

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II
SCUOLA DI DOTTORATO IN INGEGNERIA INDUSTRIALE



DOTTORATO DI RICERCA IN INGEGNERIA DEI SISTEMI MECCANICI

XXII CICLO

DIPARTIMENTO DI INGEGNERIA MECCANICA PER L'ENERGETICA

TESI DI DOTTORATO

**APPLICAZIONI DI SISTEMI DI VISIONE AD UNA CELLA
ROBOTIZZATA**

RELATORI

Ch.mo Prof. Ing. Giovanni Adiletta

Ch.mo Prof. Ing. Cesare Rossi

CANDIDATO

Ing. Salvatore Strano

COORDINATORE DEL DOTTORATO

Ch.mo Prof. Ing. Raffaele Tuccillo

INDICE

INTRODUZIONE.....	1
-------------------	---

CAPITOLO 1

I sistemi di visione

1.1 La visione artificiale.....	4
1.2 Sistemi di visione: classificazione e principio di funzionamento.....	6
1.3 Lenti sottili e sue aberrazioni.....	10
1.4 Il modello della telecamera.....	13
1.5 La trasformazione prospettica.....	14
1.6 Il modello Pin-Hole e sua discretizzazione.....	19
1.7 Modello della distorsione a cinque parametri.....	25

CAPITOLO 2

Panoramica dei metodi ottici per l'acquisizione di forme tridimensionali

2.1 Introduzione.....	29
2.2 Dalla misura al modello 3D.....	32
2.2.1 Registrazione.....	33
2.2.2 Fusione geometrica.....	34
2.2.2.1 Integrazione di maglie.....	34
2.2.2.2 Fusione volumetrica.....	35
2.3 Metodi ottici per l'acquisizione di forma.....	36
2.3.1 Depth from focus/defocus.....	37
2.3.1.1 Fochettatura.....	37
2.3.1.2 Sfocamento.....	37
2.3.2 Shape from texture.....	38
2.3.3 Shape from shadin.....	40
2.3.4 Stereo fotometrico.....	40
2.3.5 Stereopsi.....	40
2.3.6 Shape from silhouette.....	42
2.3.7 Shape from photo-consistency.....	44

2.3.8 Structure from motion.....	46
2.3.9 Active defocus.....	47
2.3.10 Stereo attivo.....	47
2.3.11 Triangolazione attiva.....	49
2.3.11.1 Metodo a luce codificata.....	50
2.3.12 Interferometria.....	52
2.3.13 Tempo di volo.....	53

CAPITOLO 3

Ricostruzione tridimensionale con tecnica: Shape from silhouettes

3.1 Principio di ricostruzione volumetrica dalle sagome.....	55
3.1.1 Algoritmo di Szeliski.....	55
3.1.2 Metodo proposto.....	57
3.2 Elaborazione delle immagini.....	57
3.3 Calibrazione della telecamera.....	59
3.3.1 Formalizzazione del problema di calibrazione.....	61
3.3.2 Un algoritmo di calibrazione.....	65
3.3.3 Software utilizzato per la calibrazione.....	72
3.4 Descrizione dell'algoritmo proposto.....	75
3.5 Valutazione della risoluzione.....	78

CAPITOLO 4

Ricostruzione tridimensionale mediante triangolazione laser

4.1 Principio di triangolazione laser.....	80
4.2 Determinazione della linea laser.....	83
4.3 Calibrazione del sistema laser scanner.....	87
4.4 Valutazione della risoluzione.....	90
4.4.1 Errore dovuto alla discretizzazione del sensore ottico.....	90
4.4.2 Errore dovuto al modello di laser scanner.....	90
4.4.3 Precisione dovuta al sistema laser scanner.....	92
4.4.4 Scanner range.....	93

CAPITOLO 5

Utilizzo di un robot antropomorfo per la digitalizzazione di superfici

5.1 Robot.....	95
5.1.1 Struttura del robot.....	95
5.1.1.1 Azionamenti.....	98
5.1.1.2 I riduttori.....	100
5.1.1.3 I cuscinetti.....	103
5.1.2 Cinematica.....	104
5.1.2.1 Matrice di Denavit-Hartenberg.....	106
5.2 Descrizione del sistema robot – sistema di acquisizione 3D.....	110
5.2.1 Sistema robot – fotocamera – tavolo girevole.....	110
5.2.1.1 Posizionamento relativo tra i vari sistemi di riferimento.....	111
5.2.1.2 Calibrazione del sistema.....	113
5.2.1.3 Descrizione della procedura di ricostruzione tridimensionale.....	114
5.2.1.4 Risultati.....	116
5.2.2 Sistema robot – laser scanner.....	117
5.2.2.1 Calibrazione del sistema.....	118
5.2.2.2 Descrizione della procedura di ricostruzione tridimensionale.....	119
5.2.2.3 Risultati.....	121
5.3 Confronto fra le due tecniche.....	126

CAPITOLO 6

Utilizzo di un robot antropomorfo per la realizzazione di superfici digitalizzate

6.1 Descrizione della procedura.....	127
6.2 Pianificazione della traiettoria.....	130
6.3 Controllo del robot.....	143
6.3.1 Il comando “blend”.....	145
6.3.2 Scelta dei punti intermedi in una traiettoria continua.....	148
6.3.3 Software di gestione della scheda di controllo assi.....	150
6.4 Risultati.....	153
CONCLUSIONI.....	156

BIBLIOGRAFIA.....157

INTRODUZIONE

Attualmente la procedura di replicazione di un oggetto non noto è divisa in varie fasi: si utilizzano prima macchine per il rilievo di forme e successivamente celle di fabbricazione del prototipo, in fine sono utilizzati altri strumenti per la verifica di qualità.

Il rilievo della forma di oggetti tridimensionali è un processo che ha, recentemente, assunto un ruolo significativo in svariate applicazioni industriali. Gli attuali strumenti per il rilievo di forma consentono di riprodurre con un'elevata accuratezza e ripetibilità le caratteristiche geometriche di oggetti anche complessi. Per questo motivo il loro impiego è in continua crescita: dal Reverse Engineering (RE) per l'analisi della concorrenza al controllo di qualità di prodotti industriali; dalla creazione di modelli da inserire in ambienti creati attraverso la realtà virtuale, alla ricostruzione o al restauro di manufatti appartenenti al campo dei beni culturali. Per una così vasta gamma di applicazioni sono stati sviluppati ed adottati sistemi basati su differenti principi di funzionamento (sensori meccanici, ottici, laser). La procedura di RE comporta in genere di ricavare il modello geometrico dal modello fisico, attraverso due fasi: la digitalizzazione tridimensionale del modello fisico e la realizzazione del suo modello geometrico elettronico.

Per quanto riguarda la produzione del modello fisico si possono distinguere le tecniche classiche:

- Asportazione di truciolo
- Formatura
- Deformazione

e le tecniche di Rapid Prototyping and Manufacturing (RP&M):

- Accrescimento

Nella procedura per asportazione di truciolo il prototipo viene costruito per asportazioni successive di materiale partendo da un grezzo di lavorazione (lavorazioni di tornitura, fresatura, elettroerosione, ecc.); il modello CAD del prototipo viene utilizzato per generare e verificare il percorso dell'utensile (CAM).

Nella tecnica di deformazione il prototipo viene costruito deformando plasticamente in stampi i semilavorati (a caldo e/o a freddo) ed il modello CAD del prototipo viene utilizzato per progettare e definire il modello CAD degli stampi ed inoltre per generare e verificare il percorso dell'utensile (CAM) per la produzione degli stampi. Nella procedura per formatura il prototipo viene ottenuto depositando in stampi il materiale allo stato liquido o sotto forma di polvere ed il modello CAD del prototipo viene utilizzato per progettare e definire il modello CAD degli stampi ed inoltre per generare e verificare il percorso dell'utensile (CAM) per la produzione degli stampi. Quindi si può dire che tutte le tecniche classiche di produzione prototipi richiedono step successivi alla definizione del modello CAD, per generare e verificare i percorsi utensile (CAM). Ciò può essere necessario per costruire il prototipo stesso (asportazione di truciolo) o gli stampi necessari per costruirlo (deformazione o formatura).

Le tecniche RP&M permettono di ottenere i prototipi per accrescimento (addizione e successiva aggregazione di materiale sotto forma di particelle, filamenti o strati) direttamente dal modello CAD, con tempi e costi decisamente inferiori rispetto ai prototipi prodotti con tecniche tradizionali anche se si può utilizzare una gamma limitata di materiali. Le più comuni sono:

- Stereolithography (SL)
- Selective Laser Sintering (SLS)
- Fused Deposition Modeling (FDM)
- Laminated Object Manufacturing (LOM)
- Solid Ground Curing (SGC)
- 3D Printing

In questa tesi è stato sviluppato un sistema capace di acquisire e realizzare superfici di interesse, con la possibilità di verificare il risultato ottenuto mediante l'acquisizione della forma della copia realizzata. La motivazione di questo lavoro consiste nel vantaggio di avere un unico sistema mediante il quale è possibile realizzare un prototipo a partire da un oggetto del quale non si dispone di un modello CAD.

