso di manufatti monolitici l'informazione aggiuntiva sull'interno data dalla tomografia è di fatto superflua) rendono lo SfS particolarmente interessante e degno d'approfondimento.

Come accennato, esistono molte tecniche di ricostruzione SfS, che si differenziano in base alla strategia attraverso la quale il modello 3D finale viene ricostruito, ad esempio mesh adattive³, che stimano l'oggetto a partire da una precisione molto scarsa, per poi collassarvi ricreandone i dettagli ovvero le *features* [14], oppure isocurve, utilizzate in questa tesi, che rappresentano l'oggetto attraverso la superficie di una sua approssimazione, data dal *Visual Hull*.

¹Trattasi di requisito non banale, se si pensa che alcuni manufatti possono tranquillamente essere spessi qualche centimetro ed essere costituiti di argento od oro e che la tomografia lavora in trasmissione.

²Non si richiede trasmissione, ma è sufficiente la sola maschera dell'oggetto.

³Trattasi di superfici 3D caratterizzate da un insieme di vertici la cui funzione di connettività è costante [13].

3.1 Shape from Silhouette: Visual Hull

Uno degli approcci possibili nella ricostruzione 3D è quello dell'intersezione volumetrica, presentato dal Prof. Aldo Laurentini, che consiste nell'intersezione dei volumi limitati dalle silhouette dell'oggetto. Come visibile in 3.1, tali volumi sono ricavati dalla proiezione da un punto di osservazione V della sagoma (o silhouette) dell'oggetto. L'insieme dei punti di visuale V definisce una regione dello spazio detta regione di osservazione e indicata con R , ovvero $R = \{V_i\}$. Le proiezioni creano delle coniche al cui interno è contenuto il volume dell'oggetto, ma affinché sia possibile ottenere abbastanza informazioni per una ricostruzione 3D è necessario porre un'ulteriore ipotesi: che l'oggetto sia ripreso da un certo numero di angolazioni.

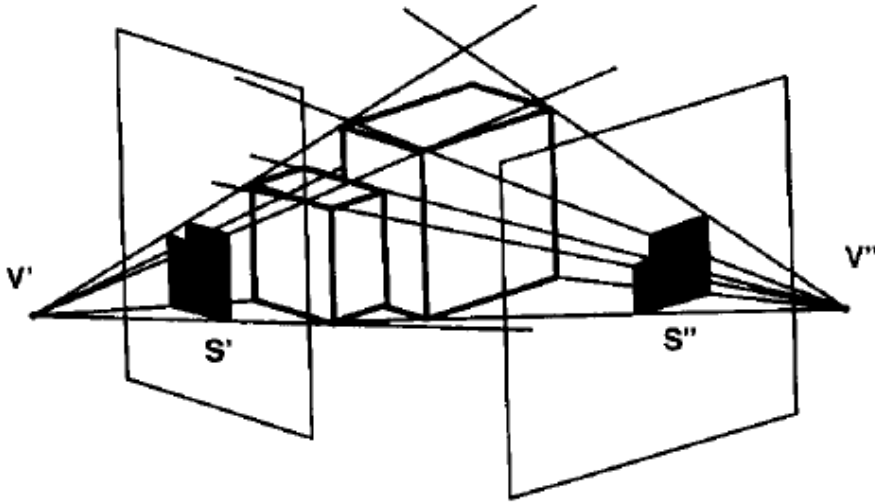


Figura 3.1: ricostruzione di oggetti attraverso l'intersezione volumetrica.

Il volume trovato con queste intersezioni si dice *Visual Hull* ed è così definito:

”Il Visual Hull $VH(S, R)$ di un oggetto S relativo a una regione di visuale R è una regione dello spazio tridimensionale E^3 tale che $\forall P \in VH(S, R) \subset E^3$ e \forall punto di vista $V \in R$, la semiretta che inizia da V passante per P interseca almeno un punto di S .”

Questa definizione, proposta da Laurentini stesso in [15], esprime in modo rigoroso l'idea che il Visual Hull ¹ consista nell'involucro volumetrico di tutte

¹per brevità da qui in poi VH.

le possibili coniche circoscritte di S e centrate nei punti V . Tuttavia, la sola definizione non è sufficiente per caratterizzare il VH: di seguito si riportano le sue proprietà generali.

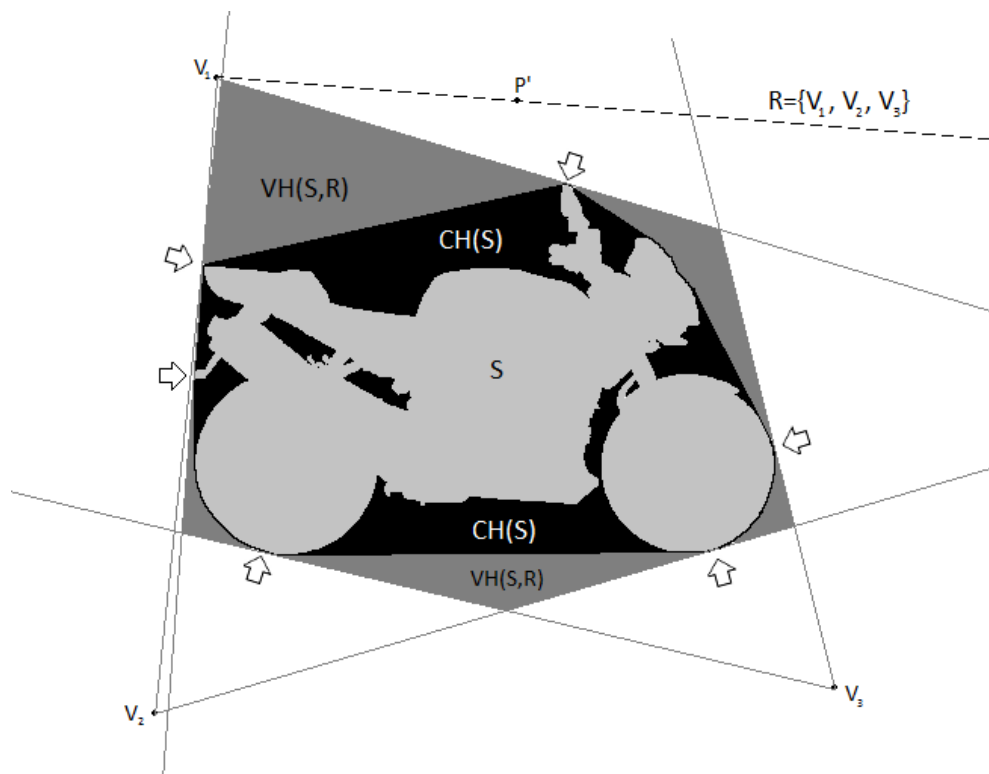


Figura 3.2: Convex Hull e Visual Hull di un oggetto S . Le frecce evidenziano i punti appartenenti a $sa(S,R)$.

Proprietà 3.1.1 $VH(S,R)$ è l'oggetto massimale silhouette-equivalente ad S rispetto a R , ovvero $VH(S,R)$ è la migliore approssimazione di S ottenibile attraverso intersezioni di volumi: dati due oggetti distinti S ed S' , questi sono distinguibili dalla stessa R se e solo se $VH(S,R) \neq VH(S',R)$.

In base alla definizione di VH, tutti i punti di S sono proiettati all'interno di VH stesso, mentre non tutti i suoi punti appartengono a S , quindi $S \leq VH(S,R)$ ¹; inoltre, come visibile in figura 3.2, dato un punto $P' \notin VH(S,R)$, esiste almeno un $V' \in R$ (nel caso in figura è il vertice V_1) tale che la semiretta con origine in V' e passante per P' non interseca S , ovvero P' non appartiene a un oggetto equivalente rispetto alla silhouette di S definita da V' . Questo rende anche $VH(S,R)$ la migliore approssimazione ottenibile attraverso l'intersezione volumetrica.

¹Laurentini considera l'oggetto S come unione del suo volume e della sua superficie.

Proprietà 3.1.2 Se $R \supset R'$, allora $VH(S, R) \leq VH(S, R')$

Aumentando i punti di vista l'intersezione avviene tra un maggior numero di coniche. Quindi è possibile rimuovere un maggior numero di punti e ridurre il VH corrispondente, migliorando l'approssimazione. Un esempio di questa proprietà è visibile in figura 3.3

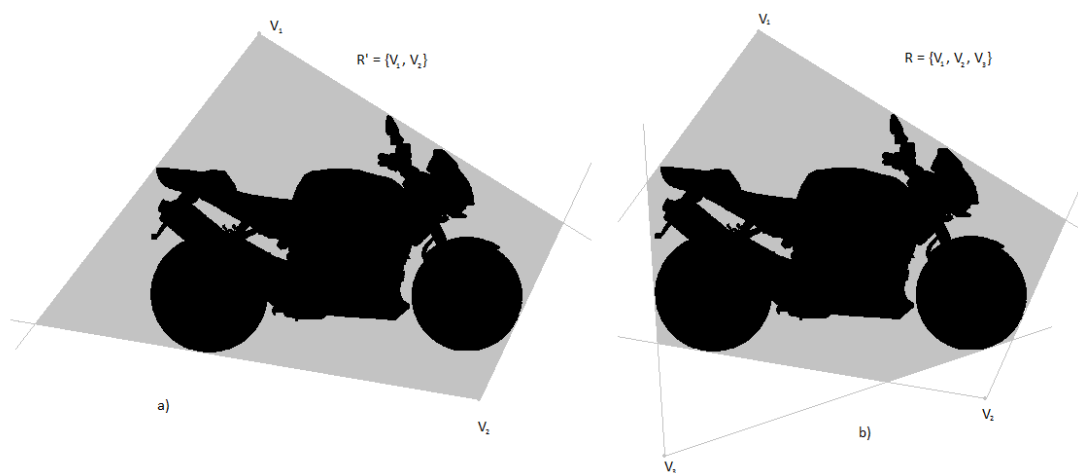


Figura 3.3: A parità di oggetto, l'estensione del VH è inversamente proporzionale alla dimensione delle regioni di visuale: in a) il VH è relativo a due punti di vista, in b) il VH è relativo a tre.

A questo punto è possibile dare una serie di altre definizioni utili al problema in esame:

- $sa(S, R)$, la superficie *silhouette-active* di S relativa a R , cioè la parte della superficie di S che appartiene anche al contorno di VH , ovvero la superficie che è possibile ricostruire esattamente usando le silhouette ottenute da R (facendo riferimento alla figura 3.2, appartengono a $sa(S, R)$ i punti evidenziati con le frecce);
- $si(S, R)$, la superficie *silhouette-inactive* di S , cioè la superficie che presenta tutte le caratteristiche (tipicamente concavità) di S che non modificano nessuna silhouette ottenuta da R (facendo riferimento alla fig. 3.2, appartengono a $si(S, R)$ i dettagli interni e la parte dietro la ruota anteriore);

La prima definizione porta direttamente alla proprietà 3.1.3, fondamentale dal punto di vista operativo:

Proprietà 3.1.3 $sa(S,R)$ determina univocamente $VH(S,R)$ di un oggetto.

Tale proprietà discende direttamente dal fatto che tutti i coni proiettivi circoscritti a un oggetto sono determinati univocamente dalla sua superficie attiva. In base alla definizione data esistono infiniti $VH(S,R)$ di un oggetto S , uno per ogni regione R . Tuttavia per quanto questo sembri diminuire l'efficacia dell'idea, è evidente che ad un oggetto S è possibile associare un VH principale, al quale all'atto pratico si può fare riferimento. Tale VH è rappresentato dal *Convex Hull* dell'oggetto S . Come visibile in figura 3.2, il Convex Hull o Convex Envelope di un oggetto S è l'insieme minimo convesso che contiene S stesso¹ e rappresenta il VH dell'oggetto S rispetto alla regione R^∞ , cioè alla regione contenente tutti i punti di osservazione dello spazio. Si consideri ora $VH(S,R')$ relativo a una regione R' che comprenda interamente S e si supponga inoltre che $R' \in R_c = E^3 - CH(S)$, ovvero che nessun punto di osservazione appartenga al CH stesso; in queste condizioni si possono formulare le proprietà 3.1.4 e 3.1.5

Proprietà 3.1.4 $VH(S,R') \leq CH(S)$

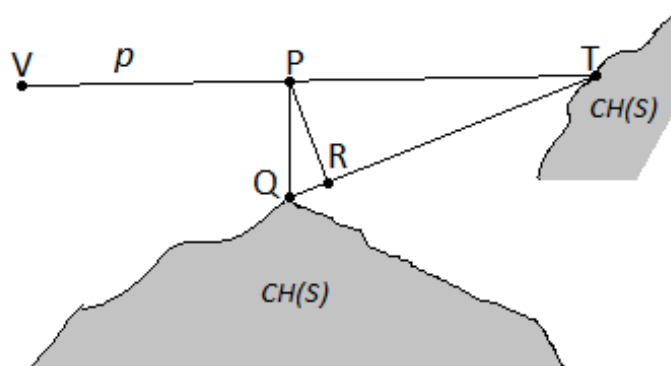


Figura 3.4: Il VH di S relativo a regioni che lo includono totalmente appartiene al CH .

Dualmente a quanto visto in 3.1.1 e come visibile in figura 3.4, i punti P del tipo $P \notin CH(S)$ creano delle linee di visuale che non intersecano S e giacciono su un piano p passante per P e normale al segmento PQ , dove Q è il punto di $CH(S)$ più vicino a P . Supponiamo per ipotesi che la linea di visuale intersechi $CH(S)$ nel punto T . Allora si avrebbe che il punto R , appartenente al CH in quanto

¹Si noti come esso sia funzione del solo oggetto S e indipendente dalla regione di visuale R considerata.

combinazione lineare di T e P e tale che $\overline{QT} \perp \overline{PR}$, sarebbe più vicino a P di Q, il che è un assurdo.

Proprietà 3.1.5 *Siano R' e $R'' \in R_c = E^3 - CH(S)$, allora $VH(S, R'') = VH(S, R')$*

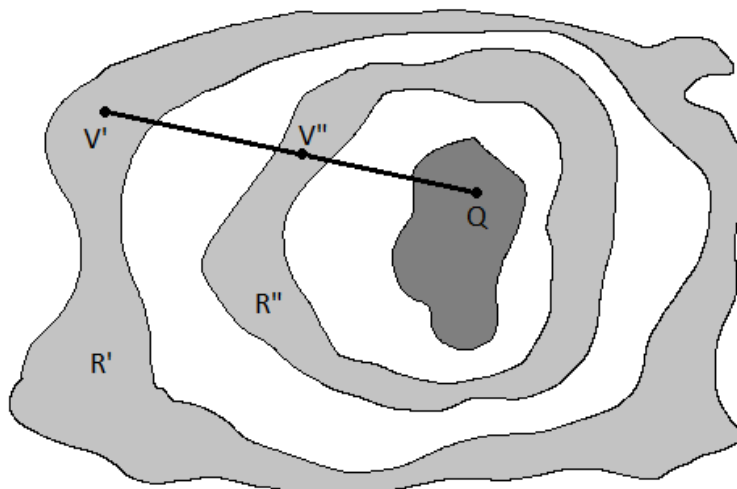


Figura 3.5: Tutti VH di S relativi a regioni che lo includono totalmente sono uguali.

Le figure 3.5 e 3.6 mostrano le relazioni tra i vari VH quando le regioni R circondano totalmente l'oggetto S. Nella prima si vede infatti come qualsiasi punto¹ appartenente a una regione R' trovi un corrispettivo in R''. Questa proprietà rimane valida finché le regioni R non si sovrappongono al CH dell'oggetto: facendo riferimento alla figura 3.6 la regione R' genera un VH che comprende il quadrilatero ABCD, mentre R'' ne genera uno che comprende solo ABCDE.

In conclusione esiste un unico VH che non penetri il CH di S e allo stesso tempo sia relativo a tutte le regioni R che lo racchiudono, tale VH si definisce *Visual Hull esterno*, o semplicemente visual hull di S, $VH(S) = VH(S, R_c)$. Poiché il CH non viene escluso dalla conclusione appena tratta, VH(S) può essere definito come il più esteso VH di S la cui regione di visuale R è limitata dal VH stesso.

Dualmente al VH esterno, è possibile definire un IVH(S), o *Visual Hull interno*, con banali riadattamenti alle considerazioni fatte finora. Com'è intuibile, esiste una relazione tra VH(S) e IVH(S), ed è la seguente:

¹Ovviamente si intende qualsiasi punto V da cui origina una semiretta che interseca il VH in un punto Q.

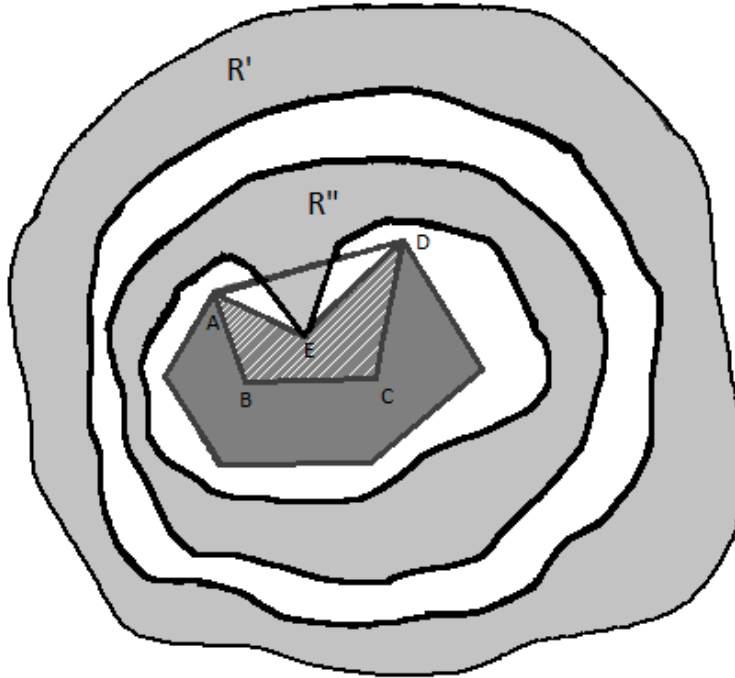


Figura 3.6: Se R'' comprende parti del CH allora i due VH sono differenti.

Proprietà 3.1.6 $IVH(S) \leq VH(S)$

Finora il VH di un oggetto è stato trattato dal punto di vista topologico, tuttavia esso nasce come risposta al problema della ricostruzione 3D e non solo dell'oggetto. È quindi evidente come sia possibile utilizzare il VH all'interno di tecniche o algoritmo della computer vision. Tuttavia, prima di fornire un algoritmo implementabile, è necessario fare alcune premesse: innanzitutto le linee di visuale passanti per un punto V vengono divise in settori di fasci di rette, definiti da Laurentini *famiglie di linee di visuale* [15]. Tali famiglie vengono utilizzate da un altro parametro, il *numero visuale*, che conta per ogni punto il numero di famiglie di linee di visuale che non intersecano alcuna parte dell'oggetto¹. Infine si definisce il segmento attivo, che consiste in una parte della linea che fa da confine tra punti i cui numeri visuali differiscono di uno. I segmenti attivi identificano quindi linee attive, limitate alle sole linee tangenti all'insieme di poligoni di cui l'oggetto in esame è composto in due punti, e intersecanti in nessuno. Si definisce quindi direzione di attraversamento positiva il lato della linea il cui numero visuale è maggiore, negativa quello in cui è minore.

¹I punti appartenenti al VH avranno quindi per definizione numero visuale nullo.

3.1.1 Euristica Generale

Fatte tali premesse, l'algoritmo generale per calcolare $VH(SP)^1$, cioè il VH corrispondente a un generico insieme di poligoni appartenenti a un piano, è il seguente:

1. Determinare tutti i segmenti attivi;
2. Costruire una nuova partizione del piano usando i segmenti trovati al punto precedente;
3. Iniziando da una regione interna a SP (il cui numero visuale è zero), verificare il numero visuale delle altre partizioni, aggiungendo o sottraendo 1 a seconda della direzione di attraversamento del segmento attivo;
4. Unire tutte le regioni il cui numero visuale è nullo.

Il lettore accorto avrà notato che questo algoritmo è valido per i soli insiemi di poligoni, e quindi per immagini 2D; tuttavia è facilmente estendibile al caso 3D, ovvero all'analisi di poliedri, ripetendo l'algoritmo per ogni piano supportato dalle facce dell'oggetto in esame.

3.1.2 Euristica Applicata

In questa tesi l'algoritmo implementato è leggermente diverso, in quanto sfrutta l'uso delle matrici proiettive per semplificare l'euristica teorica; nel caso specifico le immagini a disposizione forniscono esattamente il CH dell'oggetto e quindi si è implementato quanto concerne la parte finale (cioè l'unione di regioni con numero visuale nullo); inoltre, avendone la possibilità, sono stati utilizzati direttamente i voxel 3D invece di ripetere l'analisi per ogni piano supportato dalle facce dell'oggetto. L'algoritmo può essere così descritto:

1. Per ogni immagine I_i *segmentazione* [16], ovvero identificazione dell'area $\tilde{S}_i = P_i(S)$ proiezione della superficie S dell'oggetto nell'immagine i-esima;
2. Creazione di un volume arbitrario V, da cui verranno eliminate di volta in volta le regioni non vincolate dalle proiezioni;

¹Set di Poligoni.

3. PREMESSA TEORICA

3. Suddivisione del volume V in unità elaborabili, per esempio attraverso creazione di una griglia 3D uniforme. Ad ogni sottounità, o voxel, v di V può essere associato un valore 0 o 1 che ne rappresenta la consistenza. Inizialmente viene posto $V(v) = 1 \forall v$.
4. Applicazione dell'euristica seguente per ogni voxel v appartenente a V : $\exists I_i : P_i(v) \notin S_i \Rightarrow V(v) = 0$;
5. Una volta terminata l'assegnazione di tutti i valori $V(v)$, identificare la iso-superficie $S_0 = x : \lim_{x^* \rightarrow x} V(x^*) = 1^-$

L'iso-curva S_0 così ottenuta è la migliore approssimazione della superficie di S ottenibile con questo metodo.

3.2 Geometria Proiettiva

Nel passaggio dal dominio tridimensionale a quello bidimensionale come avviene per esempio facendo una fotografia, le dimensioni dell'oggetto ripreso vengono inevitabilmente alterate. Questo comporta che in un successivo ritorno alla rappresentazione 3D vi siano ambiguità anche sostanziose, che impongono l'applicazione di modelli prospettici. Tali modelli essenzialmente caratterizzano deformazioni quali le alterazioni prospettiche o rototraslazioni dovute al sistema ottico utilizzato nell'acquisizione dell'immagine; queste sono oggetto di studio di una specifica branca della matematica, la geometria proiettiva. Essa comprende non solo l'analisi delle trasformazioni di cui sopra, ma anche la creazione di modelli che approssimino quanto meglio la situazione reale, come quello qui adottato, noto come *pinhole camera*.

3.2.1 Pinhole Camera

Il sistema utilizzato per questa tesi è composto di tre componenti, la sorgente di radiazione, l'oggetto bersaglio e il rivelatore. Poiché la sorgente è approssimabile con un punto nello spazio che la identifica, il modello che ne fornisce una rappresentazione immediata è quello della pinhole camera, visibile in figura 3.7. Essa rappresenta il modello prospettico più semplice, caratterizzato da un piano Π , detto piano immagine, e dal punto di presa C . Inoltre ben riflette la situazione di questa tesi: al trittico sorgente X - oggetto - rivelatore corrisponde quello asse ottico (originante in c) - piano immagine (Π) - punto P .

Un'altra importante distinzione da farsi è quella tra parametri intrinseci e parametri estrinseci del sistema¹. I parametri intrinseci sono visibili sempre in figura 3.7, e consistono nella *lunghezza focale* f_c nelle coordinate del *punto principale* (x_c, y_c) . È importante notare come questi valori, essendo indipendenti dall'oggetto ritratto, siano in pixel. Questi parametri sono importanti per poter giungere alla *Matrice Intrinseca* \tilde{K} :

¹Distinzione particolarmente importante in quanto lo schema della pinhole camera generalmente appartiene all'ottica, mentre qui verrà applicato a un problema di *imaging*.

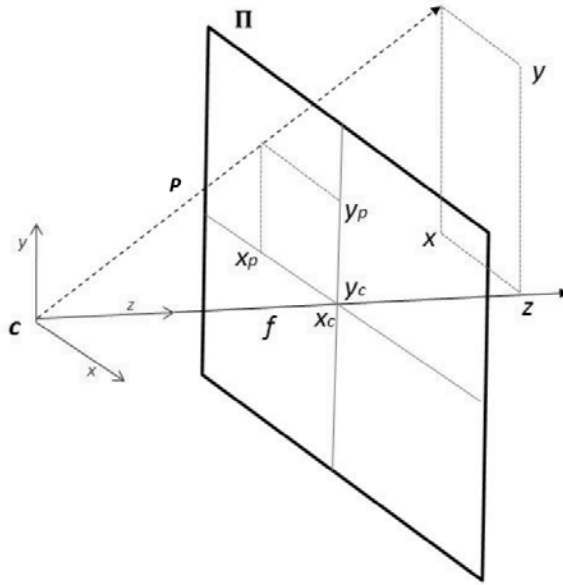


Figura 3.7: Modello proiettivo di una pinhole camera.

$$\tilde{K} = \begin{bmatrix} f & 0 & x_c \\ 0 & f & y_c \\ 0 & 0 & 1 \end{bmatrix} \implies \begin{bmatrix} i_p \\ j_p \\ 1 \end{bmatrix} \times \tilde{K} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

L'equazione qui riportata, pur corrispondendo al modello della pinhole camera, non tiene conto del caso reale, in cui bisogna tener conto che il sensore della camera consiste in una matrice di pixel, ognuno di dimensione (p_x, p_y) . Inoltre ciascun pixel, idealmente perfettamente rettangolare, può presentare deformazioni dovute a svariati motivi; per tener conto di queste irregolarità si introduce un fattore s che misura lo skew, cioè l'angolo di inclinazione rispetto alla verticale del lato corrispondente del pixel. La matrice intrinseca \tilde{K} diventa quindi

$$K = \begin{bmatrix} f_x & s & \tilde{x}_c \\ 0 & f_y & \tilde{y}_c \\ 0 & 0 & 1 \end{bmatrix}$$

dove $f_x = \frac{f}{p_x}$, $f_y = \frac{f}{p_y}$, indicano la lunghezza focale misurata in altezza e larghezza dei pixel, $s = \frac{\tan(\alpha)f}{p_y}$ è il valore del fattore di skew come definito in precedenza e $(\tilde{x}_c = \frac{x_c}{p_x}, \tilde{y}_c = \frac{y_c}{p_y})$ sono le coordinate del punto principale, sempre espresse in pixel.

La matrice intrinseca K descrive completamente il contributo di una pinhole camera a ottiche fisse per ogni immagine che essa cattura. Tuttavia, volendo ricostruire un oggetto da una successione di immagini, occorre tenere in conto il contributo dato da e relativo all'oggetto stesso, cioè bisogna calcolare i parametri estrinseci del sistema. Essi sono legati al passaggio dal sistema di riferimento della camera a quello dell'oggetto, che introduce una trasformazione di rototraslazione nel passare dal punto (x, y, z) alla sua rappresentazione (i_p, j_p) ¹. Detta R questa trasformazione e definendo \mathbf{c} come il centro ottico del nuovo sistema, si introduce la nuova relazione

$$\begin{bmatrix} \bar{J} \\ 1 \end{bmatrix} \propto {}^2KR(\mathbf{x} - \mathbf{c})$$

dove \bar{J} è il vettore 2D delle coordinate in pixel del punto proiezione di \mathbf{x} . Questa relazione può essere riscritta tramite l'introduzione della *Matrice di Proiezione* $P \in \mathcal{M}_{34}$:

$$\begin{bmatrix} \bar{J} \\ 1 \end{bmatrix} \propto P \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$$

dove

$$P = KR \begin{bmatrix} I_3 & -\mathbf{c} \end{bmatrix} \quad (3.1)$$

L'equazione 3.1 lega la terna del punto 3D alla sua proiezione 2D, ma per il problema opposto, cioè trovare il punto nello spazio di cui \bar{P} è proiezione bisogna fare altre considerazioni. Innanzitutto si noti che la matrice di proiezione P può essere scritta come unione di due matrici:

$$P = \begin{bmatrix} M & \mathbf{m} \end{bmatrix}, \text{ dove } \begin{cases} M = KR \in \mathcal{M}_{33} \\ \mathbf{m} = -KR\mathbf{c} \in \mathcal{M}_{31} \end{cases}$$

Quella qui presentata è la forma con cui si descrive comunemente la pinhole camera, che presenta inoltre il vantaggio di poter esprimere il centro ottico \mathbf{c} come $\mathbf{c} = -M^{-1}\mathbf{m}$.

Tale relazione permette infatti di esprimere il punto in forma vettoriale. Come visibile in figura 3.7, permette di determinare il vettore tridimensionale v_j che soddisfa la relazione $l : \mathbf{x} = \mathbf{c} + \lambda \mathbf{v}_j$, cioè il vettore che collega il centro proiettivo

¹Dove le due coordinate sono in pixel.

²Il simbolo \propto in questa sede rappresenta l'uguaglianza in senso prospettico [17].

c alla proiezione sul piano immagine \bar{P} , che giace sulla linea l . Tale vettore è definito dall'equazione

$$v_j = M^{-1} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix}$$

ed è la composizione di tre vettori: a_i, a_j, a_c , definiti, basandosi sempre sulla figura 3.7, come segue¹:

- a_i è il vettore che giace sulla linea che collega y_c a y_p ;
- a_j è il vettore che giace sulla linea che collega x_c a x_p ;
- a_c è il vettore che giace sulla linea che collega \mathbf{c} alla sua proiezione (x_c, y_c) .

Ergo

$$\mathbf{v}_j = \mathbf{a}_i i + \mathbf{a}_j j + \mathbf{a}_c = M^{-1} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix} = R^T K^{-1} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix}$$

cioè i vettori a_i, a_j, a_c costituiscono le colonne di M^{-1} . Ma se consideriamo le righe r_1, r_2, r_3 di M^{-1} , allora valgono le relazioni

$$\begin{cases} \mathbf{v}_j = \frac{\mathbf{x}-\mathbf{c}}{\mathbf{r}_3(\mathbf{x}-\mathbf{c})} \\ \mathbf{a}_i = \frac{\mathbf{r}_1^T}{f_x} \\ \mathbf{a}_j = \frac{\mathbf{r}_2^T}{f_y} \end{cases}$$

che portano al risultato finale

$$\mathbf{v}_j = \frac{\mathbf{r}_1^T}{f_x} i + \frac{\mathbf{r}_2^T}{f_y} j + \left(\mathbf{r}_3^T - \mathbf{r}_1^T \frac{x_c}{f_x} - \mathbf{r}_2^T \frac{y_c}{f_y} \right) \quad (3.2)$$

L'equazione 3.2 è molto importante in quanto permette, una volta note le coordinate del punto immagine \bar{P} e la parametrizzazione della pinhole camera, di definire il vettore V_j e quindi la linea l che connettono il punto nello spazio \mathbf{x} , il centro ottico \mathbf{c} e la proiezione \bar{P} .

Quanto scritto in questo capitolo permette di modellizzare il passaggio dalle coordinate tridimensionali e metriche dell'oggetto a quelle bidimensionali e in pixel della sua rappresentazione (e viceversa). La ricostruzione SfS tuttavia si

¹Si noti come siano tutti e tre vettori tridimensionali.

basa sulla sovrapposizione di informazioni ottenibili da più immagini per poter arrivare al VH che rappresenta l'oggetto 3D: è quindi evidente come sia necessario introdurre una qualche relazione tra le immagini acquisite. In questa tesi tale legame è rappresentato dal fatto che le immagini ritraggano lo stesso oggetto di volte in volta ruotato ed è modellizzato dalla matrice fondamentale.

3.2.2 Matrice Fondamentale

La *matrice fondamentale* F del sistema rappresenta il legame esistente tra due prospettive dello stesso oggetto, ovvero tra due punti corrispondenti. Essa è definita come segue:

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} : \mathbf{x}'^T F \mathbf{x} = 0 \quad (3.3)$$

dove $\mathbf{x} = [i, j, 1]^T$ e $\mathbf{x}' = [i', j', 1]^T$ sono i due punti corrispondenti di due immagini le cui coordinate sono note a priori.