L'apparecchiatura messa a punto quindi consente il rilievo di superfici al fine di poterle riprodurre, mediante lavorazione meccanica. Lo sviluppo di una tale

apparecchiatura, si pone l'obiettivo di rendere flessibile ed al tempo stesso, ripetibile, un sistema di acquisizione di forme, conferendogli un moto controllato fedelmente da un robot. Contrariamente agli apparati di acquisizione statici, una tale strumentazione consente di muovere il sistema di visione rispetto all'oggetto da analizzare, senza introdurre problematiche di matching dei dati. D'altra parte in tale applicazione, il sistema di visione diviene un dispositivo integrato, avente funzioni di trasduttore di posizione, di riconoscimento di forma e volume, creando così un miglioramento di prestazioni del robot se il sistema stesso viene inserito nell'anello di controllo, e gli conferisce la capacità di utilizzare informazioni nel suo spazio operativo per poter riprodurre le superfici osservate.

Nella prima parte della tesi sono stati studiati diversi metodi ottici di acquisizione di forma ed è stata sviluppata con l'ausilio di un robot industriale la sperimentazione in laboratorio di tali tecniche affrontando problemi come la calibrazione del sistema, la velocità della procedura e l'accuratezza. Nella seconda parte si è sviluppata la procedura di fabbricazione delle superfici ricostruite, in particolare, utilizzando lo stesso robot ed adottando la tecnica di fabbricazione per asportazione di truciolo, è stato sperimentato il processo di riproduzione del modello fisico. Le varie fasi sopra riportate sono state correlate da software interamente sviluppati in casa che permettono di gestire le operazioni in modo semplice e rapido.

CAPITOLO 1

I sistemi di visione

1.1 La visione artificiale

Tra tutte le abilità sensoriali, la visione è largamente riconosciuta come quella con le maggiori potenzialità. Le capacità dei sistemi biologici sono formidabili: l'occhio raccoglie, infatti, una banda di radiazioni elettromagnetiche rimbalzate su diverse superfici e provenienti da fonti luminose diverse ed il cervello elabora questa informazione formando il quadro della scena come noi la percepiamo.

La visione artificiale è il processo che estrae informazioni da una scena analizzando l'immagine di quella scena [Rosenfeld, 1988]. L'analisi è finalizzata a scoprire cosa è presente nella scena e l'esatta collocazione degli oggetti. In genere, la scena è illuminata da una o più sorgenti di luce, naturale o artificiale; la telecamera (il sensore preposto alla percezione visiva) è dotata di un dispositivo ottico in grado di formare un'immagine della scena su un piano: in esso, per ogni punto, la luminosità ed il colore sono correlati alla luce ricevuta dal sensore e proveniente da una certa direzione.

Per poter essere sottoposta ad elaborazione per mezzo di un computer digitale, l'immagine deve essere campionata (o come si suole dire, discretizzata), ottenendo una matrice di valori (*pixel*), ognuno legato alla luminosità o al colore in quel punto o alla media in un intorno. Il valore di ogni pixel può essere uno scalare (livello di grigio) o un vettore di n componenti (tipicamente $n = 3$), ognuna delle quali legata ad una particolare banda spettrale, e tutti insieme atti a caratterizzare l'informazione di colore del pixel.

È da sottolineare che tutti i mezzi di visione artificiale non tentano in nessun modo di replicare la visione umana. Il tentativo è destinato, infatti, al fallimento per la intrinseca differenza tra i due hardware.

Riprendendo la definizione di visione precedentemente data, l'immagine, o una sequenza di immagini della scena, costituisce l'input dell'intero processo e le informazioni estratte ne costituiscono l'output. In taluni casi, l'input può essere dato da una coppia di immagini relative alla stessa scena, ma acquisite da due angolazioni diverse.

Storicamente vi sono tre differenti approcci alla visione artificiale:

- ↵ L'approccio gerarchico [Marr, 1982];
- ↵ La visione attiva [Bajcsy, 1988];
- ↵ Le routines visuali [Ullman, 1985].

I tre approcci ricalcano direttamente le architetture robotiche: l'approccio di Marr è sicuramente correlato alle architetture deliberative [Albus *et al.*, 1987], laddove la visione attiva è legata alle architetture controllistiche, mentre le routines visuali risentono molto dell'influenza delle architetture basate sui comportamenti [Arkin, 1998].

Bisogna comunque osservare che un sistema di visione reale utilizza in genere tutti e tre gli approcci citati.

Si definisce “Sistema di visione” un insieme di dispositivi atti alla cattura permanente di immagini reali, su di un supporto qualsiasi. Esso è composto essenzialmente da un sistema ottico di puntamento, cui è attribuita la funzione di convogliare il fascio luce che gli oggetti riflettono, da un supporto sensibile, in genere un sensore di visione in grado di trasformare un fascio luminoso incidente in un segnale elettrico e da un sistema di acquisizione capace di trasformare tale segnale in un'immagine “catturata” su di un opportuno supporto.

Generalmente, l'obiettivo della visione è quello di riconoscere gli oggetti presenti nella scena e determinarne la posizione nello spazio rispetto ad un riferimento. In linea di principio, la tipologia delle informazioni estratte dipende dall'applicazione e dalle sue finalità ed esigenze. Quando è possibile, si evita di effettuare una ricostruzione tridimensionale della scena, cercando di estrarre solo le informazioni utili allo scopo.

In ambito robotico, la visione trova applicazione nella manipolazione degli oggetti, nel riconoscimento degli oggetti, del volume di lavoro, nell'acquisizione delle informazioni necessarie per spostarsi da un punto all'altro dello spazio di lavoro, evitando gli ostacoli eventualmente presenti nel suo percorso. In nessuna di esse vi è la reale necessità di una completa descrizione dell'ambiente, ma una serie di informazioni necessari per costruire una scena virtuale nel quale il robot può interagire in condizioni di sicurezza.

1.2 Sistemi di visione: classificazione e principio di funzionamento

L'utilizzo di sistemi di visione si riscontra nei campi più disparati. Le prime applicazioni rilevabili sono senz'altro nell'ambito della sorveglianza remota di ambienti, con evoluzione dal semplice monitoraggio a funzionalità di riconoscimento di oggetti in moto. Ma è nella robotica che un sistema di visione diviene un dispositivo integrato, avente funzioni di trasduttore di posizione, di riconoscimento di forma e volume, creando così un miglioramento di prestazioni del robot soprattutto nei casi in cui il sistema di visione venga inserito nell'anello di controllo, ovvero interagisca quasi istantaneamente con il robot. Si tratta, di ottimizzare e incrementare il volume di informazioni a disposizione del robot; da questo punto di vista ha senso parlare di visione artificiale.

Per semplificare è possibile suddividere la visione artificiale, in visione di basso livello e di alto livello. La prima si occupa di estrarre determinate proprietà fisiche dell'ambiente visibile, come profondità, forma tridimensionale, contorni degli oggetti. I processi di visione di basso livello sono tipicamente paralleli, spazialmente uniformi e relativamente indipendenti dal problema e dalla conoscenza a priori associata a particolari oggetti. Viceversa, la visione di alto livello si occupa dell'estrazione delle proprietà delle forme, e di relazioni spaziali, di riconoscimento e classificazione di oggetti.

La non perfetta ripetibilità dei movimenti comandati ad un meccanismo robotico è l'esempio lampante di quello che un sistema di visione può mettere in risalto facilitando le eventuali correzioni da effettuare. L'integrazione di dispositivi di percezione nei sistemi robotici è motivato e favorito dalla continua necessità di aumentare la loro versatilità e la loro sfera di applicazione.

Il processo di visione del robot, comunemente chiamato 'machine vision' o 'computer vision', è sicuramente il passo più grande verso il perfezionamento della sua automazione.

La percezione visiva per un sistema robotizzato è, dunque, una caratteristica che influenza sensibilmente le sue capacità di interagire con l'ambiente circostante, ma ancor prima di arrivare all'interazione con l'ambiente, la visione risulta un ottimo e, per certi versi insostituibile, mezzo di valutazione ed analisi dell'operato del robot stesso.

In questa ottica si inseriscono il sistema di acquisizione delle immagini tramite telecamera digitale che è stato messo a punto, e l'indagine che si è eseguita sul robot sperimentale del laboratorio di robotica del D.I.M.E [1,2,4,6].

Un apparato di acquisizione di immagini funziona raccogliendo la luce riflessa dagli oggetti della scena e creando una immagine bidimensionale. Un sistema di visione è principalmente rappresentato da una telecamera.