Dato un numero sufficiente di coppie corrispondenti, è possibile imbastire un sistema di equazioni che permetta di calcolare F [18]. Infatti introducendo il vettore riga \mathbf{f} , composto dalle righe di F , cioè $\mathbf{f} = (f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33})$, la relazione 3.3 è esprimibile come il seguente prodotto interno tra vettori:

$$(x'x, x'y, x'y', y'x, y'y, y'x, y, 1)\mathbf{f} = 0 \quad (3.4)$$

Se si suppone di avere n punti corrispondenti, per ciascuno di essi vale il prodotto 3.4: è quindi possibile creare un sistema di equazioni che portano a

$$\mathbf{A}\mathbf{f} = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix} \mathbf{f} = 0 \quad (3.5)$$

Il sistema 3.5 così introdotto è un sistema lineare omogeneo, che ha soluzione se e solo se la matrice A ha almeno rango 8, e in questo caso la soluzione è unica. Solitamente però la matrice A ha rango 9,¹ per cui la soluzione trovata non sarà quella esatta ma bensì quella ai minimi quadrati. La soluzione ai minimi quadrati per \mathbf{f} è il “vettore singolare”, che corrisponde al minimo valore singolare di A , ricavabile dalla SVD² $A = UDV^T$ come ultima colonna della matrice V .

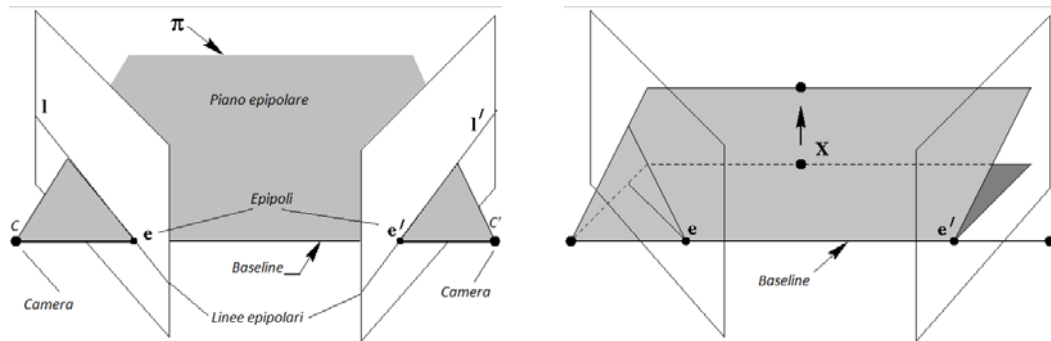
¹Dovuto al rumore nelle coordinate dei punti.

²**Singular Value Decomposition:** presa una generica matrice A , la sua SVD è una sua fattorizzazione per cui $A = UDV^T$, dove U e V sono matrici ortogonali e D è una matrice diagonale a valori non-negativi.

3.2.3 Matrici Proiettive

Come già asserito, la matrice fondamentale F lega tra loro punti corrispondenti di immagini differenti attraverso la relazione $x'^T F x = 0$. È facilmente immaginabile come questo concetto possa essere espanso dalla corrispondenza di singoli punti alla corrispondenza tra prospettive, argomento di cui si occupa la geometria epipolare. La geometria epipolare è in realtà la geometria proiettiva intrinseca tra due visuali, ed è ben rappresentata dalla matrice fondamentale stessa. [25] Prima di proseguire è necessario dare alcune definizioni:

- L'**epipolo** è il punto di intersezione tra il piano immagine e la linea, detta *baseline*, che congiunge le due camere, identificate dai rispettivi centri di visuale;
- Un **piano epipolare** è un piano che contiene la baseline¹;
- Una **linea epipolare** è l'intersezione tra un piano epipolare e il piano immagine.



(a) Entità fondamentali della geometria epipolare.

(b) Man mano che il punto 3D \mathbf{X} si sposta, il piano epipolare “ruota” attorno alla baseline: questa famiglia di piani si definisce *pencil epipolare*.

Figura 3.8: Geometria Epipolare.

L'utilizzo geometria epipolare ben si presta all'esperimento di questa tesi in quanto permette l'espressione della matrice fondamentale F in funzione di due

¹Esiste una famiglia di piani epipolari, caratterizzata da un unico parametro. Tale famiglia si definisce *pencil*.

matrici di proiezione P e P' relative a due immagini successive dell'oggetto. Infatti, dato un punto del piano immagine \mathbf{x} corrispondente al punto nello spazio \mathbf{X} , il raggio che lo collega al centro ottico \mathbf{C} attraverso P è soluzione di $P\mathbf{X} = \mathbf{x}$. Similmente al modello della pinhole camera, si vuole rappresentare queste soluzioni come dipendenti da un unico parametro, cioè

$$\mathbf{X}(\lambda) = P^+\mathbf{x} + \lambda\mathbf{C} \quad (3.6)$$

dove P^+ è la matrice pseudo-inversa¹ di P , \mathbf{C} il vettore soluzione di $P\mathbf{C} = 0$ e λ uno scalare che parametrizza il raggio.

Il raggio soluzione della 3.6 comprende in particolare due punti: il centro ottico \mathbf{C} (per $\lambda = \infty$) e il punto $P^+\mathbf{x}$ (per $\lambda = 0$). Questi punti vengono mappati dalla seconda camera (di matrice proiettiva P') rispettivamente nei punti $P'P^+\mathbf{x}$ e $P'\mathbf{C}$, formando così la linea epipolare $\mathbf{l}' = (P'\mathbf{C}) \times (P'P^+\mathbf{x})$. Poiché il punto $P'\mathbf{C}$ è l'epipolo nella seconda immagine, introducendo la notazione

$$e' = [e_1, e_2, e_3]^\top \Rightarrow [e']_\times = \begin{bmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{bmatrix}$$

si arriva al seguente risultato:

$$\mathbf{l}' = (P'\mathbf{C}) \times (P'P^+\mathbf{x}) = [e']_\times(P'P^+\mathbf{x}) = Fx$$

dove la matrice fondamentale F ha forma

$$F = [e']_\times P'P^+$$

.

¹In un sistema che rappresenta il problema di ricostruzione, la matrice pseudo-inversa P^+ di P è la matrice che permette di trovare una soluzione e contemporaneamente minimizza l'errore dovuto al rumore di osservazione. Essa ha la seguente espressione: $P^+ = (P^\top P)^{-1} P^\top$.

L'espressione così trovata è importante perché permette di esprimere \mathbf{F} come funzione di matrici intrinseche e di rotazione: ricordando infatti l'espressione delle matrici di proiezione¹ $P = K \begin{bmatrix} I & 0 \end{bmatrix}$, $P' = K' \begin{bmatrix} R & \mathbf{t} \end{bmatrix}$ si arriva a

$$P^+ = \begin{bmatrix} K^{-1} \\ 0^\top \end{bmatrix} \quad \mathbf{C} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

e quindi a

$$\begin{aligned} \mathbf{F} &= [P'\mathbf{C}]_{\times} P' P^+ = [K'\mathbf{t}]_{\times} K' R K^{-1} = K'^{-\top} [\mathbf{t}]_{\times} R K^{-1} = \\ &K'^{-\top} R [R^\top \mathbf{t}]_{\times} K^{-1} = K'^{-\top} R K^\top [K R^\top \mathbf{t}]_{\times} \\ &= \mathbf{K}'^{-\top} \mathbf{R} \mathbf{K}^\top [\mathbf{e}]_{\times} \quad (3.7) \end{aligned}$$

in quanto gli epipoli sono

$$\mathbf{e} = P \begin{pmatrix} -R^\top \mathbf{t} \\ 1 \end{pmatrix} = K R^\top \mathbf{t} \quad \text{ed} \quad \mathbf{e}' = P' \begin{pmatrix} 0 \\ 1 \end{pmatrix} = K' \mathbf{t}.$$

L'equazione 3.7 permette quindi di ricavare la matrice fondamentale note quella intrinseca e quella di rotazione da un'immagine all'altra: è evidente che la stessa equazione può essere invertita per ricavare la matrice di rotazione nota quella fondamentale, obiettivo di questa tesi. Tuttavia alla matrice che tiene conto della rotazione tra un'immagine e l'altra e che per ipotesi di lavoro è costante, bisogna accostare un'ulteriore matrice che consideri la rotazione dell'oggetto rispetto al sistema di riferimento assoluto. Le metodologie necessarie a trovare tali matrici sono argomento dell'autocalibrazione del sistema.

¹Si è supposta la matrice P relativa alla camera sita al centro del sistema di riferimento.

3.3 Autocalibrazione

Obiettivo di questa tesi è l'estrazione automatica dei parametri estrinseci del sistema di acquisizione ai fini della ricostruzione SfS¹. Come visto nel capitolo 3.1, lo SfS necessita, per poter ricostruire un oggetto, di informazioni ridondanti, ottenute da molteplici immagini con punti di vista differenti. A tal scopo in questa tesi si è deciso di far ruotare l'oggetto durante l'acquisizione, in modo da ottenere diverse silhouette, ognuna separata dalla successiva da un angolo costante. Una volta avute a disposizione le immagini, si è proceduto alla segmentazione, ovvero alla ricerca di punti o curve che possedessero caratteristiche tali da fornire la ridondanza richiesta.

Tra le possibili famiglie di punti adatti allo scopo, è stata scelta quella degli spigoli (o corner), cioè i punti dove le features dell'immagine presentano curvatura particolarmente accentuata. La scelta è ricaduta su tali punti in quanto spesso proiezione della superficie attiva. Da un'immagine alla seguente la curvatura di tali punti generalmente si mantiene pur modificandosi leggermente per effetti prospettici, e quindi essi permettono di stabilire relazioni tra un'immagine e l'altra. La scelta di tali punti è particolarmente delicata, in quanto un detector troppo sensibile porterebbe a falsi positivi e correlativamente a tempi di elaborazione elevati, mentre uno troppo approssimato non fornirebbe dati sufficienti. Il detector scelto, come si vedrà nelle prossime pagine, consente un buon equilibrio tra sensibilità e onerosità di calcolo. Successivamente all'individuazione è necessario mettere i punti trovati in relazione, in modo da avere una successione ordinata ovvero di seguire lo spigolo man mano che l'oggetto ruota.

3.3.1 Fast Chord-to-Point Distance Accumulation Corner Detector

Per gli scopi di questa tesi, elaborare tutti i punti di ogni silhouette a disposizione sarebbe incredibilmente oneroso, per cui è sufficiente trovare almeno un numero minimo di features che permettano di ricavare le informazioni necessarie. Le features di un'immagine sono quei punti o quelle caratteristiche che ne permettono l'elaborazione topologica: possono essere per esempio le linee di variazione cromatica, le zone d'ombra o, come nel caso in esame, gli spigoli o *corner*.

¹Shape from Silhouette.

In letteratura esistono principalmente tre famiglie di *corner detector*: contour-based, intensity-based e model-based. I primi estraggono gli spigoli ricavando le curve planari dell'immagine attraverso un rivelatore di bordi e successivamente cercando i punti di curvatura adatta, i secondi stimano una misura ritenuta indicatrice di un punto d'interesse lavorando direttamente sui pixel e gli ultimi cercano gli spigoli confrontando l'immagine con un modello predefinito. I detector contour-based possono essere suddivisi ulteriormente in altre due categorie, a scala multipla e a scala singola. Questi ultimi presentano sfide non indifferenti per trovare la scala che si adatti a tutti i corner trovati, ragion per cui i primi sono quelli maggiormente diffusi. Generalmente usano da una a tre diverse scale di smoothing per trovare gli spigoli e successivamente ne usano una piccola per localizzarli con maggior precisione.

La tecnica di stima della curvatura più diffusa è basata sullo *curvature space-scale* o CSS, che utilizza uno smoothing gaussiano di scala σ . Tuttavia il CSS presenta due grossi svantaggi, innanzitutto utilizza derivate fino al secondo ordine per la localizzazione dei punti delle curve, rendendo la curvatura stimata molto sensibile a rumore e imprecisa nel caso di curve ad alta variazione locale, secondariamente la scelta di σ non è banale, in quanto valori troppo bassi non effettuerebbero uno smoothing sufficiente, mentre valori troppo alti altererebbero eccessivamente la curva originale. [19] Per risolvere questi problemi sono stati ideati i detector basati sulla *chord-to-point distance accumulation* (CPDA), che non utilizzano alcuna derivata e sono meno sensibili a variazioni locali o rumore. Il generico detector CPDA segue cinque passi principali:

1. *Estrazione e selezione dei bordi*
2. *Smoothing della curva*
3. *Stima della curvatura*
4. *Raffinazione dell'insieme di candidati*
5. *Insieme finale di corner rivelati*

I cinque passi di tale euristica verranno analizzati in dettaglio nei paragrafi seguenti.

Estrazione e selezione dei bordi

Il detector CPDA estrae le curve planari attraverso un rilevatore di bordi di Canny con soglie $low = 0.2$ e $high = 0.7$.¹ Poi avviene una prima scrematura delle curve trovate, in quanto vengono mantenute le sole curve la cui lunghezza n rispetta il limite $n > (a+l)/\alpha$, dove a e l sono altezza e larghezza dell'immagine e α è il parametro che limita la lunghezza del bordo, pari a 25.

Smoothing della curva

Ogni singola curva $\Gamma(t) = (x(t), y(t))$ viene smussata attraverso una convoluzione gaussiana

$$G(t, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-t^2}{2\sigma^2}\right)$$

dove σ si dice *scala di smoothing*. Per migliorare il posizionamento viene scelto a seconda della lunghezza della curva un σ diverso, 1 se $n \leq 100$, 2 se $100 < n \leq 200$ e 3 se $n > 200$.

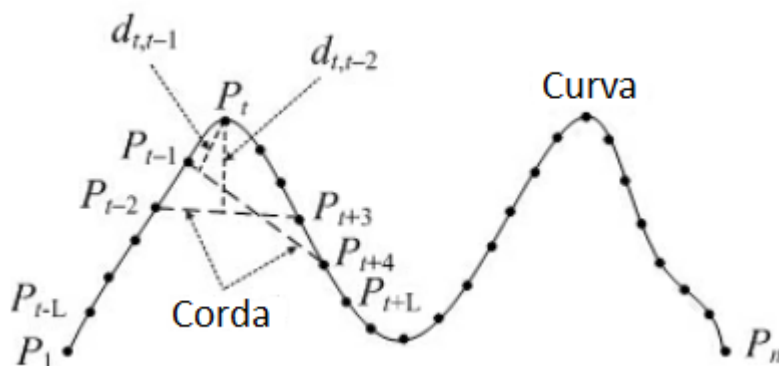


Figura 3.9: Chord-to-point distance accumulation per una corda di lunghezza $L=5$.

Stima della curvatura

Questo passo, raffigurato in figura 3.9, è il nucleo del detector CPDA. Una corda di lunghezza L viene fatta scorrere lungo ciascuna curva trovata al passo precedente e per ogni suo punto P_t le distanze perpendicolari $d_{t,t-1}$ tra esso e la corda vengono sommate, calcolandone così la curvatura $h_L(t)$, ovvero

¹NB: I valori dei parametri utilizzati in questa tesi sono alternativamente valori determinati da altri studi e valori determinati empiricamente.

$$h_L(t) = \sum_{j=t-L+1}^{t-1} d_{t,j}$$

Il detector CPDA esegue questo calcolo con tre corde di diversa lunghezza, rispettivamente $L_1 = 10$, $L_2 = 20$, $L_3 = 30$, cui corrispondono tre valori di curvatura $h_1(t)$, $h_2(t)$, $h_3(t)$: questo permette di ovviare alle ambiguità non scalabili. Una volta ottenute le curvature, queste vengono normalizzate secondo la seguente formula:

$$h'_j(t) = \frac{h_j(t)}{\max(\text{abs}(h_j))}, \text{ per } 1 \leq t \leq n \text{ e } 1 \leq j \leq 3$$

Infine, la curvatura totale $H(t)$ viene ottenuta secondo la formula

$$H(t) = \prod_{j=1}^3 h'_j(t), \text{ con } 1 \leq t \leq n.$$

Questo approccio ha un grande vantaggio: moltiplicando le tre curvature si accentua la differenza tra vertici deboli e forti, in quanto questi ultimi hanno valori di curvatura superiori ai primi. Ciò permette di eliminare quelli deboli attraverso una singola soglia sul valore di curvatura.

Raffinazione dell'insieme di candidati

A questo punto vengono raggruppati tutti i massimi locali della funzione $H(t)$ per tutte le curve appartenenti all'insieme dei candidati. Si noti come un massimo locale possa essere indifferentemente forte, debole o falso. Le ultime due categorie dovrebbero essere escluse dalla lista dei corner rilevati. A questo punto vengono introdotti una seconda soglia sul valore di curvatura $T_h = 0.2$, sotto la quale il corner viene scartato, e un controllo sull'angolo δ formatosi tra il corner e i suoi vicini¹; se tale angolo è maggiore del valore di soglia $\delta_{Th} = 157^\circ$ allora il corner viene scartato come falso.