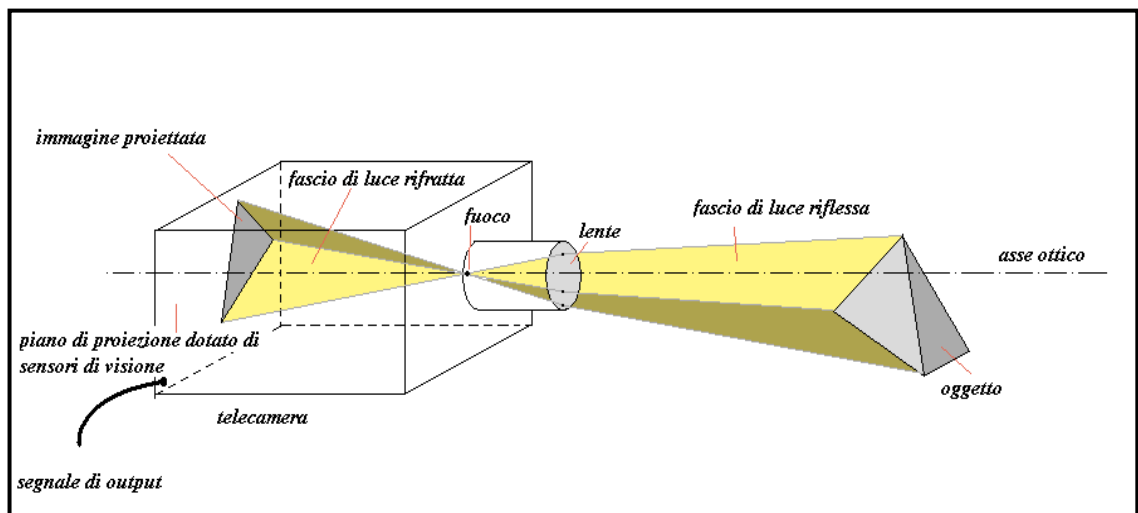


Fig. 1.1

La luce proveniente da una sorgente incide sulle superfici dell'oggetto, il fascio di luce riflessa, dovuta alla riflessione del fascio di luce di partenza, incidono sulla superficie della lente convergente, vengono rifratti a causa del passaggio da un mezzo all'altro (aria-vetro, vetro-aria), e convogliati in un punto detto fuoco. I raggi rifratti, obbligati a passare per uno stesso punto, mutano la propria direzione, e successivamente incidono sul piano di proiezione, dove si trovano i sensori di visione. Proprio a causa della deflessione del fascio di luce, si ha che le proiezioni dell'oggetto su tale piano avranno dimensioni funzioni della geometria del sistema ottico, degli indici di rifrazione dei mezzi attraversati e della distanza dell'oggetto dall'obiettivo; è di banale constatazione, guardando una fotografia o un filmato, che oggetti più lontani appaiono più piccoli.

Questo è il modello più semplice della formazione dell'immagine: la telecamera *pin-hole* (letteralmente fotocamera a foro di spillo).

Sia P un punto della scena, di coordinate (X, Y, Z) e sia P' la sua proiezione sul piano immagine (o piano retina), di coordinate (X', Y', Z') : Se f (distanza focale), è la distanza del foro (o centro di proiezione) O dal piano immagine allora dalla similarità dei triangoli si ottiene:

$$X' = -f \cdot \frac{X}{Z} \quad Y' = -f \cdot \frac{Y}{Z} \quad Z' = -f \quad (1.1)$$

Si noti che l'immagine risulta invertita rispetto alla scena, sia destra-sinistra che sopra-sotto, come indicato dal segno meno. Queste equazioni definiscono il processo di formazione dell'immagine che prende il nome di proiezione prospettica. Equivalentemente possiamo modellare la proiezione prospettica ponendo il piano immagine davanti al centro di proiezione, eliminando così il segno negativo. Questo accorgimento fu per la prima volta raccomandato dal pittore rinascimentale Leon Battista Alberti, accreditato come scopritore della prospettiva, con il suo trattato “*De Pictura*” del 1435.

La divisione per Z è responsabile dell'effetto di scorcio per cui la dimensione dell'immagine di un oggetto varia in ragione della sua distanza dall'osservatore.

Con questa tecnica si avrà un'immagine proiettata sul piano dei sensori; in altri termini, un fascio di luce incide sul piano retina e, quindi, ciascuno dei sensori viene investito da una certa quantità di energia associata al treno di onde luminose. Si ricorda che un sensore è un dispositivo in grado di convertire la grandezza fisica che si vuole rilevare o misurare, in un'altra grandezza, in genere elettrica, legata alla prima da una relazione che in prima approssimazione è di proporzionalità. Nel caso dei sensori di visione, la grandezza da rilevare è l'energia luminosa posseduta dal fascio di luce incidente, che viene convertita in una grandezza elettrica processabile per via analogica o digitale.

Si nota esplicitamente che il piano retina è composto da una fitta rete di sensori, e quindi ci sarà un insieme di segnali analogici ognuno dei quali proviene da un singolo rettangolino sensibile.

Un sistema di acquisizione di immagini è costituito da un'unità chiamata *frame grabber* (scheda di acquisizione), una *telecamera* ed un *calcolatore host*.

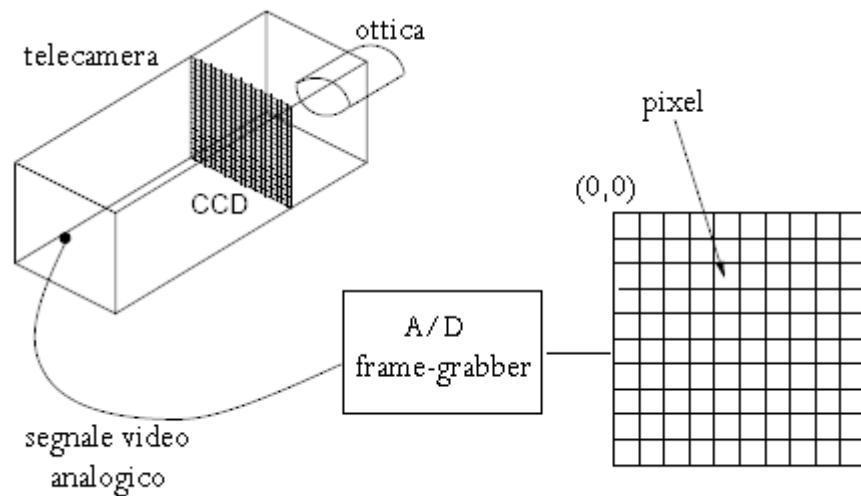


Figura 1.2

La telecamera è composta dall'ottica (che approssimiamo con una lente sottile), e da una matrice di CCD (Charged Copuled Device) che costituisce il piano immagine (un CCD misura circa 1 cm^2 e contiene circa 5×10^5 elementi). Possiamo vedere quest'ultima come una matrice $n \times m$ di celle rettangolari fotosensibili, ciascuna dei quali converte l'intensità della radiazione luminosa incidente in un potenziale elettrico. L'uscita della telecamera a CCD è un segnale elettrico analogico, ottenuto leggendo il potenziale degli elementi della matrice CCD per righe. Il segnale video (analogico) viene letto dal frame grabber che lo digitalizza, convertendolo in una matrice $N \times M$ di valori interi. Gli elementi della matrice prendono il nome di pixel (picture element).

Indicheremo con $I(u,v)$ il valore dell'immagine (luminosità) nel pixel individuato dalla riga v e colonna u (sistema di coordinate $u-v$ con l'origine nell'angolo in alto a sinistra). La luminosità $I(p)$ di un pixel p nell'immagine è proporzionale alla quantità di luce che la superficie S_p riflette verso la telecamera (S_p è la superficie che si proietta nel pixel p).

La dimensione $n \times m$ della matrice CCD non è necessariamente la stessa della immagine $N \times M$ (matrice dei pixel); per questo motivo la posizione di un punto del piano immagine è diversa se misurata in elementi CCD piuttosto che in pixel.

Naturalmente la relazione tra pixel ed elementi CCD, immaginando che esse sono associate alla stessa realtà fisica, e che in fondo al pixel corrisponde un'area rettangolare sul CCD array (l'impronta del pixel), sono semplici relazioni di proporzionalità [1,2,6]:

$$\begin{cases} u_{pixel} = \frac{N}{n} \cdot u_{CCD} \\ v_{pixel} = \frac{M}{m} \cdot v_{CCD} \end{cases} \quad (1.2)$$

1.3 Lenti sottili e sue aberrazioni

Nella descrizione del principio di funzionamento di un sistema di visione si è fatto riferimento al modello pin-hole. Un oggetto viene proiettato su un piano immagine attraverso il centro ottico secondo i principi matematici della geometria proiettiva; il modello è anche detto “a proiezione prospettica”, ed è quello che abbiamo assunto come modello della telecamera. Il pin-hole non è realisticamente realizzabile, dato che la sua dimensione non potrà mai essere esattamente puntiforme. Esso non modella gli effetti di aberrazione ottica, né le distorsioni tipiche delle immagini reali. Si studiano così le lenti (sottili) ed il loro modello matematico.

Una lente, essendo più grande di un foro di spillo, riesce a raccogliere più luce. Il rovescio della medaglia è, tuttavia, quello che non tutta la scena può essere a fuoco allo stesso tempo. L'approssimazione che si compie per l'ottica del sistema di acquisizione, che in generale è molto complessa, essendo costituita da più lenti, è quella della lente sottile. Le lenti sottili godono delle seguenti proprietà:

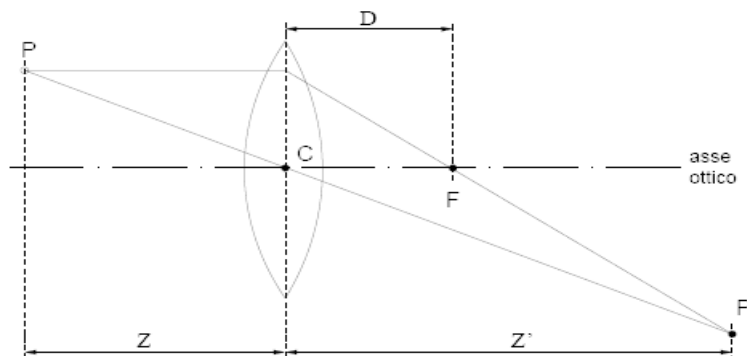


Fig. 1.3

- 1) I raggi paralleli all'asse ottico incidenti sulla lente vengono rifratti in modo da passare per un punto dell'asse ottico chiamato fuoco F ;
- 2) I raggi che passano per il centro C della lente sono inalterati.

L'approssimazione di lente sottile è valida quando lo spessore della lente è molto più piccolo rispetto alle distanze in gioco: tale situazione è quasi sempre verificata nelle applicazioni tradizionali.

La distanza del fuoco F dal centro della lente C prende il nome di distanza focale D ; essa dipende dai raggi di curvatura delle due superfici della lente e dall'indice di rifrazione del materiale costituente.

Dato un punto della scena P , è possibile costruirne graficamente l'immagine P' (o punto coniugato), servendosi di due raggi particolari che partono da P : il raggio parallelo all'asse ottico, che dopo la rifrazione passa per F ed il raggio che passa inalterato per C . Grazie a questa costruzione ed alla similarità dei triangoli, si ottiene la “formula dei punti coniugati”, o “equazione della lente sottile”:

$$\frac{1}{Z} + \frac{1}{Z'} = \frac{1}{D} \quad (1.3)$$

L'immagine di un punto della scena distante Z dalla lente, viene prodotta ad una distanza dalla lente Z' , che dipende dalla profondità Z del punto e dalla distanza focale D della lente. Se si cambia Z con Z' si ottiene l'immagine sfocata del punto, ovvero un cerchio, che prende il nome di cerchio di confusione. Finché il cerchio di confusione non supera le dimensioni dell'elemento fotosensibile, l'immagine risulta a fuoco. Esiste un intervallo di profondità, per il quale i punti sono a fuoco, definito *profondità di campo*. Per mettere a fuoco oggetti a distanze diverse, le lenti dell'occhio cambiano forma (focale), mentre le lenti delle telecamere traslano nella direzione Z .

La telecamera pin-hole ha una profondità di campo infinita, mentre per telecamere reali, la profondità di campo è inversamente proporzionale al diametro dell'iride. La

luce che viene raccolta, invece, è direttamente proporzionale al diametro dell'iride [2].

Le lenti reali sono affette da una serie di aberrazioni, dovute ad un differente comportamento rispetto alle lenti sottili, per cui il modello prospettico presentato è applicabile con una buona precisione, su una porzione dell'immagine.

Si è notato, infatti, che i raggi paralleli all'asse ottico incidenti sulla lente vengono tutti rifratti in modo da passare per uno stesso punto dell'asse ottico che abbiamo chiamato fuoco. In realtà, anche nel caso di raggi parassiali quello che si verifica è mostrato in figura 1.4:

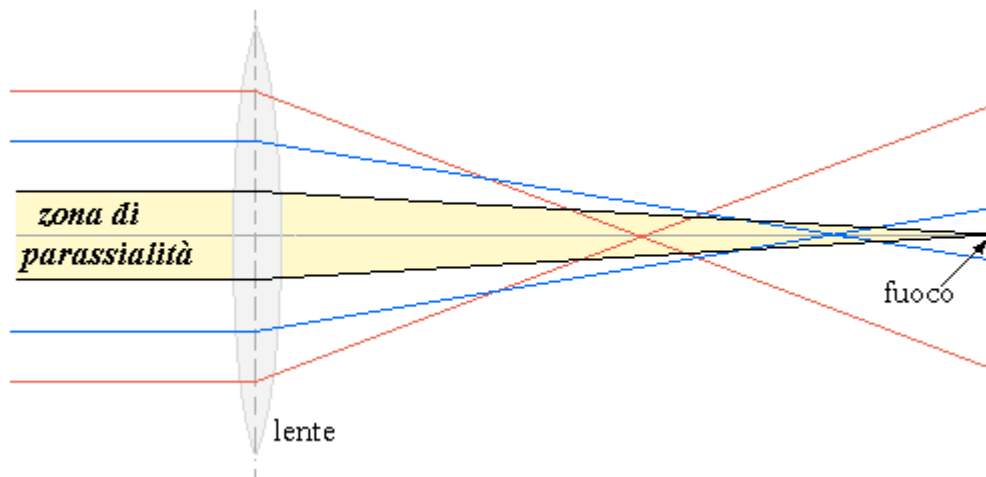


Fig. 1.4

i raggi convergono in punti diversi a seconda del punto della lente per cui passano i raggi incidenti, salvo per una zona, denominata *zona di parassialità*, contenuta in un intorno dell'asse ottico in cui tale effetto non è rilevabile.

Questo discostamento dal comportamento teorico è detto *aberrazione sferica*, ed è dovuta al fatto che l'approssimazione dell'ottica geometrica non è sempre valida, perchè la meccanica ondulatoria prevede fenomeni più complessi; ad esempio la direzione di propagazione dei raggi luminosi non può essere ritenuta rigorosamente rettilinea, al variare dall'angolo di incidenza tra raggio-lente varia l'angolo di rifrazione, e per angoli prossimi a 90° tale rifrazione non si realizza; inoltre il fenomeno dipende dalla geometria delle lenti e dagli indici di rifrazione, ed è inversamente proporzionale ad esso.

In ogni caso, i sistemi ottici componenti i principali sistemi di visione presentano strutture in grado di compensare l'effetto indesiderato dell'aberrazione sferica: abbinando una lente convergente con una divergente è possibile sommare il risultato di due aberrazioni sferiche di segno opposto, che quindi si annullano. Un altro metodo consiste nell'utilizzare superfici lenticolari a curvatura non costante, ad esempio paraboloidi. Una lente che da sola corregge l'aberrazione sferica viene detta *asferica* [1].

1.4 Il modello della telecamera

Si supponga di avere una telecamera che inquadrì una scena: si fissi in essa un sistema di riferimento cartesiano tridimensionale $\{O,x,y,z\}$ con origine arbitraria. L'immagine prodotta dalla telecamera è bidimensionale, per cui si fissi un sistema cartesiano bidimensionale $\{u,v\}$ ad essa riferito.

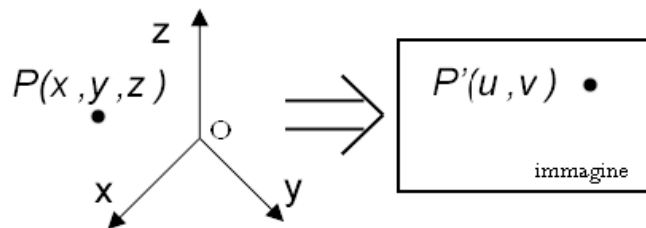


Fig. 1.5

Sia P un punto all'interno della scena e P' il punto ad esso corrispondente nell'immagine prodotta dalla telecamera. In generale, ogni punto della scena avrà il suo punto corrispondente all'interno dell'immagine. I “punti-scena” ed i “punti-immagine” sono legati tra loro mediante una trasformazione puramente geometrica. Realizzare il modello della telecamera significa trovare le relazioni geometriche che legano il modo in cui un “punto-scena” si trasforma in un “punto-immagine”; esse altro non sono che equazioni che legano i due mondi. Partendo da considerazioni puramente geometriche, relative al cambiamento di riferimento, attraverso il calcolo di un'opportuna trasformazione prospettica, si giunge all'espressione del modello finale, in cui verranno incluse anche le distorsioni, che la lente introduce per effetto

di un comportamento non ideale delle stesse; dovuta alla fisicità della lente stessa che l'allontana dal modello comportamentale delle lenti sottili.

Realizzare il modello della telecamera significa, determinare la trasformazione anzidetta, servendosi di tutti i parametri che servono per descrivere completamente il modello: i parametri interni (parametri intrinseci), e i parametri di posizione (parametri estrinseci), rispetto ad un opportuno riferimento esterno, della telecamera.

1.5 La trasformazione prospettica

Siano $\{O, x, y, z\}$ e $\{\Omega, \xi, \eta, \zeta\}$ 2 terne levogire arbitrariamente disposte, sia P , un punto dello spazio le cui coordinate cartesiane nella terna $\{O, x, y, z\}$ siano espresse dal vettore $\{w\}$, le coordinate cartesiane nel riferimento $\{\Omega, \xi, \eta, \zeta\}$ sono date dal vettore seguente:

$$\{w_c\} = [R] \cdot \{w\} + \{t\} \quad (1.4)$$

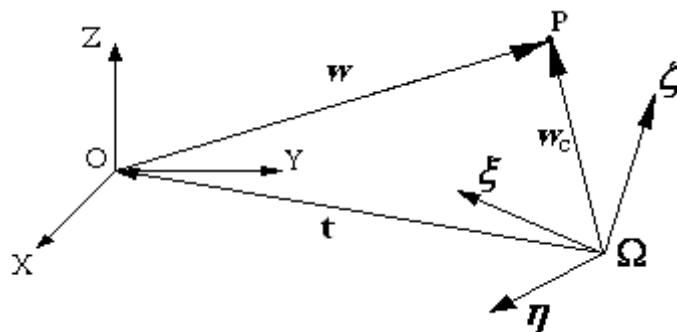


Fig. 1.6

dove

- ⊗ $\{w\}$ è espresso nel riferimento $\{O, x, y, z\}$;
- ⊗ $\{w_c\}$ è espresso nel riferimento $\{\Omega, \xi, \eta, \zeta\}$;
- ⊗ $\{t\}$ è il vettore che va dal punto Ω al punto O , ed è espresso in terna $\{\Omega, \xi, \eta, \zeta\}$;
- ⊗ $[R]$ è la matrice che mi definisce la rotazione tra i due riferimenti.