Insieme finale di corner rivelati

Dopo aver rimosso angoli deboli e falsi positivi, il gruppo finale viene ulteriormente raffinato considerando che potrebbero esserci corner tra gli estremi di un

¹Oppure con le estremità della curva qualora non siano disponibili vicini.

cappio: in questo caso viene verificato l'angolo con un estremo della curva e se questo è minore di δ_{Th} , allora il corner viene aggiunto all'insieme finale. Infine vengono aggiunti, se presenti, solamente i corner a T lontani da quelli già rilevati, dove "lontano" corrisponde a quanto si trova all'esterno di un'area di 5x5 pixel centrata sul punto.

FCPDA

Poichè l'algoritmo del detector CPDA esegue la stima della curvatura per tutti i punti di tutte le curve trovate dopo l'applicazione del filtro di Canny, è intuibile come nella sua forma base questo sia un corner detector decisamente esoso in termini di risorse necessarie. Una soluzione a questo problema è quella di ridurre l'insieme dei punti dove calcolare la curvatura sfruttando il seguente fatto: una curva con cambi di direzione significanti è affetta maggiormente dallo smoothing gaussiano che non una dritta, ovvero lo smoothing gaussiano interessa maggiormente la prima rispetto alla seconda. [20] Questo comporta che la distanza di un punto tra la curva originale e quella smussata sia elevata in un intorno della posizione del corner, e bassa se non ci sono corner vicini, ragion per cui i massimi della funzione distanza F_d ¹ offrono direttamente un insieme di candidati.

Tuttavia, poiché la curva d'origine potrebbe essere rumorosa o avere un'alta variazione locale e quindi generare falsi positivi, conviene usare due scale di smoothing, $\sigma_1 < \sigma_2$. Un piccolo valore di σ_1 garantisce una buona localizzazione dei candidati, mentre un alto valore di σ_2 ne indica la robustezza. È quindi chiaro che la scelta di questi due parametri è particolarmente delicata².

La funzione distanza F_d su cui si basa questa modifica all'algoritmo CPDA è definita come la distanza euclidea tra le coppie punti delle curve smussate:

¹Che consiste nelle distanze punto-punto tra la curva originale e quella smussata.

²In questo ambito sono stati scelti i valori di 3.0 e 4.0 per raggiungere un compromesso, in quanto valori troppo piccoli di σ_1 rendono il posizionamento nuovamente sensibile al rumore, mentre valori troppo grandi di σ_2 potrebbero cancellare candidati idonei.

$$\begin{aligned}\Gamma(t, \sigma_1) &= \Gamma(t) \otimes G(t, \sigma_1) \\ \Gamma(t, \sigma_2) &= \Gamma(t) \otimes G(t, \sigma_2) \\ \mathbf{F}_d(\mathbf{t}, \sigma_1, \sigma_2) &= [x(t) \otimes G(t, \sigma_2) - x(t) \otimes G(t, \sigma_1)]^2 + \\ &\quad + [y(t) \otimes G(t, \sigma_2) - y(t) \otimes G(t, \sigma_1)]^2 \\ &= (\dots) \\ &= [\mathbf{x}(\mathbf{t}) \otimes \mathbf{G}_d(\mathbf{t}, \sigma_1, \sigma_2)]^2 + [\mathbf{y}(\mathbf{t}) \otimes \mathbf{G}_d(\mathbf{t}, \sigma_1, \sigma_2)]^2\end{aligned}$$

dove $G_d(t, \sigma_1, \sigma_2)$ è la distanza tra due funzioni gaussiane.

A questo punto viene preso come insieme di candidati direttamente quello dei massimi della funzione F_d . In questo modo si è ridotto di molto il numero di punti sui quali calcolare la somma delle distanze dalle corde, e con esso il tempo di esecuzione del codice, da cui il nome di questa variante, cioè detector FCPDA o *Fast Chord-to-Point Distance Accumulation*.

3.3.2 Cernita delle features

L'algoritmo FCPDA precedentemente introdotto ricava per ogni immagine un insieme di punti, ciascuno corrispondente a una feature dell'immagine stessa. Evidentemente non tutti i corner rilevati saranno adatti a un'elaborazione successiva¹, ragion per cui è necessario introdurre una relazione tra le varie features trovate. L'obiettivo cui si mira è ottenere un numero di sequenze di punti che permetta di ricavare la matrice fondamentale del sistema. Poiché tali punti permettono di seguire l'evolversi della feature tra le varie immagini, sono stati definiti *ereditari*. Inoltre giacché da questo passo in poi le immagini iniziali non vengono più considerate, per immagine I_i si intende l'insieme delle features ovvero degli ereditari ricavato tramite l'algoritmo FCPDA dalla silhouette i -esima.

Inizialmente viene calcolata per ogni coppia di immagini consecutive (I_i, I_{i+1}) la matrice $D(i, i+1)$, cioè la distanza euclidea² tra ciascun ereditario di I_i e tutti quelli di I_{i+1} . Per costruzione quindi si ha che vale la seguente proprietà:

Proprietà 3.3.1 *la riga k di $D(i, i + 1)$ fa riferimento all'ereditario k -esimo di I_i , ogni colonna j fa riferimento all'ereditario j -esimo di I_{i+1} .*

Una volta costruita la matrice $D(i, i + 1)$, il minimo di ogni riga identifica univocamente, secondo la proprietà 3.3.1, l'elemento di I_{i+1} più vicino a I_i e quindi l'ereditario successivo più probabile. Questo rende sensata l'idea di costruire una sequenza per ogni ereditario che ne rappresenta l'evoluzione coerente da una silhouette all'altra ovvero rappresenta la concatenazione di ereditari successivi, ragion per cui tali sequenze sono state definite *catene*. Poiché immagini differenti presentano features differenti, è necessario operare una scelta che tenga conto dei seguenti fattori:

- corner scomparsi in seguito a rotazioni dell'oggetto;
- corner apparsi in seguito a rotazioni dell'oggetto;
- catene che muoiono sullo stesso ereditario generando ambiguità (*catene confluenti*);

¹Il detector FCPDA non fa distinzione sulla qualità della feature rilevata, né può verificare se questa rimanga visibile per un numero sufficiente di immagini.

²Poiché le immagini possono avere dimensioni rispettivamente r, c con $r \neq c$, si ha che $D(i, i + 1) \in \mathcal{M}_{r \times c}$.

- catene che iniziano da ereditari di immagini successive alla prima (*catene orfane*).

I corner che compaiono o scompaiono rappresentano due manifestazioni differenti dello stesso problema, in quanto entrambi sono dovuti alla rotazione dell'oggetto non potendo essere risolti tali punti vengono comunque considerati affidandosi alla ridondanza dell'informazione. Fortunatamente costituiscono una parte relativamente piccola di tutti gli errori, in quanto è poco probabile che una feature scompaia in una foto per essere sostituita da un'altra, diversa, in quella successiva, ragion per cui non è stata implementata una soluzione specificatamente dedicata.

Tuttavia esiste uno scenario, le catene confluenti, in cui il seguente fattore va preso in considerazione: supposto di avere una feature f_1 che nell'immagine successiva sparisce per vari motivi, è tutt'altro che improbabile che ce ne sia un'altra, f_2 , sufficientemente vicina da essere considerata dall' algoritmo come ereditario successivo più probabile. In questo caso si avrebbe un'unica catena di ereditari invece che due catene distinte, il che porterebbe a errori nella ricostruzione della matrice fondamentale. Si riposta un esempio di tale ambiguità in figura 3.10.

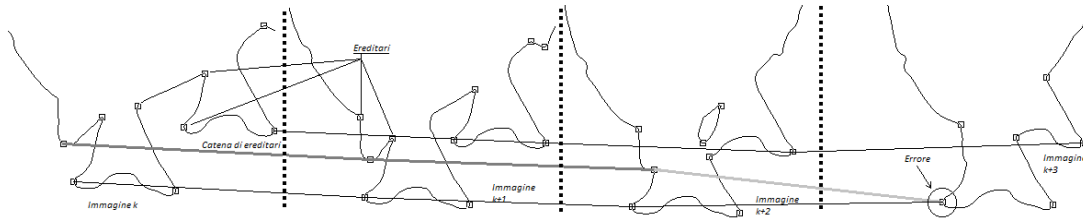


Figura 3.10: Un errore nella ricostruzione della catena. Nei primi due passi la catena, in grigio scuro, è ricostruita correttamente, ma nel passaggio dall'immagine $k+2$ alla $k+3$ la feature seguita scompare (si confonde con la maschera dell'oggetto), tuttavia l'algoritmo prosegue scegliendo ancora il più vicino (il tratto in grigio chiaro), unendo erroneamente due catene separate.

La soluzione proposta in questa tesi si basa sulla considerazione che i punti di una catena sono separati l'uno dall'altro al di sotto di una soglia Δ_{th} , quindi è sensato supporre che se f_1 dista da f_2 maggiormente di quella quantità, allora le due features vanno considerate separatamente; empiricamente tale valore Δ_{th} è stato fissato pari a 16 pixel.

Con questa cernita si garantisce la possibilità di separare le catene confluenti, in modo da eliminare il maggior numero possibile di ambiguità. L'ultimo punto,

quello delle catene orfane, è risolto, come si vedrà nei prossimi paragrafi, ignorando temporaneamente gli ereditari in eccesso per poi riprenderli in considerazione in una successiva iterazione dell'euristica. Il fatto di recuperare tutte le catene è inoltre utile per una verifica a posteriori dell'elaborazione: qualora la scelta della catena da parte dell'algoritmo non dovesse avere un buon esito ai fini della ricostruzione, confrontando le varie catene di ereditari è possibile, entro certi limiti, correggere i risultati a posteriori, senza essere costretti a ripetere le scansioni e l'elaborazione, semplicemente scegliendo un'altra catena.

Come ultimo passo per la determinazione della matrice fondamentale, è necessario scegliere tra quelle disponibili le catene che meglio si adattano all'elaborazione per l'*8-point algorithm*, descritto nella sezione seguente. L'algoritmo richiede un numero minimo, otto, di features che relazionano tra loro le immagini: poiché per costruzione le catene contengono esattamente le features necessarie all'algoritmo, a questo punto è sufficiente selezionare quelle di lunghezza pari o superiore al necessario per poter procedere con il calcolo.

Algoritmo di cernita

La descrizione delle varie fasi dell'algoritmo è la seguente:

1. Estrazione tramite FCPDA delle features, o corner, di tutte le immagini;
2. Creazione della matrice $dist_{i,i+1}$ relativa a ogni coppia di immagini consecutive;
3. Tramite elaborazioni matematiche iterative, estrazione dell'ereditario dell'immagine successiva più vicino a ciascuna feature di quella corrente;
4. Costruzione delle catene secondo questo schema:
 - (a) Marcatura di tutti i punti di tutte le immagini a -1;
 - (b) Per ogni elemento dell'immagine costruzione della catena corrispondente e marcatura degli ereditari visitati a 1;
 - (c) Una volta terminata l'immagine corrente, passaggio a quella successiva;
 - (d) Per ogni elemento non marcato, costruzione della catena e marcatura fintantoché si incontrano punti non marcati, in caso contrario procedere con il prossimo punto iniziale;
 - (e) Ripetizione dei vari passaggi fino all'esaurimento dei punti.
5. Filtraggio e separazione delle catene;
6. Scelta catene idonee.

3.3.3 Normalized 8-point algorithm

Una volta determinate le features delle varie immagini e le catene idonee corrispondenti, è possibile procedere alla fase successiva, cioè alla valutazione delle trasformazioni in atto nel sistema. Quello che viene presentato è l'algoritmo più semplice per l'estrazione della matrice fondamentale, [21] che consiste nella costruzione e soluzione di un sistema lineare di equazioni create ad hoc. Pur essendo estremamente semplice, garantisce risultati eccellenti, a patto di porre molta cura nel pretrattamento dei dati in ingresso, che vanno normalizzati adeguatamente (di qui il Normalized del nome). L'essenza dell'algoritmo a 8 punti, nella sua variante normalizzata, segue questi passi:

1. Normalizzazione delle coordinate di entrambe le immagini rappresentabile secondo le trasformazioni $\hat{x}_i = T x_i$ e $\hat{x}'_i = T' x'_i$;
2. Calcolo della matrice fondamentale F secondo
 - (a) Soluzione del sistema lineare $Af = 0$ definito in 3.5 (ricavo di \hat{F});
 - (b) Sostituzione della \hat{F} con F_{SVD} in modo che $\det(F_{SVD})=0$;
3. Denormalizzazione della matrice F_{SVD} secondo la $F = T'^T F_{SVD} T$.

La F così trovata è la matrice fondamentale relativa alle coordinate originali x_i e x'_i .

La trasformazione di normalizzazione scelta dev'essere tale da garantire che la distanza media tra un punto generico e l'origine sia pari a $\sqrt{2}$, di modo che ciascun punto sia il più prossimo possibile a quello di coordinate $(1, 1, 1)^T$. Tale normalizzazione è necessaria per evitare di dover confrontare, al momento della soluzione del sistema 3.5, valori nell'ordine di 10^4 con valori unitari¹, confronto che porterebbe a variare di pochissimo alcuni valori e moltissimo altri, con conseguenze devastanti sull'errore finale all'atto della sostituzione di A con la sua SVD [22]. Per questo motivo la normalizzazione o pre-condizionamento dei dati è un passo fondamentale e inevitabile nell'algoritmo a 8 punti.

¹Situazione tutt'altro che ipotetica: supponendo di avere due punti di coordinate $(100,100,1)$ e $(100,100,1)$, all'interno di A si avrebbero esattamente i valori citati. Inoltre nel caso preso in esame da questa tesi, le coordinate spaziano all'interno di immagini 1000×1100 , potenzialmente quindi si confronterebbero valori nell'ordine di 10^6 con valori unitari.

Capitolo 4

Fase Sperimentale

L'esperimento condotto in questa tesi consiste nell'implementazione X della ricostruzione SfS di un soprammobile in argento sfruttando, per acquisire le immagini richieste dalla tecnica, la radiazione X. A causa della pericolosità della radiazione impiegata, le acquisizioni sono state fatte all'interno di un bunker schermato che ha richiesto alcuni accorgimenti: innanzitutto la sorgente è stata controllata dall'esterno del bunker tramite l'apposita unità di controllo; per poter rivelare le silhouette è stato impiegato un sensore collegato a un elaboratore, anche questo all'esterno del bunker; infine si è posto il bersaglio su un rototraslatore controllato da remoto, in modo da poterlo ruotare ottenendo così diverse silhouette, con la condizione di acquisirle separandole di un angolo costante.

4.1 Strumenti

In figura 4.1 è possibile vedere il sistema utilizzato per questa tesi, costituito principalmente da tre componenti: la sorgente di radiazione a sinistra, il rototraslatore al centro e il rivelatore X a destra.¹ La zona del rototraslatore è anche dove sono stati posti gli oggetti bersaglio di questa tesi, cioè un soprammobile in argento, per la precisione un galeone, e una scatola di cartone utilizzata per l'estrazione dei parametri intrinseci del sistema, detta *calibration box*.

La sorgente a disposizione è la *Hamamatsu L8601-01*, trattasi di una sorgente a microfocus con finestra in berillio controllata in potenza con i seguenti dati

¹Questa disposizione rende evidente il motivo per cui il modello pinhole originale è stato adattato come scritto nella sezione 3.2.