Le colonne della matrice R rappresentano le componenti dei vettori $\hat{x}, \hat{y}, \hat{z}$ nel riferimento $\{\Omega, \xi, \eta, \zeta\}$; reciprocamente, le righe di tale matrice sono le componenti dei vettori $\hat{\xi}, \hat{\eta}, \hat{\zeta}$ nel riferimento $\{O, x, y, z\}$.

In simboli tale matrice vale:

$$[R] = \begin{bmatrix} \xi_x & \xi_y & \xi_z \\ \eta_x & \eta_y & \eta_z \\ \zeta_x & \zeta_y & \zeta_z \end{bmatrix} \quad (1.5)$$

Le matrici di rotazione godono della proprietà di ortogonalità, e pertanto l'inversa di tale matrice coincide con la trasposta.

Volendo fornire una notazione più compatta della (1.4), bisogna passare alle coordinate omogenee:

$$\{\tilde{w}_c\} = \begin{Bmatrix} w_\xi \\ w_\eta \\ w_\zeta \\ 1 \end{Bmatrix} = [T] \cdot \{\tilde{w}\} = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ \zeta_x & \zeta_y & \zeta_z & t_\zeta \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{Bmatrix} w_x \\ w_y \\ w_z \\ 1 \end{Bmatrix} \quad (1.6)$$

La matrice $[T]$ che compare nella (1.6) è una matrice di trasformazione omogenea. Di questa matrice ciò che interessa è il vettore costituito dai primi tre elementi della quarta riga, che costituiscono la matrice di trasformazione prospettica. Tale trasformazione consiste nel proiettare punti, luoghi geometrici, o più in generale figure spaziali di qualsiasi tipo, su un piano, tenendo conto sia del loro orientamento sia della loro distanza da detto piano.

Si pone il problema, quindi, di ricavare l'espressione della trasformazione prospettica nel caso semplice in cui si voglia rappresentare un generico punto dello spazio tridimensionale su di un piano di proiezione. Il primo passo per fare ciò è calcolare la distanza di tale punto dal piano di proiezione. Questo può essere fatto effettuando un semplice prodotto scalare. Sia \hat{n} un versore normale al piano R nel riferimento

immagine: esso coincide con uno dei versori della terna immagine. Si faccia riferimento alla figura 1.7:

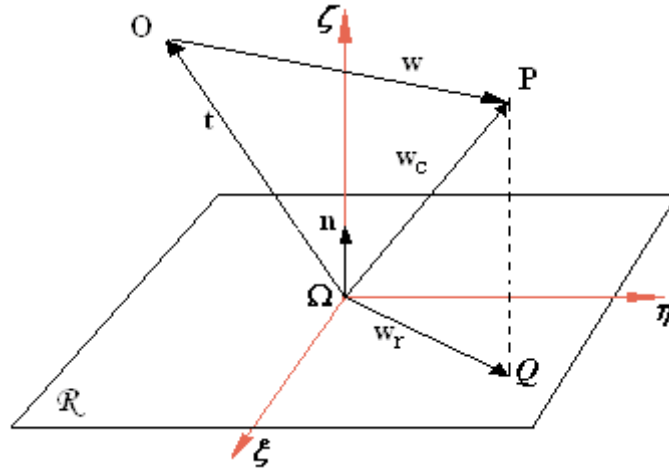


Fig. 1.7

È facile vedere che la distanza del punto P dal piano R è individuata dal segmento PQ , il cui valore può essere ottenuto come prodotto scalare tra \hat{n} e w_c e avvalendosi della proprietà di ortogonalità di $[R]$:

$$d = \{\hat{n}\}^T \cdot \{w_c\} = \{\hat{n}\}^T \cdot ([R] \cdot \{w\} + \{t\}) = ([R]^{-1} \cdot \{\hat{n}\})^T \cdot \{w\} + \{\hat{n}\}^T \cdot \{t\} \quad (1.7)$$

e con l'ausilio delle coordinate omogenee la (1.7) può scriversi:

$$d = \begin{Bmatrix} [R]^{-1} \cdot \{\hat{n}\} \\ \{\hat{n}\}^T \cdot \{t\} \end{Bmatrix}^T \{\tilde{w}\} = \{\tilde{N}\}^T \cdot \{\tilde{w}\} \quad (1.8)$$

dove il prodotto $[R]^{-1} \cdot \{\hat{n}\}$, individua le prime tre componenti del vettore $\{\tilde{N}\}$, e $\{\hat{n}\}^T \cdot \{t\}$ ne rappresenta la quarta componente. Per esteso:

$$\{\tilde{N}\} = \begin{Bmatrix} N_x \\ N_y \\ N_z \\ t_n \end{Bmatrix} = \begin{Bmatrix} \xi_x \cdot n_\xi + \eta_x \cdot n_\eta + \zeta_x \cdot n_\zeta \\ \xi_y \cdot n_\xi + \eta_y \cdot n_\eta + \zeta_y \cdot n_\zeta \\ \xi_z \cdot n_\xi + \eta_z \cdot n_\eta + \zeta_z \cdot n_\zeta \\ t_\xi \cdot n_\xi + t_\eta \cdot n_\eta + t_\zeta \cdot n_\zeta \end{Bmatrix} \quad (1.9)$$

I primi tre elementi di N corrispondono alle componenti del versore normale al piano di proiezione, scritte nel riferimento di partenza, mentre il quarto elemento rappresenta la distanza fra le due terne misurata proprio lungo tale normale.

A partire dal vettore $\{w_c\}$, che individua le coordinate del punto P in terna telecamera, è possibile determinare il vettore $\{w_r\}$: esso individua le coordinate del punto Q nel riferimento $\{\Omega, \xi, \eta, \zeta\}$, ed è, pertanto, sempre un vettore del piano retina R . In maniera del tutto generale, è possibile esprimere tale vettore come prodotto di un'opportuna matrice $[S]$ per il vettore $\{w_c\}$:

$$\{\tilde{w}_r\} = \begin{Bmatrix} w_\xi \\ w_\eta \\ w_\zeta \\ 1 \end{Bmatrix} = [S] \cdot \{\tilde{w}_c\} = \begin{bmatrix} n_\eta + n_\zeta & 0 & 0 & 0 \\ 0 & n_\xi + n_\zeta & 0 & 0 \\ 0 & 0 & n_\xi + n_\eta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{Bmatrix} w_\xi \\ w_\eta \\ w_\zeta \\ 1 \end{Bmatrix} \quad (1.10)$$

Ovviamente una delle componenti del vettore $\{w_r\}$ sarà nulla, essendo esso un vettore planare: nel caso in cui il versore \hat{n} coincida (come in figura 1.7) con l'asse ζ , sarà allora $w_\zeta=0$. Ciò si traduce, infatti, nell'espressione della matrice S , che presenterà sempre una riga nulla, e con riferimento al caso illustrato in figura 1.7, la riga nulla sarà la terza.

Sostituendo nella (1.10) l'espressione di $\{w_c\}$, si ottiene:

$$\{\tilde{w}_r\} = [S] \cdot \{\tilde{w}_c\} = [S] \cdot [T] \cdot \{\tilde{w}\} \quad (1.11)$$

Non resta, quindi, che effettuare la scalatura, ultimo passo del procedimento complessivo. Il senso di questa operazione è da ricercarsi nell'effetto della prospettiva stessa: un oggetto in una certa posizione, quando viene rappresentato in

prospettiva, appare più o meno “grande” a seconda della sua distanza dal punto di osservazione.

Si definisce, il vettore $\{\tilde{p}\}$ di *proiezione prospettica*, che si ottiene a partire dal vettore $\{w_r\}$, scalandolo della distanza d ; anche $\{\tilde{p}\}$ avrà una coordinata sempre nulla a seconda di quale piano coordinato, nel riferimento della telecamera, sarà assunto come piano di proiezione:

$$\{\tilde{p}\} = \begin{Bmatrix} w_\xi \\ w_\eta \\ w_\zeta \\ d \end{Bmatrix} = \begin{Bmatrix} (n_\eta + n_\zeta) \cdot w_\xi \\ (n_\xi + n_\zeta) \cdot w_\eta \\ (n_\xi + n_\eta) \cdot w_\zeta \\ \{\tilde{N}\}^T \cdot \{\tilde{w}\} \end{Bmatrix} = [T_p] \cdot \{\tilde{w}\} \quad (1.12)$$

La matrice $[T_p]$ viene detta matrice di trasformazione prospettica; sia assegnato un vettore contenente le coordinate di un punto nel riferimento $\{O,x,y,z\}$, essa restituisce le coordinate dello stesso punto nel piano immagine, opportunamente scalate di un valore pari alla distanza d del punto dal piano.

È interessante effettuare la decomposizione nelle due matrici:

$$[T_p] = [S] \cdot [T_{OL}] = \begin{bmatrix} (n_\eta + n_\zeta) & 0 & 0 & 0 \\ 0 & (n_\xi + n_\zeta) & 0 & 0 \\ 0 & 0 & (n_\xi + n_\eta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ \zeta_x & \zeta_y & \zeta_z & t_\zeta \\ N_x & N_y & N_z & t_n \end{bmatrix} \quad (1.13)$$

La matrice $[T_{OL}]$ effettua un cambio di coordinate ed una scalatura in “3D”, così come un olografo, attraverso la luce laser, costruisce un’immagine scalata dell’oggetto reale in 3 dimensioni. Nell’ipotesi che il versore \hat{n} coincida con l’asse ζ , l’espressione della $[T_{OL}]$ diventa:

$$[T_p] = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ N_x & N_y & N_z & t_n \end{bmatrix} \quad (1.14)$$

e conseguentemente l'espressione del vettore di proiezione prospettica sarà:

$$\{\tilde{p}\} = \{w_\xi, w_\eta, 0, d\}^T$$

È evidente in questa espressione una perdita di informazione, che si traduce sempre in una riga nulla, che dipende da quale piano coordinato è scelto come piano di proiezione: infatti, delle tre componenti del versore normale $\{n\}_r$, una è sempre uguale a 1, mentre le altre due sono nulle. Questo comporta la perdita di rango della matrice $[S]$, e quindi, indirettamente, la non invertibilità della $[T_p]$, ovvero l'impossibilità di ricavare le coordinate nello spazio del punto P a partire dalle sue coordinate nell'immagine.

Tale matrice rappresenta, dunque, la formalizzazione matematica del fatto che, nella rappresentazione prospettica della realtà, si ha sempre la perdita dell'informazione riguardante la profondità, ed è, pertanto, impossibile riconoscere un oggetto a partire da una sua immagine bidimensionale [1,5,6].

1.6 Il modello Pin-Hole e sua discretizzazione

Una telecamera è un dispositivo in grado di catturare in un'immagine piana una porzione di spazio detto *campo*. Nel caso in cui, note le coordinate spaziali di un punto, mediante un opportuno formalismo, si riescano a calcolare le coordinate dello stesso punto nel piano dell'immagine, è possibile concludere che è stato trovato un modello matematico della telecamera, e che il formalismo adottato ne è la rappresentazione; di fatto, sviluppare un modello della telecamera non significa altro che farne una simulazione. I risultati di tale simulazione non potranno, tuttavia, coincidere perfettamente con ciò che avviene realmente, in quanto per tutti i modelli ci sarà sempre un grado di approssimazione e di semplificazione, senza cui non sarebbe possibile elaborare un metodo di calcolo.

Il modello più semplice in assoluto è il modello pin-hole, ovvero il modello a foro stenopeico, che approssima l'ottica della telecamera a quella di una camera oscura.

Si consideri un punto nello spazio P , di coordinate $\{w_\xi, w_\eta, w_\zeta\}^T$, e la sua proiezione su R attraverso C , di coordinate $P' = \{u, v\}^T$. Ovviamente P' è espresso nel riferimento immagine, mentre P rappresenta il punto dello spazio riferito alla terna telecamera.

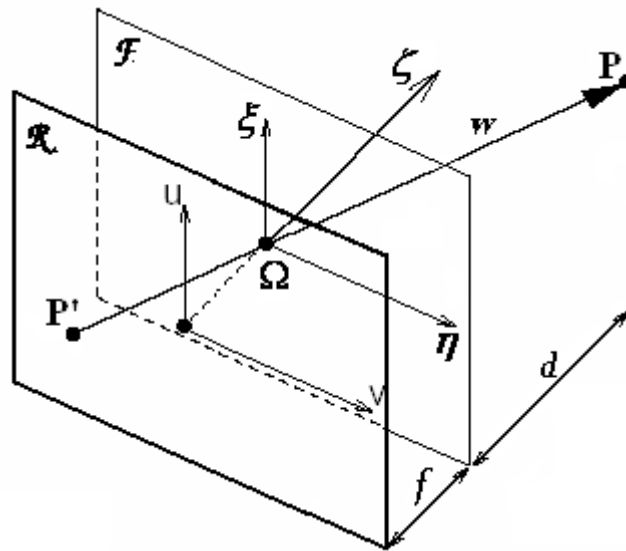


Fig. 1.8

Si possono scrivere le semplici relazioni di proiezione prospettica, ponendo il piano immagine davanti al centro di proiezione (eliminando così il segno negativo):

$$\begin{cases} u = f \cdot \frac{w_\xi}{w_\zeta} \\ v = f \cdot \frac{w_\eta}{w_\zeta} \end{cases} \quad (1.15)$$

In altri termini, si è resa la trasformazione che il vettore posizione di un punto P , in un riferimento parallelo a quello del piano immagine, subisce attraverso una lente sottile: essa consiste in una scalatura mediante un fattore pari a d/f , dove d rappresenta la distanza del punto P dal piano della telecamera (e nel caso semplice di $\hat{n} \equiv \zeta$, risulta $d=w_\zeta$). Ciò è in perfetta analogia con la trasformazione prospettica illustrata precedentemente.

Se si assume, per semplicità, che il versore \hat{n} coincide con l'asse ξ , senza per questo ledere la generalità dei risultati, si ottiene la relazione:

$$\{\tilde{p}\} = \begin{Bmatrix} w_\xi^* \\ w_\eta^* \\ 0 \\ d \end{Bmatrix} = [T_p] \cdot \{\tilde{w}\} = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ N_x & N_y & N_z & t_n \end{bmatrix} \cdot \begin{Bmatrix} w_x \\ w_y \\ w_z \\ 1 \end{Bmatrix} \quad (1.16)$$

già vista nel precedente paragrafo (relazione 1.13).

In definitiva per tener conto della trasformazione operata dalla lente sottile, che ricordiamo opera una scalatura mediante un fattore pari a d/f , introduciamo una nuova matrice, $[F]$ che chiameremo *matrice focale*:

$$\begin{Bmatrix} u^* \\ v^* \\ 0 \\ d/f \end{Bmatrix} = [F] \cdot \{\tilde{p}\} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/f \end{bmatrix} \cdot \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ N_x & N_y & N_z & t_n \end{bmatrix} \cdot \begin{Bmatrix} w_x \\ w_y \\ w_z \\ 1 \end{Bmatrix} \quad (1.17)$$

Tale relazione consente di calcolare le coordinate (u,v) di un *punto-immagine* P' a partire dalle coordinate di un *punto-scena* P . Questa espressione è, però, ancora incompleta, in quanto non tiene conto della “discretizzazione” che si ottiene per effetto della dimensione finita dei sensori [2,5,6].

La discretizzazione tiene conto del fatto che:

1. Il centro ottico della telecamera non coincide con il centro fisico del CCD ma ha coordinate (u_0, v_0) .
2. Le coordinate di un punto nel sistema di riferimento standard della telecamera sono misurate in pixel: si deve introdurre quindi un fattore di scala.
3. La forma dei pixel non è quadrata: occorre considerare due fattori di scala diversi, lungo u (δ_u) e lungo v (δ_v).
4. A causa di sfasamenti nella scansione di righe successive dello schermo, gli assi di riferimento immagine u,v non sono ortogonali ma inclinati di ϑ .

I primi tre punti vengono presi in esame mediante l'introduzione, nelle formule (1.15), della traslazione del centro ottico e della riscalatura *indipendente* degli assi u e v :

$$\begin{cases} u = \frac{1}{\delta_u} \cdot f \cdot \frac{w_\xi}{w_\zeta} + u_0 \\ v = \frac{1}{\delta_v} \cdot f \cdot \frac{w_\eta}{w_\zeta} + v_0 \end{cases} \quad (1.18)$$

dove si sono indicati con

- ⌘ (u_0, v_0) le coordinate del punto principale;
- ⌘ $\delta_u (\delta_v)$ la dimensione efficace del pixel lungo la direzione u (v); le dimensioni fisiche di δ sono in $pixel \cdot m^{-1}$.

Dopo questo aggiornamento delle equazioni, il modello proposto, sempre nell'ipotesi semplificativa che il versore \hat{n} coincida con l'asse ζ , diventa:

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} = \frac{1}{d} \cdot \begin{bmatrix} f/\delta_u & 0 & 0 & u_0 \\ 0 & f/\delta_v & 0 & v_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{Bmatrix} w_\xi \\ w_\eta \\ 0 \\ d \end{Bmatrix} = [K] \cdot \{\tilde{p}\} \quad (1.19)$$

La matrice $[K]$ è detta *matrice dei parametri intrinseci* della telecamera, poiché i suoi elementi dipendono solo dalla distanza focale e dalle caratteristiche del sensore di visione.

Il modello più generale in realtà deve prevedere anche la possibilità che gli assi u, v non siano ortogonali, ma inclinati di un angolo ϑ . La matrice K nella forma più generale possibile è, quindi, la seguente:

$$[K] = \begin{bmatrix} f/\delta_u & (f/\delta_u) \cdot \cot \vartheta & 0 & u_0 \\ 0 & f/(\delta_v \cdot \sin \vartheta) & 0 & v_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1.20)$$

Nei moderni sistemi di acquisizione si è cercato di attenuare la disomogeneità lungo u e v della matrice di CCD, per cui si può considerare $\vartheta = \pi/2$, così in questo caso la (1.20) coincide con la matrice K presente nella (1.19).