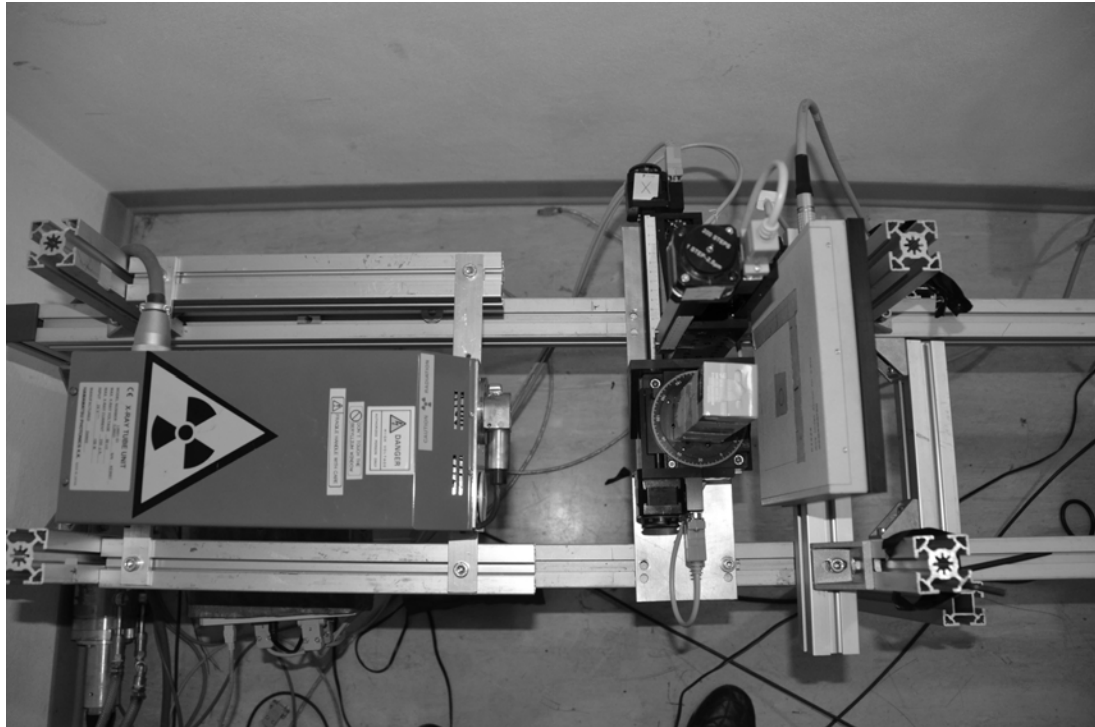


Figura 4.1: Vista dall'alto del sistema sperimentale utilizzato.

tecniche [23]:

- diametro dello spot compreso tra $7\mu\text{m}$ a 10W e $5\mu\text{m}$ a 4W ;
- angolo della radiazione emessa pari a 39° ;
- range operativi di $20\text{-}90\text{ kV}$ e $0\text{-}250\ \mu\text{A}$ vincolati entro 10W per preservare la sorgente stessa dal degrado dovuto a surriscaldamento

Per gli scopi di questa tesi si è scelto di utilizzare la sorgente alla massima potenza possibile per ridurre ambiguità dovute al supporto e alla scatola nelle foto, rispettivamente, del galeone e della calibration box.

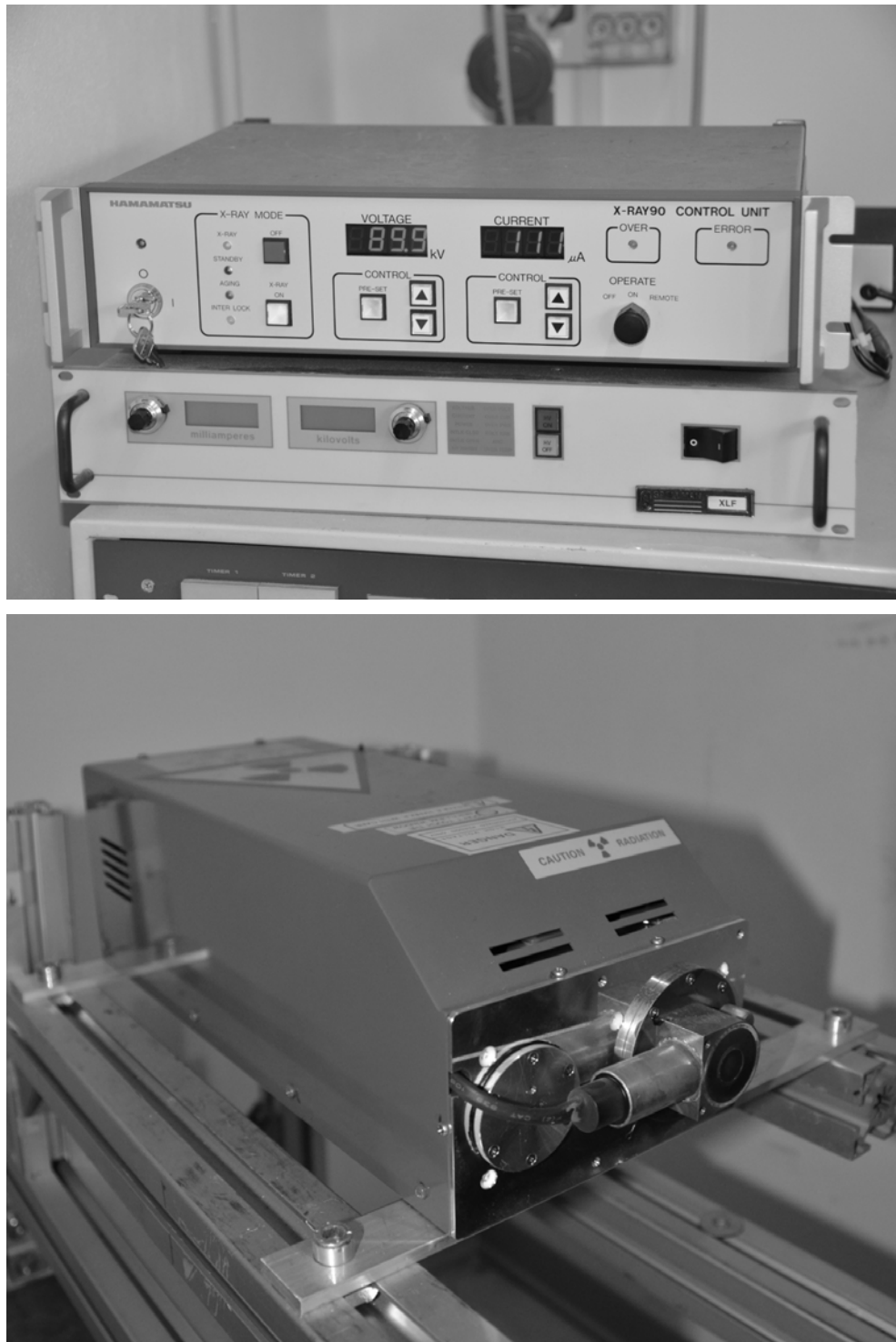


Figura 4.2: Pannello di controllo e lato della finestra d'uscita della sorgente Hamamatsu L8601-01.

4. FASE SPERIMENTALE

Nonostante l'algoritmo Sfs preveda di raccogliere una serie di immagini attorno all'oggetto, per la peculiare natura del sistema¹ è stato necessario ruotare l'oggetto tenendo fissi sorgente e rivelatore. La rotazione è stata affidata a un sistema Standa di movimento motorizzato a tre assi (orizzontale, verticale e rotazionale) composto da due attuatori lineari 8MT175, usati per centrare l'oggetto in modo tale che rimanesse sempre all'interno dell'area sensibile del rivelatore, e dal rotatore 8MR191, capace di effettuare rotazioni complete con una precisione pari a 0.01° a un passo massimo di 8 rotazioni al minuto.

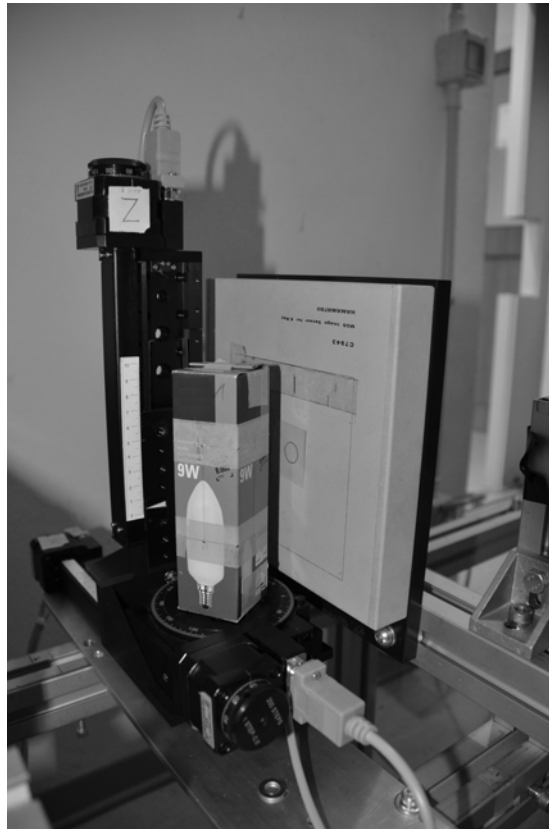


Figura 4.3: Il rototraslatore Standa con la calibration box e il rivelatore Hamamatsu C7943

Infine, le immagini sono state raccolte dal rivelatore flat panel Hamamatsu C7943, costituito da una scheda sensori e una di controllo, inserite in un case in alluminio spesso un millimetro a scopo protettivo. La scheda sensori è composta da uno scintillatore in CsI e da una matrice bidimensionale di fotodiodi

¹Che da un lato richiede un ambiente protetto per poter eseguire l'esperimento, dall'altro presenta l'impossibilità di muovere la sorgente.

da 1248x1248 pixel, specialmente realizzati per un'alta sensibilità verso la luce verde emessa dallo scintillatore all'assorbimento di raggi X. Inoltre sono presenti 8 charge sensitive amplifier array, ciascuno dotato di 156 amplificatori con un registro orizzontale. Il rivelatore offre un frame rate di 7 fps con range dinamico di 12 bit e un'area attiva massima di 124.8 mm x 124.8 mm [26]. Il pixel size è di $100\mu m$, tuttavia con questo sensore è possibile raggiungere risoluzioni nell'ordine di $10\mu m$, tramite tecniche di contrasto di fase. [24]

A fronte di un'area massima disponibile pari, in pixel, a 1216 x 1232, si è ridotta la dimensione di acquisizione a immagini di larghezza 1000 pixel e altezza 1100.

4.2 Target

L'esperimento si divide in tre fasi principali, due iniziali di acquisizione e una, finale, di elaborazione. Sono state compiute due serie di acquisizioni in quanto una necessaria per scansire l'oggetto da ricostruire, un galeone in argento di circa 6 x 8 cm (scelto perché dotato di una struttura complessa e molti dettagli) e una per acquisire le immagini necessarie all'estrazione dei parametri intrinseci del sistema.

Per ottenere la prima serie di immagini, quelle del galeone, è stato inizialmente posizionato l'oggetto al centro del rotatore, poi accesa la sorgente e quindi salvato quanto rivelato dal sensore a passi di 5 gradi lungo un giro completo, ottenendo così 73 immagini, con l'ultima uguale alla prima. Lo stesso si è fatto per la seconda serie di acquisizioni: quest'ultima consiste nelle proiezioni di una calibration box (cioè una scatola di cartoncino dotata di spilli situati a distanze conosciute) fatta ruotare a passo costante di 5 gradi.

Il confronto tra il poliedro estratto dalle teste degli spilli, o *tiepoints*, nelle proiezioni e quello corrispondente a realtà permette di risalire alla distanza focale e al centro ottico del sistema, ovvero ai parametri intrinseci, richiesti nella successiva elaborazione dei dati.

4.2.1 Modello Matematico: *Calibration Box*

La calibration box (da qui in avanti CB) utilizzata ha come unico scopo quello di ricavare i parametri intrinseci del sistema di imaging adottato. Per raggiungere tale obiettivo si è pensato di dotare la CB di due serie da 8 spilli, in modo da ottenere un poliedro di dimensioni note che ruoti con il sistema di riferimento. Le dimensioni della CB sono tali che gli spilli o più precisamente le capocchie (dette anche *tie-points*) rimangano all'interno dell'area di acquisizione per tutti gli scatti effettuati.

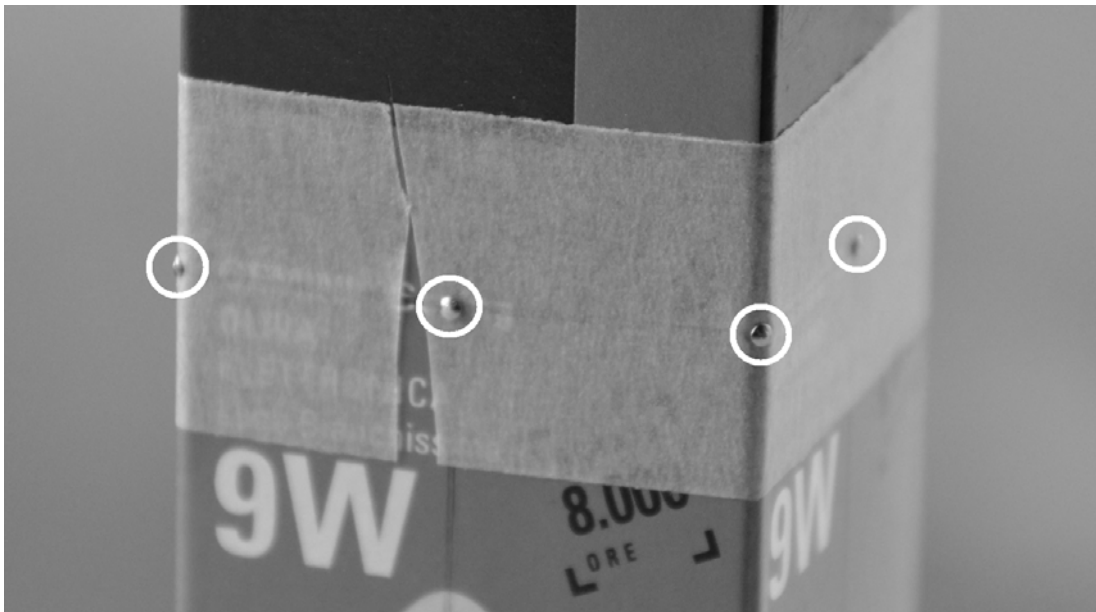


Figura 4.4: Particolare della calibration box, in evidenza le teste degli spilli, corrispondenti ai tiepoints.

Una volta posizionata la CB e acquisite le immagini, viene effettuata una segmentazione in modo da isolare gli spilli; successivamente si identificano i centri delle capocchie sfruttando il fatto che la loro forma è quasi ellittica e infine un semplice controllo sulle coordinate divide le due serie in alta e bassa. A questo punto l'elaborazione si fa più delicata: per poter relazionare gli spilli da un'immagine alla seguente è necessario sia trovare la posizione dello stesso spillo in due immagini consecutive che determinarne la posizione relativamente agli altri. Un filtro di Kalman sulle coordinate provvede a mantenere l'identificazione dello spillo mentre un algoritmo semiautomatico verifica che la numerazione segua la rotazione, in quanto lo spillo i -esimo può trovarsi a destra o sinistra di quello

$i + 1$ -esimo a seconda dell'angolo di ripresa, quindi è necessario verificare che non ci siano scambi di posizione.

Una volta trovate e assegnate le teste, queste vengono divise in due gruppi, $B_{up} = \{\mathbf{x}_{up,1}; \mathbf{x}_{up,2}; \dots; \mathbf{x}_{up,n}\}$ e $B_{dw} = \{\mathbf{x}_{dw,1}; \mathbf{x}_{dw,2}; \dots; \mathbf{x}_{dw,n}\}$ ¹, dove $\mathbf{x}_{,i}$ rappresenta la coppia di coordinate del tie-point. A questo punto si suddivide ulteriormente ogni sequenza a seconda della dimensione di B_{dw} ,² iniziando da quelle complete (la maggior parte). Poiché la distanza tra due teste consecutive è breve, ha senso ipotizzare che localmente seguano un'orbita ellittica, che si può sfruttare per integrare la posizione mancante nell'immagine successiva. Iterando il processo si arriva al punto in cui tutte le sequenze sono complete e, grazie alla forte ridondanza, con un errore trascurabile.

Come ultimo passo nell'ordinamento si sfrutta la topologia del sistema: siccome vale la relazione

$$x_{ik} = P_k \begin{bmatrix} X_i \\ 1 \end{bmatrix}$$

(dove x_{ik} sono le coordinate dell' i -esimo spillo nella k -esima immagine, P_k è la matrice di proiezione e X_i rappresenta le coordinate 3D dell' i -esimo spillo nel sistema di riferimento noto) si può definire una matrice L_i così composta:

$$L_i = \begin{bmatrix} X_i^\top & 1 & \mathbf{0}_{1 \times 4} & -x_{ik} \begin{bmatrix} X^\top & 1 \end{bmatrix} \\ \mathbf{0}_{1 \times 4} & X_i^\top & 1 & \end{bmatrix}$$

La sovrapposizione delle varie L_i permette di ricavare la matrice globale \mathbf{L} , che a sua volta permette di determinare la matrice proiettiva P_k , ricavabile come il minor autovettore della matrice $L^\top L$.

¹Si noti come n possa essere diverso da 8 a causa di sovrapposizioni.

²Le medesime considerazioni valgono per B_{up} .

4.2.2 Modello Matematico: Procedura

Una volta disponibili le coordinate 3D \mathbf{X} della calibration box e quelle ad esse associate per ogni immagine J_i , per ciascuna di queste ultime è possibile estrarre la matrice P_i tale che

$$J_i \propto P_i X_i \text{ con } P_i = KR \begin{bmatrix} I & C_i \end{bmatrix}$$

Tuttavia le matrici così trovate non sono corrette, in quanto la matrice K dei parametri intrinseci non è vincolata ad essere unica per tutte le immagini. Per poter imporre tale condizione, si compiono le seguenti ipotesi: le coordinate della calibration box devono innanzitutto riportarsi tramite una rototraslazione $[R_b, t_b]$ nel sistema di riferimento del rotatore; in tale sistema quindi, all'acquisizione i -esima, esse sono ruotate lungo l'asse verticale di un angolo costante, rotazione espressa dalla matrice R_a ; infine è necessario riportarsi tramite un'ulteriore rototraslazione $[R_s, t_s]$ nel sistema di riferimento della camera, nel quale prende luogo l'effettiva proiezione. In figura 4.5 sono visibili le tre trasformazioni $[R_b, t_b]$, R_a e $[R_s, t_s]$.

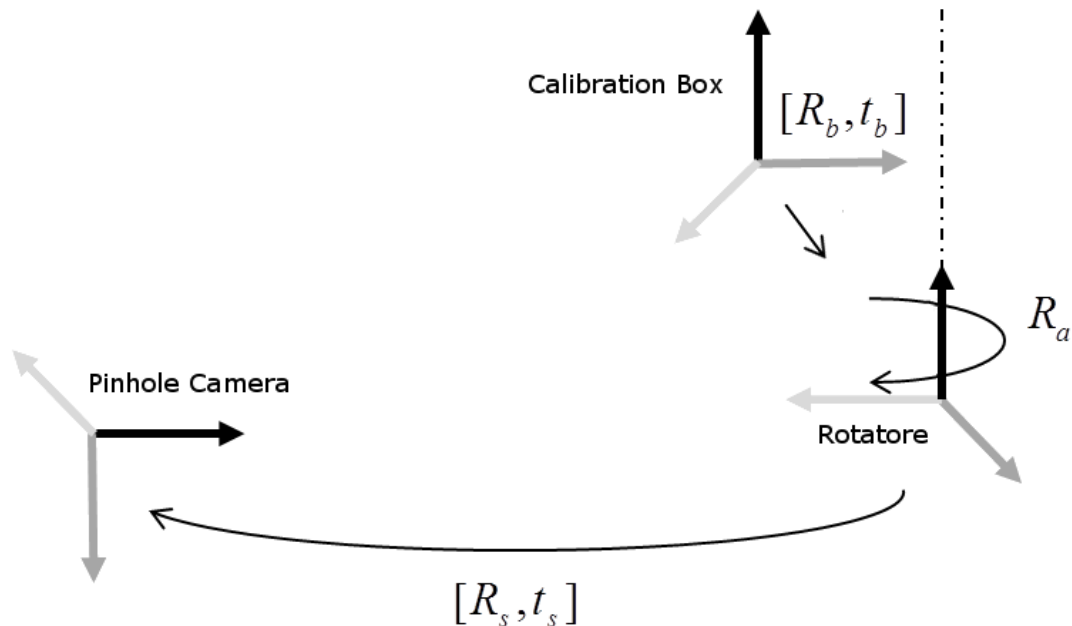


Figura 4.5: Il modello considerato nell'esperimento.

Il modello considerato assume quindi la seguente forma:¹

$$P_i = K [R_s R_a^{i-1} (R_b \mathbf{X} + t_b) + t_s]$$

A questo punto si ha che, secondo il modello, vale sia $R_i = R_s R_a^{i-1} R_b$ che

$$C_i = R_b' t_b + R_b' R_a^{i-1} R_s' t_s. \quad (4.1)$$

Con queste relazioni si arriva al risultato cercato, cioè

$$P_i = K R_i \begin{bmatrix} I & C_i \end{bmatrix}$$

Rumore di acquisizione

L'acquisizione dei tie-points è soggetta a rumore, per ridurre il quale è sufficiente analizzare ogni tie-point della calibration box assieme al relativo andamento nelle acquisizioni: essendo questi punti in rotazione attorno ad un asse, la loro proiezione rappresenta un'omotetia.² Quindi si può scrivere per ciascuno di essi la seguente relazione:

$$J_i \propto H X_i$$

dove H è una matrice 3x3. Tale espressione permette di ridurre il rumore di fondo, ottenendo così i parametri mostrati in figura 4.6, dove sono presenti per ogni immagine delle 72 considerate³ sia i parametri intrinseci f_x, f_y, c_x, c_y, s_k che quelli estrinseci C_i, R_i .

Vincolo di acquisizione circolare

Come visibile in figura 4.6 i parametri intrinseci risultano particolarmente variabili durante l'acquisizione, ragion per cui è necessario imporre dei vincoli. Essendo la meno rumorosa tra le disponibili, la posizione dei centri camera è indubbiamente quella che più si presta a fornire il vincolo richiesto: poiché le acquisizioni sono

¹Si noti come l'indice i , relativo alla singola acquisizione, compaia solo all'esponente della matrice R_a .

²Trattasi di trasformazione geometrica che dilata e/o contrae le distanze mantenendo invariati gli angoli o gli spigoli.

³Le acquisizioni totali sono 73, ma la prima e l'ultima sono identiche.

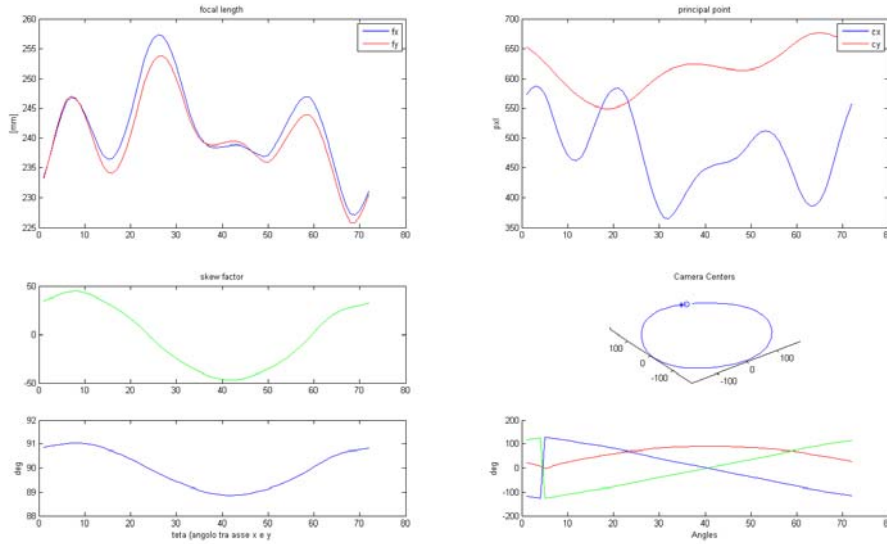


Figura 4.6: La matrice K dopo l'analisi a punto singolo.

avvenute attraverso una rotazione costante intorno ad un asse, i centri devono necessariamente trovarsi su un cerchio. Questa condizione rappresenta il nuovo vincolo, evidente anche dall'equazione 4.1.

È quindi lecito vincolare i C_i a trovarsi su un cerchio perfetto, identificando in tal modo univocamente l'asse del rotatore (ed eliminando così la dipendenza da R_b) per poi stimare nuovamente gli altri parametri, ottenendo i risultati illustrati in figura 4.7.

Rotazione

Con il precedente vincolo si è eliminata la componente sull'asse Z del rotatore, ottenendo quindi che

$$R'_i R_{i+1} = R_a$$

Utilizzando la scomposizione di Rodrigues, si nota in figura 4.8 che tale valore risulta sufficientemente stabile. In particolare la matrice stimata ha rotazione nulla su due assi e attorno ai 5 gradi sugli altri, in accordo con le previsioni.

A questo punto è possibile vincolare tale rotazione sostituendo a X l'espressione $R_a^{i-1}(X - C_i)$. Questo permette di mantenere costante la rotazione residua (definita da R_s), come illustrato in figura 4.9.

4. FASE SPERIMENTALE

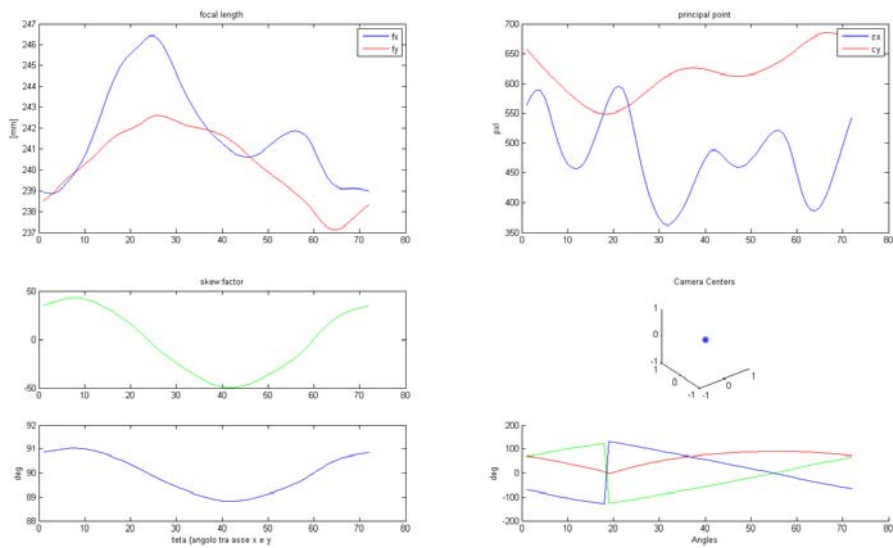


Figura 4.7: Stima della Matrice K e dei centri C.

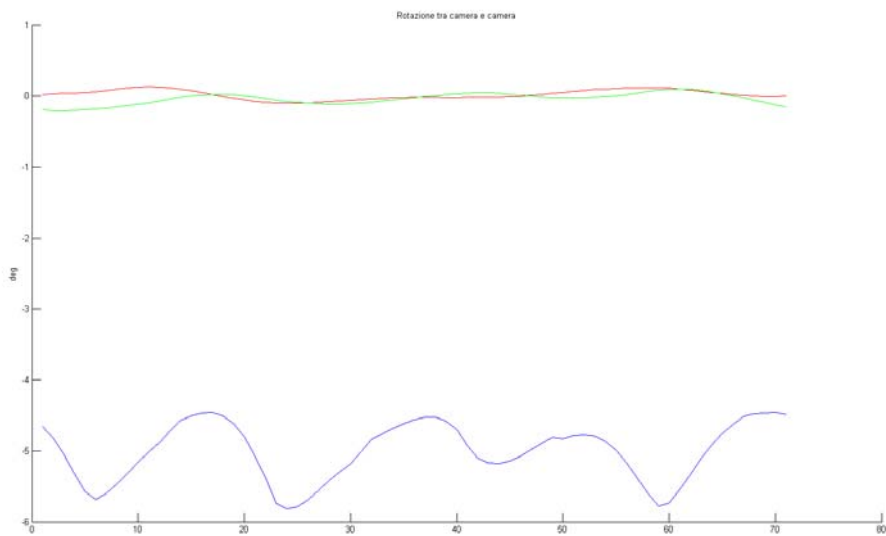


Figura 4.8: Rotazione tra un'acquisizione e l'altra.

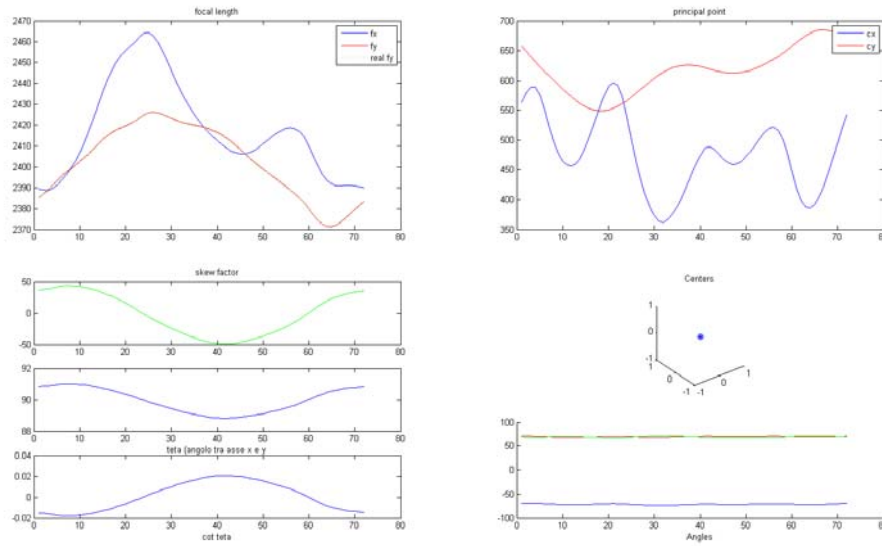


Figura 4.9: Stima delle matrici K , R_a e dei centri C .

È possibile imporre ulteriormente tale rotazione, ottenendo così un ulteriore miglioramento della stima, come visibile in figura 4.10.

Skew factor e correzione finale

In questa sede è possibile agire anche sull'ultimo parametro, lo skew factor, imponendolo nullo. Tuttavia gli errori residui sulla focale e sul centro proiettivo rimangono presenti, ma considerato il loro andamento armonico è evidente che siano dovuti non a rumore ma a imprecisioni della valutazione delle dimensioni della calibration box. Pertanto è possibile svolgere un procedimento a ritroso per correggere le coordinate della posizione “reale” delle teste degli spilli fino a massimizzare la convergenza, come visibile in figura 4.11.

4. FASE SPERIMENTALE

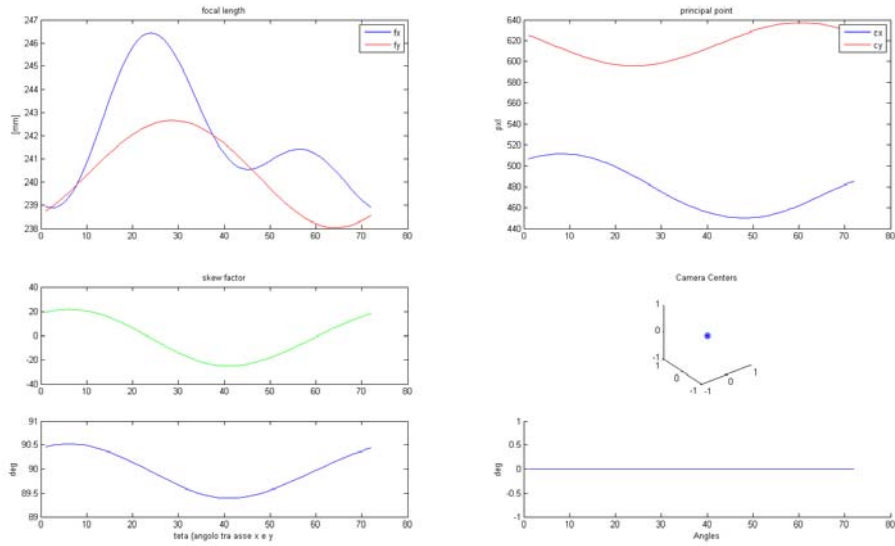


Figura 4.10: Stima delle matrici K , R_a , R_s e dei centri C .

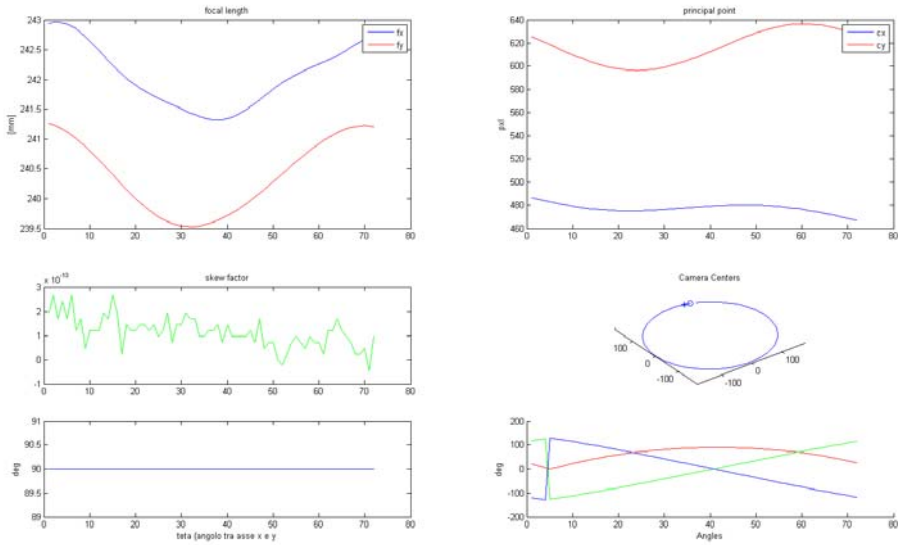


Figura 4.11: Stima di K' con le posizioni dei tiepoints corrette.