In definitiva, il modello “*pin-hole*” della telecamera trova la sua espressione generale nella relazione:

$$\begin{Bmatrix} u \\ v \\ w \\ 1 \end{Bmatrix} = \frac{1}{d} \cdot [K] \cdot [S] \cdot [T_{OL}] \cdot \{\tilde{w}\} = [M] \cdot \{\tilde{w}\} \quad (1.21)$$

A primo membro di tale espressione, compare il vettore $\{u, v, w, 1\}$, a giustificazione del fatto che nel caso più generale possibile, il piano di proiezione può essere uno qualunque tra i piani coordinati del riferimento telecamera $\{\Omega, \xi, \eta, \zeta\}$. Nell’ipotesi che il versore \hat{n} coincida con l’asse ζ , l’espressione del modello diventa:

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} = \frac{1}{d} \cdot \begin{bmatrix} f/\delta_u & 0 & 0 & u_0 \\ 0 & f/\delta_v & 0 & v_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ \zeta_x & \zeta_y & \zeta_z & t_\zeta \\ N_x & N_y & N_z & t_n \end{bmatrix} \cdot \begin{Bmatrix} w_x \\ w_y \\ w_z \\ 1 \end{Bmatrix} \quad (1.22)$$

e semplificando ulteriormente si ha:

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} = \frac{1}{d} \cdot \begin{bmatrix} f/\delta_u & 0 & 0 & u_0 \\ 0 & f/\delta_v & 0 & v_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ N_x & N_y & N_z & t_n \end{bmatrix} \cdot \begin{Bmatrix} w_x \\ w_y \\ w_z \\ 1 \end{Bmatrix} = \frac{1}{d} \cdot [K] \cdot [T_P] \cdot \{\tilde{w}\} \quad (1.23)$$

tale relazione può essere scritta anche nella seguente forma

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} = [M] \cdot \{\tilde{w}\} \Leftrightarrow [\tilde{\phi}] = [M] \cdot \{\tilde{w}\} \quad (1.24)$$

avendo indicato con $[\tilde{\phi}]$ il vettore colonna $\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix}$.

In tale espressione:

- ↖ $[K]$ è la matrice dei parametri intrinseci della telecamera; contiene informazioni sull'ottica e la sensoristica della telecamera.
- ↖ $[T_{OL}]$ è la matrice dei parametri estrinseci della telecamera; tali parametri contengono le informazioni riguardanti la sua configurazione cinematica nello spazio.

Un'altra importante osservazione è relativa alla matrice $[S]$ che compare nell'espressione della $[T_P]$. Gli elementi di tale matrice, pur avendo un significato geometrico ben preciso (sono le componenti del versore normale al piano retina), non vengono considerati come parametri estrinseci, poiché il loro valore è binario, dipende solo dalle convenzioni adottate.

Nell'espressione più compatta $\{\tilde{\phi}\} = [M] \cdot \{\tilde{w}\}$, dove con $\{\tilde{\phi}\}$ indichiamo il vettore che individua le coordinate del punto nel piano immagine, la matrice $[M]$ governa la proiezione dei punti nello spazio (riferiti al sistema mondo), sul piano immagine. È da notare che M è composta, da 12 elementi, ma dipende da 11 parametri indipendenti. Infatti, è possibile, ricorrendo ad una normalizzazione della matrice, eliminare la dipendenza da un parametro.

Per omogeneità di notazione con i principali algoritmi di calibrazione della telecamera, la matrice $[K]$ può essere scritta nel seguente modo:

$$[K] = \begin{bmatrix} f/\delta_u & 0 & 0 & u_0 \\ 0 & f/\delta_v & 0 & v_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & 0 & u_0 \\ 0 & f_v & 0 & v_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1.25)$$

In questi termini, la distanza focale può essere espressa in pixel e rappresentata dal vettore $\{f\} = \{f_u \ f_v\}^T$. Le ragioni di tale posizione sono da ricercarsi nella maggiore praticità nell'utilizzare un numero di parametri in gioco inferiore, utilizzando in luogo dei 3 parametri f, δ_u, δ_v solo le due componenti del vettore $\{f\}$ [3,5,6].

1.7 Modello della distorsione a cinque parametri

Il modello pinhole della telecamera ad 11 parametri non è ancora sufficiente a descrivere in maniera completa la trasformazione geometrica che subiscono le immagini acquisite. Le considerazioni finora fatte sono state desunte dal modello delle lenti sottili: ma le lenti sottili costituiscono un'astrazione matematica.



Fig. 1.9

Per quanto i progressi compiuti dall'ottica abbiano portato alla costruzione di lenti sempre più prossime a queste ipotesi di idealità, in esse è sempre presente una componente di distorsione radiale (vedi foto), e una tangenziale.

Tra le due, la componente dominante è quella radiale, che è dovuta alle dimensioni reali della lente, mentre la distorsione tangenziale è più lieve, ed è causata dal decentramento dell'asse della lente.

Ragionando in un sistema di coordinate polari con origine nel centro dell'immagine, la distorsione radiale agisce sui pixel dell'immagine in funzione della sola distanza dal centro ρ , mentre quella tangenziale dipende anche dall'angolo ϑ .

Negli ultimi anni sono stati sviluppati diversi modelli matematici di distorsione al fine di correggerla e tornare così alle ipotesi di lente sottile: tale processo di correzione prende il nome di *compensazione della distorsione*.

L'inclusione degli effetti prodotti dalla distorsione nel modello può renderlo più preciso, anche se i suoi effetti non sono apprezzabili, soprattutto nel caso in cui la maggior parte degli oggetti inquadrati si trova ad una distanza non troppo piccola.

La distorsione può essere pensata come una funzione non lineare che si applica ai singoli pixel dell'immagine originaria, non distorta, di coordinate (u, v) , che dipende, oltre che dalle stesse coordinate, anche dalla distanza dal centro $\rho = \sqrt{u^2 + v^2}$, e dall'angolo $\vartheta = \tan^{-1}(v/u)$.

Se indichiamo con u_d e v_d le coordinate distorte, la relazione che esprime la trasformazione da quelle non distorte è:

$$\begin{cases} u_d \\ v_d \end{cases} = f(u, v, \rho, \vartheta) \quad (1.26)$$

La maggior parte dei modelli matematici è basata sulla stima dei coefficienti dello sviluppo in serie di Taylor di f , arrestato al secondo ordine, e non tengono quasi mai in conto della distorsione tangenziale: pertanto forniscono, pertanto, soltanto due parametri di distorsione, che sono sufficienti a correggere fenomeni di aberrazione molto lievi, risultando, comunque invece inefficaci per alterazioni enormi come quella in figura 1.9.

L'algoritmo di compensazione, sia della distorsione radiale che della distorsione tangenziale, di seguito presentato è basato su cinque parametri, piuttosto che due: tre per la distorsione radiale e due per quella tangenziale.

Il modello utilizzato è stato sviluppato dalla Oulu University (Finlandia), ed è lo stesso adottato dal Matlab Calibration ToolBox.

La trasformazione operata dall'algoritmo è, dunque, la seguente:

$$\begin{Bmatrix} u_d \\ v_d \end{Bmatrix} = \underbrace{(1 + kc_1 \cdot \rho^2 + kc_2 \cdot \rho^4 + kc_5 \cdot \rho^6)}_{\text{applica la distorsione radiale}} \cdot \begin{Bmatrix} u \\ v \end{Bmatrix} + \underbrace{\begin{Bmatrix} 2 \cdot kc_3 \cdot u \cdot v + kc_4 \cdot (\rho^2 + 2 \cdot u^2) \\ 2 \cdot kc_4 \cdot u \cdot v + kc_3 \cdot (\rho^2 + 2 \cdot u^2) \end{Bmatrix}}_{\text{applica la distorsione tangenziale}}$$

applica la distorsione radiale

applica la distorsione tangenziale

dove kc sarà convenzionalmente indicato come un vettore di cinque componenti così ordinate: $kc = \{kc_1, kc_2, kc_3, kc_4, kc_5\}$.

Il parametro più importante è kc_1 , perché è quello che influenza maggiormente l'entità della distorsione:

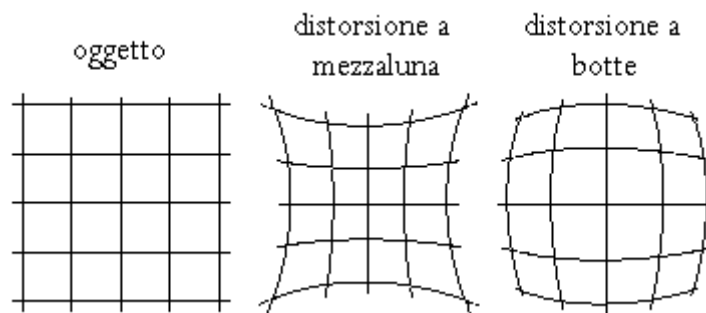


Fig. 1.10

se $kc_1 > 0$, si ottiene una distorsione di tipo mezzaluna; se invece $kc_1 < 0$ si ha una distorsione a botte, entrambe illustrate in figura 1.10.

I cinque parametri di distorsione si aggiungono ai parametri intrinseci della telecamera e, per una maggiore precisione, vanno presi in considerazione nella calibrazione delle telecamere.