4.3 Applicazioni Precedenti

La tecnica SfS utilizzata in questa tesi è già stata applicata con successo nella ricostruzione di due oggetti monolitici, con variazioni minime nella posizione di sensore e rotatore. Tuttavia c'è una differenza sostanziale con quanto compiuto in questa tesi: in entrambe le occasioni la calibration box veniva utilizzata sia per ricavare i parametri intrinseci del sistema, come anche in quest'ultimo esperimento, sia per ottenere una metrica attraverso la quale calcolare il modello 3D. Inoltre gli angoli di rotazione degli oggetti erano noti a priori.

Nel caso in esame in questa tesi, invece, la calibration box è stata acquisita separatamente al bersaglio, e si sono utilizzate le features rilevate per calcolare le catene di ereditari precedentemente introdotte ai fini di ricavare il valore della rotazione, che aveva come unico vincolo quello di essere costante.

Il primo utilizzo sperimentale della tecnica è visibile in figura 4.12, l'oggetto ricostruito è una suora di peltro [27].



Figura 4.12: Ricostruzione SfS di una suora di peltro, a sinistra una foto della suora, a destra due lati del modello 3D.

A fronte di un volume misurato di $36.2 \pm 0.2 \text{ cm}^3$, la ricostruzione 3D ha riportato un volume 36.5 cm^3 , pienamente accettabile per gli scopi di questa tecnica. In questa foto si nota inoltre come le concavità dell'oggetto vengano in gran parte perse, tuttavia, nonostante questo, è ancora pienamente possibile dedurre la

forma della suora nei suoi dettagli più importanti (la posizione del viso e del velo e il fatto che “in mano” abbia un oggetto vagamente rettangolare).

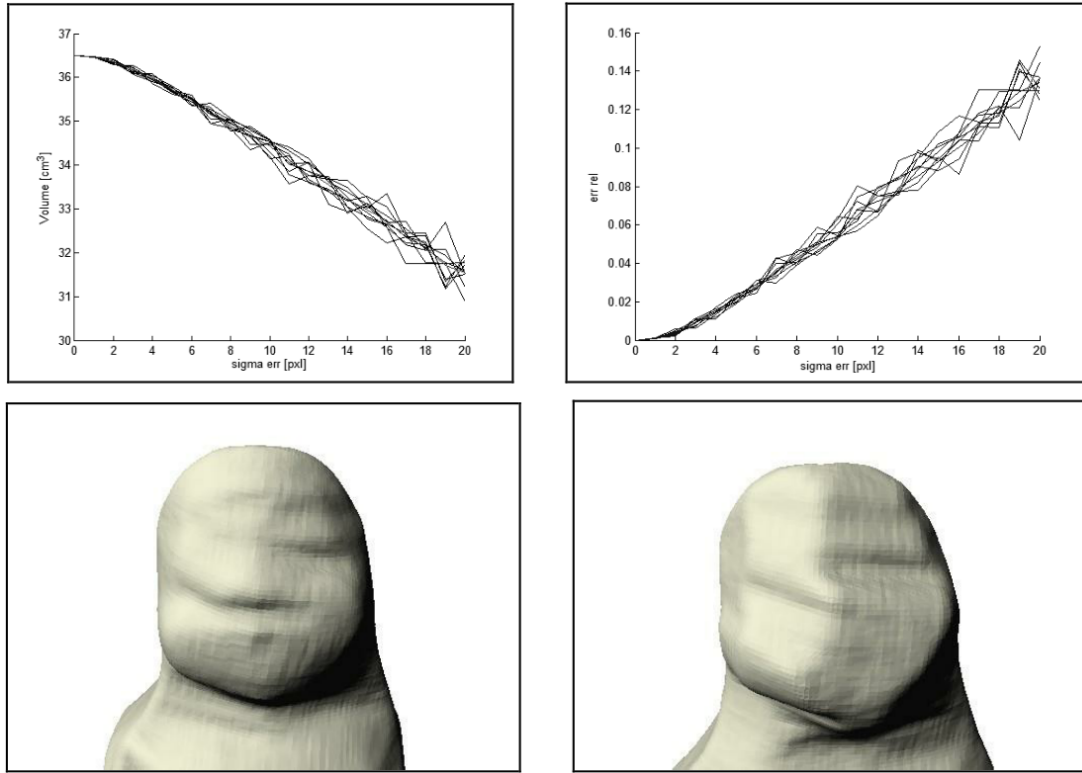


Figura 4.13: Robustezza dell’algoritmo di ricostruzione, in senso orario partendo dall’alto a sinistra: variazione del volume calcolato in funzione dell’errore introdotto nella posizione dei tiepoints; errore relativo sull’effettivo valore del volume; viso della suora prima dell’introduzione del rumore (sono visibili occhi e bocca); viso della suora in seguito all’introduzione di rumore con varianza 20 pixel sulle coordinate dei tiepoints.

Interessante è osservare il comportamento dell’algoritmo rispetto al rumore, illustrato in figura 4.13. Poiché la dimensione degli spilli in pixel risultava al massimo 50 pixel, si è supposto un errore di al più ± 30 pixel (cui corrisponde una deviazione standard σ pari a 20 pixel), caso in cui comunque la differenza rispetto al volume effettivo si è mantenuta sotto il 15%. L’andamento monotonic del volume in presenza di errore è spiegato dalla perdita di coerenza tra le silhouette, a sua volta dovuto all’addittività dell’errore nel posizionamento dei tiepoints.

Altro esempio di questa tecnica è stata la sua applicazione in ambito archeologico, [28] visibile in figura 4.14. L'oggetto da ricostruire consisteva in un manufatto paleoveneto metallico completamente circondato da roccia. Poiché il metallo assorbe totalmente i raggi X e la roccia no, è stato possibile utilizzare la tecnica SfS senza particolari accorgimenti, in quanto la silhouette dell'oggetto risultava comunque ben distinguibile dalla roccia che lo racchiudeva.

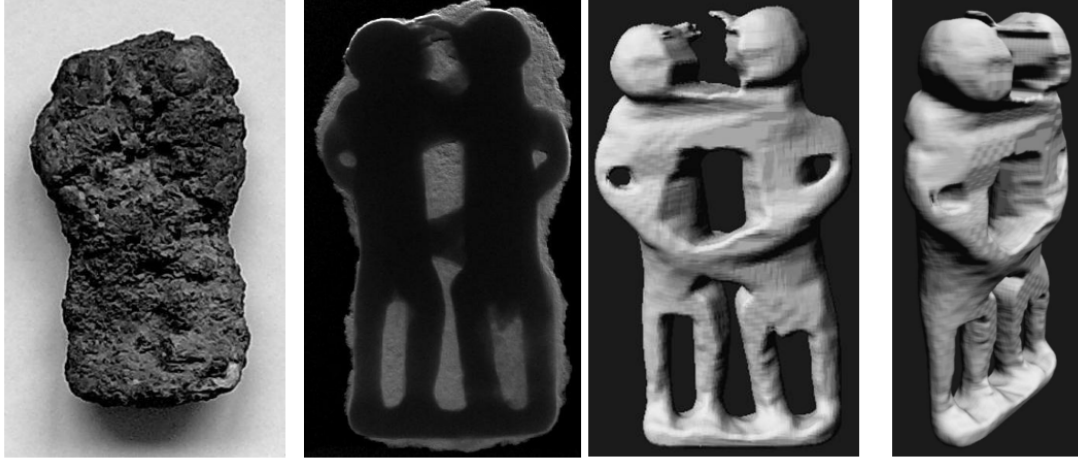


Figura 4.14: Un manufatto paleoveneto ricostruito tramite SfS. La prima immagine evidenzia come ad occhio nudo il manufatto fosse invisibile. La seconda immagine mostra una radiografia dell'oggetto e le ultime due mostrano il modello 3D ricostruito.

L'archeologia è il campo dove questa particolare tecnica mostra meglio i suoi vantaggi: la tomografia offre sì una risoluzione estremamente superiore, ma la velocità e la praticità di utilizzo del setup SfS sono molto migliori, inoltre, come visibile in figura 4.14, il dettaglio raggiunto è soddisfacente e sufficiente se non altro a dare indicazioni operative a chi dovrà eseguire un eventuale recupero del manufatto dal blocco di roccia in cui si trova.

Capitolo 5

Ricostruzione

Come già asserito nei capitoli precedenti, lo scopo di questa tesi è stimare automaticamente l'angolo di rotazione dell'oggetto direttamente dalle immagini che lo ritraggono. La stima di tale angolo, tuttavia, non è fine a sé stessa, ma è essenziale alla ricostruzione 3D finale, punto di arrivo e di verifica dell'intero esperimento.

5.1 Galeone

Il bersaglio scelto per questa tesi doveva presentare, per far risaltare i punti di forza e di debolezza dell'algorithmo implementato, le seguenti caratteristiche: innanzitutto doveva essere di dimensioni ridotte, in accordo con l'area sensibile del rivelatore; poi doveva possedere una struttura relativamente complessa ovvero con vari livelli di dettaglio, in modo da evidenziare direttamente la risoluzione ottenibile dal sistema; infine si è cercato un oggetto la cui struttura presentasse concavità, convessità e parti forate, così da mettere alla prova le prestazioni dello SfS.

Un oggetto che risponde a tutti i requisiti è quello visibile in figura 5.1.

Trattasi, come scritto in precedenza, di un galeone in argento di circa 6x8 cm, posto su un supporto plastico trasparente ai raggi X per poter essere centrato sul rivelatore. Una volta acquisite le sue silhouette, gli ereditari sono stati scelti in base alla lunghezza delle catene loro corrispondenti: qualche esempio sono le punte delle tre bandiere, gli spigoli inferiori delle vele e i piedini del supporto. Un esempio di silhouette acquisita e della sua elaborazione è visibile in figura 5.2.

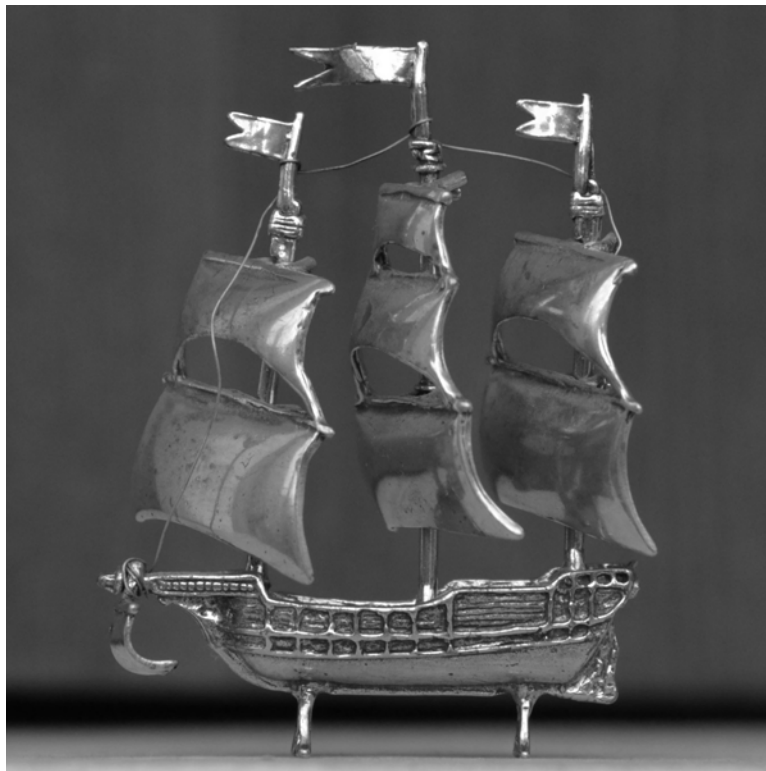
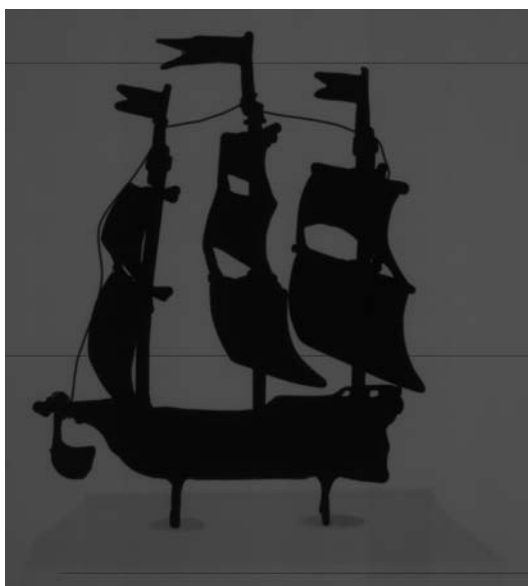


Figura 5.1: Il galeone utilizzato in questa tesi come bersaglio.



(a) Una delle 73 silhouette del galeone acquisite per la ricostruzione 3D.



(b) Le features della stessa silhouette individuate dal detector FCPDA.

Figura 5.2: Elaborazione delle silhouette.

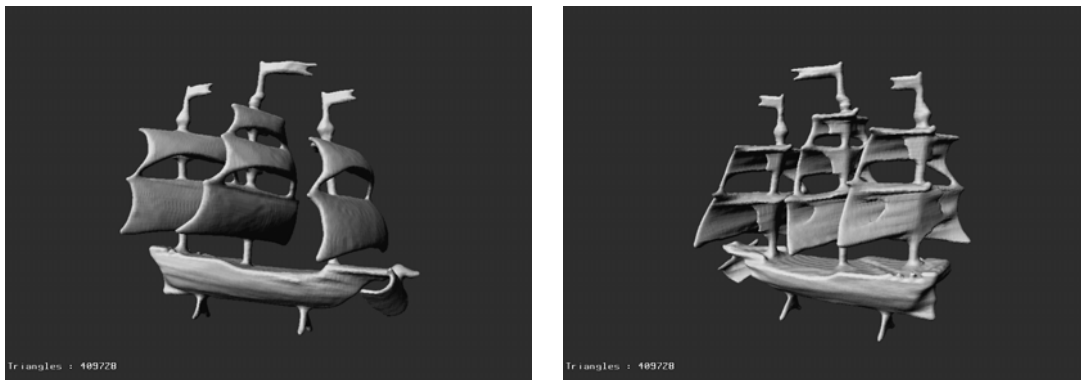
Dalla figura 5.2(a) si evince inoltre come il sensore presentasse tre linee difettose (le righe nere), che hanno reso necessario un pre-trattamento manuale delle immagini acquisite. La zona più scura alla base del galeone è la proiezione della scatola in plastica utilizzata come supporto, mentre le due chiazze scure sotto i piedini sono dovute a della plastilina utilizzata per assicurare il galeone al supporto durante le rotazioni. La presenza di queste ombre non ha creato problemi alla segmentazione delle immagini.

In figura 5.2(b) sono visibili le features così come sono state rilevate dal detector FCPDA, prima di essere sottoposte a cernita per identificare gli ereditari. La posizione di ciascuna di esse è identificata dal centro del relativo quadrato. Si noti inoltre come la riga attraversante le vele introduca una separazione che non esiste in realtà, generando così dei falsi positivi.

I passi successivi dell'algoritmo avvengono compiendo iterativamente operazioni su matrici, arrivando con successo alla ricostruzione 3D finale visibile in figura 5.3.

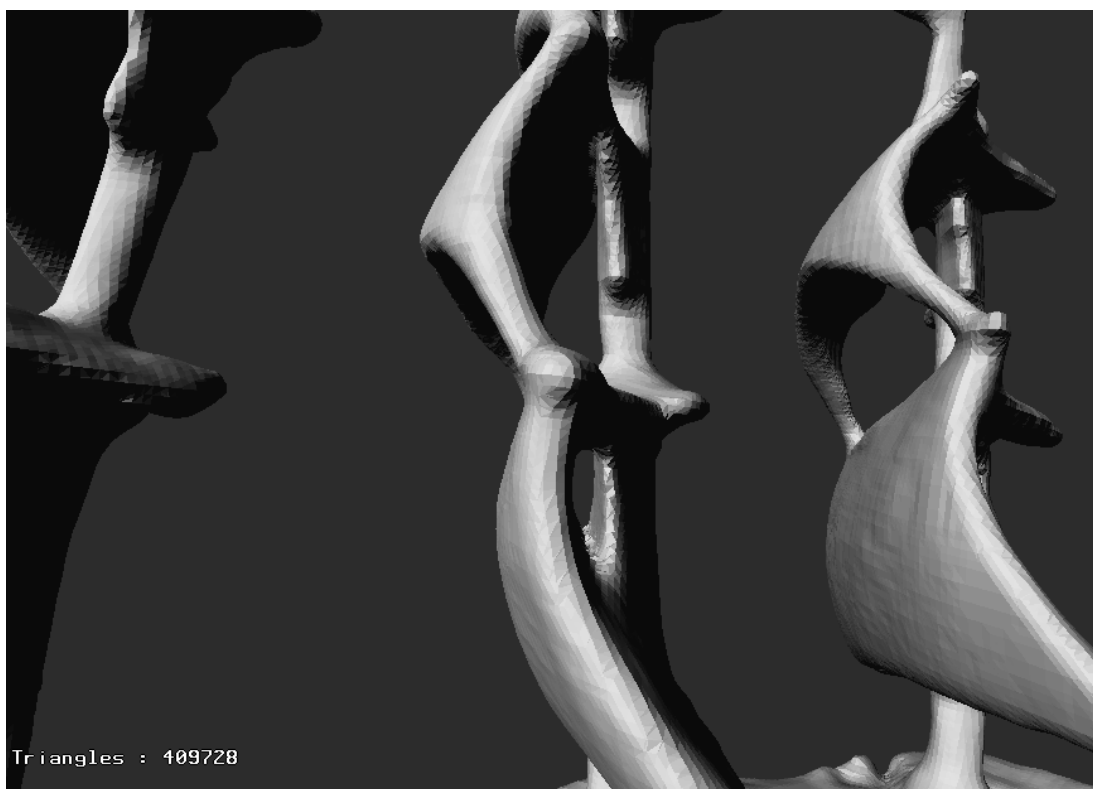
Il modello 3D mostra chiaramente le caratteristiche principale del galeone: sono visibili infatti tutte le singole vele, con anche le parti separate dai rispettivi alberi, le bandiere mostrano un accenno di curvatura (come si vede in figura 5.1, sono lavorate come se sventolassero) e la parte del timone è ben ricostruita.

Purtroppo però sono presenti anche difetti: la trama dello scafo non è visibile, ma soprattutto l'algoritmo non riesce a visualizzare il filo d'argento che collega le vele tra di loro e successivamente al pennone di prua.



(a) Vista laterale destra del modello 3D ricostruito.

(b) Vista laterale sinistra del modello 3D ricostruito.

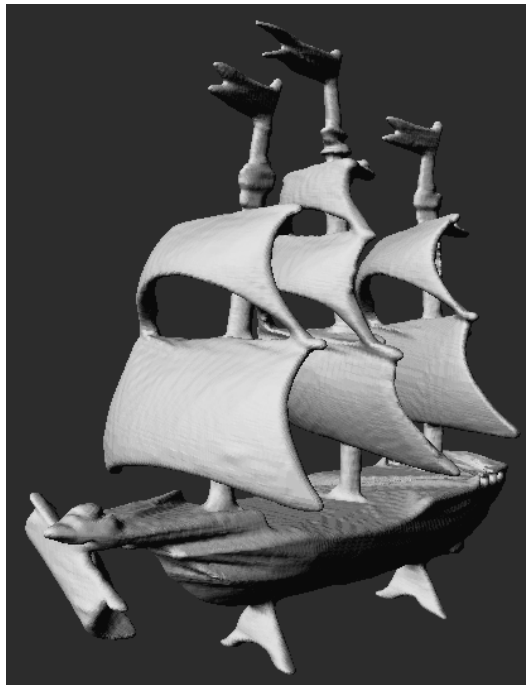


(c) Particolare della vela centrale.

Figura 5.3: Ricostruzione 3D finale.



(a) Visione 3/4 anteriore del galeone.



(b) Visione 3/4 anteriore della ricostruzione 3D.

Figura 5.4: Confronto tra il galeone originale e la sua ricostruzione.

Capitolo 6

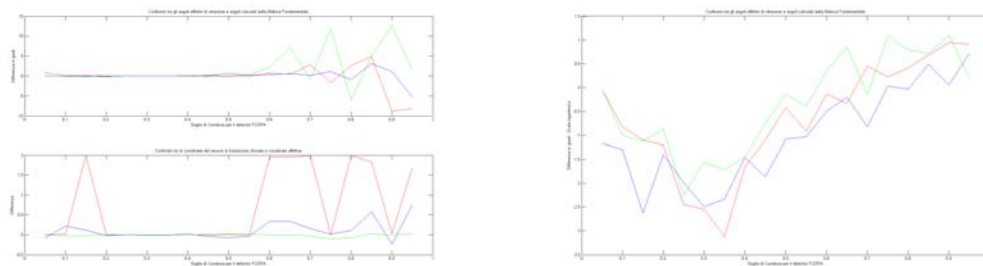
Conclusioni

In questa tesi ci si è posti l'obiettivo di estrarre i parametri estrinseci del sistema di imaging necessari alla ricostruzione 3D.

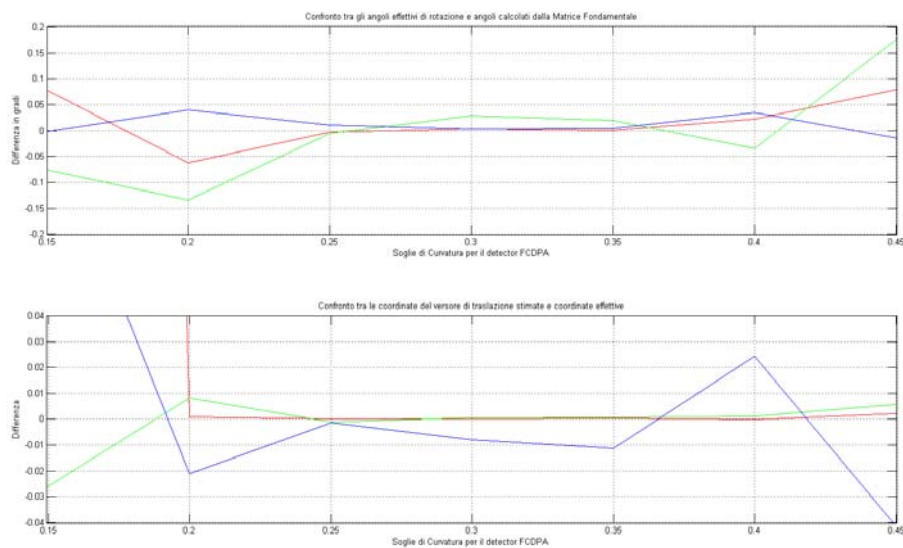
Tale obiettivo, come dimostrato dalla figura 5.3, è stato raggiunto con successo; è tuttavia necessario fare alcune annotazioni: innanzitutto la ricostruzione è sì corretta per le caratteristiche macroscopiche, ma per i dettagli più minuti la tecnica mostra qualche debolezza, è il caso delle concavità presenti come decorazioni sullo scafo o delle bandiere, specialmente vicino all'asta che le sorregge. Inoltre vi sono sezioni del modello con scarsa ridondanza di informazioni dovute alla silhouette, come ad esempio il volume subito dietro le vele, ben visibile nella figura 5.3(b).

La tesi ha dimostrato la possibilità di ridurre il processo di calibrazione del sistema di acquisizione. Un ulteriore lavoro deve essere fatto ai fini di valutare la soglia di curvatura utilizzata dal detector FCDPA dipendentemente dalla topologia del target preso in considerazione. Per il momento si assume come soglia empirica 0.32, valore scelto in quanto le acquisizioni eseguite hanno evidenziato l'esistenza di un range ottimo, come visibile in figura 6.1, compreso tra 0.15 e 0.45.

È da considerare infine che un'ipotesi fondamentale è stata introdotta per proporre l'euristica alla base di questa tesi; secondo tale ipotesi si deve avere la possibilità di un'acquisizione del target da punti di vista definibili attraverso una rototraslazione costante tra un punto di presa e il successivo. È quest'ipotesi a introdurre la ridondanza di informazione che permette di utilizzare gli algoritmi di self-calibration per la stima della matrice fondamentale nonostante l'informazione presente sulle immagini sia minima.



(a) Differenze tra parametri effettivi e stimati. (b) Differenze tra angoli di rotazione effettivi e stimati - Scala logaritmica.



(c) Ingrandimenti dei grafici precedenti nel range ottimo.

Figura 6.1: Errori relativi ai parametri estrinseci.

Anche qualora fossero disponibili i parametri intrinseci per il sistema di acquisizione, l'algoritmo è in grado di rilevare solo una visione parziale dei parametri estrinseci; in particolare non è possibile (data la scalabilità della matrice fondamentale) rilevare la dimensione del vettore di traslazione tra un sistema di acquisizione e il "successivo". Quella che viene a perdersi è effettivamente l'informazione metrica mentre quella topologica rimane valida.

Il metodo si propone pertanto come mezzo per una ricostruzione qualitativa di un target.

Appendice A

Implementazione MATLAB

Per questa tesi sono stati utilizzati i seguenti file:

MM_chain_final.m

Script che analizza la cartella dove sono salvate tutte le immagini acquisite, le elabora ad una ad una estraendone gli ereditari ed elabora le catene restituendo alla fine un cell array dove ogni cella contiene una matrice $n \times 2$ contenente le due coordinate (x e y) degli n punti appartenenti alla catena.

fast_cpda.m

Detector delle features, è l'implementazione del detector proposto in [20]. Fornisce in uscita l'immagine con indicata la posizione di tutte le features, il vettore con le coordinate e il valore di curvatura calcolato per ogni feature (quest'ultimo inutilizzato in questa tesi).

SIM1_(...).m -> SIM10_(...).m

Sono i files realizzati per la calibrazione del sistema in occasione degli esperimenti precedenti [27] [28], qui usati in versione modificata per l'estrazione della matrice fondamentale.

PhaseA_(...).m -> PhaseL_(...).m

I file utilizzati per riunire tutti i dati ottenuti dai precedenti script, restituiscono sia il modello 3D sotto forma di file WRL che le analisi degli errori e del rumore relativi ai parametri calcolati.

Altri Files

Oltre a quelli descritti sopra, sono stati usati molteplici files come “quick

fixes” o script di prova. Data la loro natura estemporanea e accessoria all’esecuzione, vengono solamente citati qui per completezza.

Bibliografia

- [1] Dana H. Ballard, Christopher M. Brown, *Computer Vision*, 1982, Prentice-Hall Inc.
- [2] AA.VV., *Encyclopedia Britannica Vol.23 Part III (World War - Zygote)*, reperito all'indirizzo http://gluedideas.com/content-collection/Encyclopedia-Britannica-Vol-23-World-War-Zygote/Nature-of_P10.html, ultima visita il 16 aprile 2012
- [3] Alexi Assmus, *Early History of X Rays*, 1995, <http://www.slac.stanford.edu/pubs/beamline/25/2/25-2-assmus.pdf>, ultima visita il 16 aprile 2012
- [4] A. Haase, G. Landwehr, E. Umbach, *Röntgen Centennial - X-rays in Natural and Life Sciences*, 1997, World Scientific Publishing Co. Pte. Ltd.
- [5] Luca Poletto, *Corso di Nanotecnologie Ottiche e Laser*, 2010, appunti personali
- [6] W. A. Kalender, *Computer Tomography*, 2005, Publicis
- [7] W. Hendee, *Biomedical Uses of Radiation*, 1999, Wiley
- [8] AA.VV., *Other Uses of X-Rays*, <http://en.wikipedia.org/wiki/X-ray>, ultima visita il 16 aprile 2012
- [9] NASA CHANDRA X-ray observatory website, <http://chandra.harvard.edu/about/>, ultima visita il 16 aprile 2012
- [10] Goddard Space Flight Center, *X-ray Astronomy Satellites & Missions*, http://imagine.gsfc.nasa.gov/docs/sats_n_data/xray_missions.html, ultima visita il 16 aprile 2012
- [11] M. Uda, G. Demortier, and I. Nakai, *X-Rays for Archeology*, 2005, Springer

-
- [12] F. Casali, *X-ray and neutron digital radiography and computed tomography for cultural heritage*, Physical Techniques in the Study of Art, Archaeology and Cultural Heritage, 2006 ,Elsevier
- [13] Hervé Delinglette, *Simplex Meshes: a General Representation for 3D Shape Reconstruction*, Marzo 1994, INRIA Sophia-Antipolis Programme 4, Rapport de recherche num. 2214
- [14] Y. Yemez, Y.Sahillioğlu, *Shape from silhouette using topology-adaptive mesh deformation*, Maggio 2009, Pattern Recognition Letters
- [15] Aldo Laurentini, *The Visual Hull Concept for Silhouette-Based Image Understanding*, Febbraio 1994, IEEE transactions on pattern analysis and machine intelligence, Vol. 16, No. 2
- [16] Asano, T., Chen, D.Z., Katoh, N., Tokuyama, T. *Polynomial-time solutions to image segmentation*, Gennaio 1996, Proc. of the 7th Ann. SIAM-ACM Conference on Discrete Algorithms
- [17] Emanuele Simioni, *Analysis of the Stereoscopic Performance of the Stereo Camera (STC) for the BEPICOLOMBO Mission: A New Method for 3D Reconstruction*, 2011, Tesi di Dottorato
- [18] Richard I. Hartley, Andrew Zisserman, *Multiple View Geometry in computer vision, 2nd edition*, 2003, Cambridge University Press
- [19] Sadegh Abbasi, Farzin Mokhtarian, Josef Kittler, *Curvature scale space image in shape similarity retrieval*, 1999 Multimedia Systems 7, Springer-Verlag
- [20] Mohammad Awrangjeb, Guojun Lu, Clive S. Fraser, Mehdi Ravanbakhsh, *A Fast Corner Detector Based on the Chord-to-Point Distance Accumulation Technique*, Dicembre 2009, Digital Image Computing: Techniques and Applications (DICTA '09)
- [21] HC Longuet-Higgins, *A computer algorithm for reconstructing a scene from two projections.*, 1981, Nature 293
- [22] Richard I. Hartley, *In defence of the 8-point algorithm*, October 1997, IEEE Transactions on Patter Analysis and Machine Intelligence
-

- [23] Hamamatsu Photonics K.K., *L8601-01 Operating Manual*, versione 1.3e
- [24] L. Poletto, M. Caldon, G. Tondello, A. Megighian, *A system for high-resolution X-ray phase-contrast imaging and tomography of biological specimens* 2008, Proc. SPIE 7078, 70781P
- [25] G. Xu, Z. Zhang, *Epipolar Geometry in Stereo, Motion and Object Recognition*, 1996, Kluwer Academic Publishers
- [26] Hamamatsu Photonics K.K., *C7943CA-22 Flat Panel Sensor Datasheet*, Giugno 2010, No. di catalogo
- [27] E. Simioni, F. Ratti, I. Calliari, L. Poletto, *Three-dimensional reconstruction with X-ray shape-from-silhouette*, 2010, SPIE Proc. Vol 7804, Development in X-ray tomography VII
- [28] E. Simioni, F. Ratti, I. Calliari, L. Poletto, *Three-dimensional modeling using x-ray shape-from-silhouette*, 1 Luglio 2011, Applied Optics Vol. 50, No. 19

Ringraziamenti

Con questa tesi si conclude un periodo intenso e importante della mia vita, e non avrei raggiunto questo notevole risultato senza l'aiuto di chi mi è stato vicino. In queste righe spero di riuscire a esprimere la mia gratitudine per quanto mi è stato dato in questi anni;

Desidero ringraziare i miei genitori Valerio e Paola e mia sorella Barbara, per il loro insostituibile supporto e affetto in tutti questi anni.

Grazie, grazie grazie alla mia ragazza Anna, che non mi ha mai abbandonato in tutti questi anni.

Ringrazio il professor Luca Poletto e il dottor Emanuele Simioni per la disponibilità e il supporto datimi per la tesi.

Un grazie corale per aver condiviso le gioie e le pene di questi anni di Ingegneria Elettronica alla compagine del DEI, ovvero Daniele, Alberto, Denis, Giulia, Davide, Alessandro, Andrea e Piero.

In questi anni patavini ho avuto il piacere e l'onore di conoscere delle persone speciali, con le quali ho diviso momenti indimenticabili (e anche il divano del soggiorno! XD), grazie mille a Tommaso, Stefano, Igor, Marco, Martina, Sara, Nicola e Federico!

Tuttavia senza il supporto degli amici storici non avrei mai potuto arrivare a oggi, pertanto un gratissimo abbraccio a Filippo, Francesco, Marco, Paolo, Mattia, Alberto, Andrea, Agostino e anche a Laura, Silvia, Claudia, Maria Chiara, Fedra, Fulvia e Angela!

In questa pagina lo spazio è troppo poco e la mia memoria è troppo scarsa per nominare singolarmente tutti quanti, ma ognuno ha un posto speciale dove più conta.

Un grazie veramente di cuore, per tutto.

Marcomattia

A mio nonno