È da notare che se si fossero considerati i soli due termini della sola distorsione radiale (come nella maggior parte dei sistemi di visione), la relazione finale che identifica i parametri di distorsione sarebbe stata molto più semplice:

$$\begin{Bmatrix} u_d \\ v_d \end{Bmatrix} = (1 + kc_1 \cdot \rho^2 + kc_2 \cdot \rho^4) \cdot \begin{Bmatrix} u \\ v \end{Bmatrix} \quad (1.27)$$

Tale relazione è computazionalmente più veloce, ma consente una minore precisione, per le ragioni già esposte; quest'ultima, infatti, è la stessa che è utilizzata dalla maggior parte di fotocamere digitali in commercio per correggere velocemente l'aberrazione radiale della lente [3,4].

CAPITOLO 2

Panoramica dei metodi ottici per l'acquisizione di forme tridimensionali

2.1 Introduzione [2]

In questo capitolo si vogliono descrivere le tecniche computazionali per stimare le proprietà geometriche (la struttura) del mondo tridimensionale (3D) a partire da proiezioni bidimensionali (2D) di esso: le immagini. Verrà introdotto il problema della acquisizione della forma (ovvero shape/model acquisition, image-based modeling, 3D photography) e si descriveranno sommariamente i passi che è necessario effettuare a valle della misura 3D per ottenere veri e propri modelli degli oggetti.

Esistono molti metodi per l'acquisizione automatica della forma di un oggetto. Una possibile tassonomia dei metodi di acquisizione della forma (shape acquisition) è illustrata in figura 2.1.

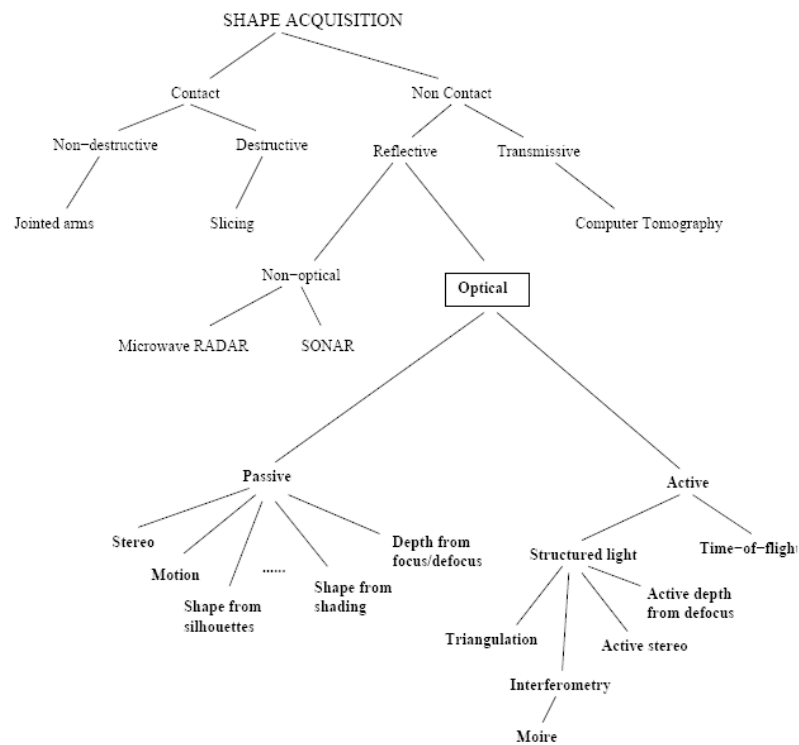


Fig. 2.1

In seguito si farà riferimento a quelli ottici riflessivi, visto che ci si vuole interessare della visione. Tra i vantaggi di queste tecniche vi è il fatto di non richiedere contatto, la velocità e la economicità. Le limitazioni includono il fatto di poter acquisire solo la parte visibile delle superfici e la sensibilità alle proprietà delle superfici come trasparenza, brillantezza, colore..

Il problema che si vuole affrontare, noto come image-based modeling o 3D photography si pone nel modo seguente: gli oggetti irradiano luce visibile; la telecamera cattura questa luce, le cui caratteristiche dipendono dalla illuminazione della scena, geometria della superficie, riflettanza della superficie, la quale viene analizzata al computer con lo scopo di inferire la struttura 3D degli oggetti. Nella figura 2.2 viene riportata una apparecchiatura per l'acquisizione di forma mediante due immagini.



Fig. 2.2

La distinzione fondamentale tra le tecniche ottiche riguarda l'impiego o meno di sorgenti di illuminazione speciali. In particolare, si distinguono i metodi attivi, che irradiano la scena con radiazioni elettromagnetiche opportune (pattern luminosi, luce laser, radiazioni infrarosse etc...), ed i metodi passivi, che si basano sull'analisi di immagini (di colore) della scena così com'è. I primi hanno il vantaggio di raggiungere risoluzioni elevate, ma sono più costosi e non sempre applicabili. I

secondi sono economici, pongono meno vincoli per l'applicabilità ma hanno risoluzioni più basse.

Il sistema visivo umano sfrutta molteplici "indizi visivi" (visual cues) per recuperare la profondità delle superfici della scena, informazione che viene persa nel processo di proiezione sulla retina (bidimensionale). Alcuni di questi sono:

- lo sfocamento
- il parallasse (o disparità, cioè la differenza tra viste differenti della scena)
- il chiaroscuro (shading),
- le tessiture.

Molti dei metodi ottici per l'acquisizione della forma restituiscono una immagine range, ovvero una immagine nella quale ciascun pixel contiene la distanza dal sensore ad un punto visibile della scena, invece della sua luminosità (figura 2.3). Una immagine range è costituita da misure (discrete) di una superficie 3D rispetto ad un piano 2D (il piano immagine del sensore, di solito) e pertanto prende anche il nome di immagine 2.5D. La superficie si potrà sempre esprimere nella forma $Z = f(X, Y)$, se il piano di riferimento è identificato con il piano XY . Un sensore range è un dispositivo che produce una immagine range.



Fig. 2.3

Nel seguito si parlerà di sensori range ottici in senso generalizzato, per intendere qualunque sistema ottico di acquisizione della forma, attivo o passivo, composto da

apparecchiature e programmi, che restituisce una immagine range della scena. La bontà di un sensore range si valuta prendendo in considerazione (ma non solo):

- **risoluzione:** il più piccolo cambiamento di profondità che il sensore può rilevare;
- **accuratezza:** differenza tra valore misurato (media di misure ripetute) e valore vero (misura l'errore sistematico);
- **precisione:** variazione statistica (deviazione standard) di misure ripetute di una stessa quantità (misura la dispersione delle misure attorno alla media);
- **velocità:** misure al secondo.

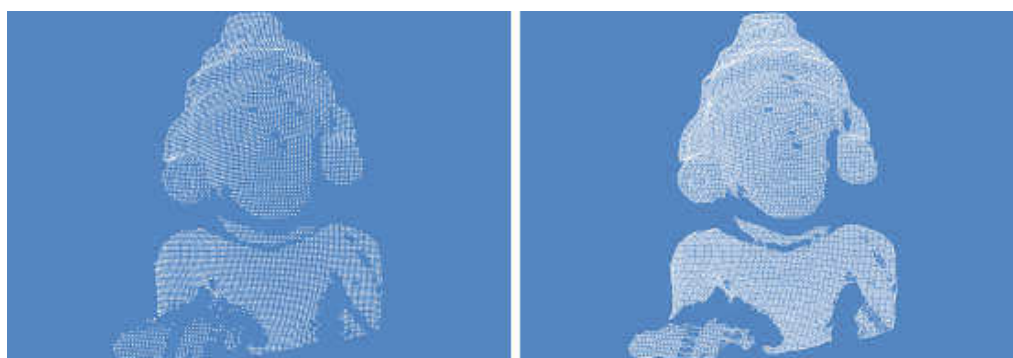
2.2 Dalla misura al modello 3D

Il recupero della informazione 3D non esaurisce però il processo di acquisizione della forma, per quanto ne costituisca il passo fondamentale. Per ottenere un modello completo di un oggetto, o di una scena, servono molte immagini range, che devono essere allineate e fuse tra di loro, ottenendo una superficie 3D (in forma di maglia poligonale). La ricostruzione del modello dell'oggetto a partire da immagini range (o viste) prevede dunque tre fasi:

- **registrazione:** (o allineamento) per trasformare le misure fornite dalle varie immagini range in un unico sistema di riferimento comune;
- **fusione geometrica:** per ottenere una singola superficie 3D (tipicamente una maglia poligonale) dalle varie immagini range;
- **semplificazione della maglia:** tipicamente i punti restituiti da un sensore range sono troppi per avere un modello maneggevole. La maglia finale risultante deve essere semplificata in maniera opportuna.

Nel seguito sarà descritta soprattutto la prima fase, si tratterà sommariamente la seconda e si tralascerà la terza. Questo perché questi argomenti sfumano gradualmente nella Grafica Computazionale ed esulano dunque dagli scopi introduttivi di questo capitolo.

Come detto in precedenza una immagine range $Z(X,Y)$ definisce un insieme di punti 3D del tipo $(X,Y,Z(X,Y))$, come in figura 2.4. Per ottenere una superficie nello spazio 3D (superficie range) è sufficiente connettere tra loro i punti più vicini con facce triangolari (figura 2.4b).



a)

b)

Fig. 2.4

In molti casi esistono discontinuità di profondità (buchi, bordi occludenti) che non si vuole vengano coperte dai triangoli che potrebbero crearsi. Per evitare di fare assunzioni non giustificate sulla forma della superficie, si deve evitare di collegare punti che sono troppo distanti (nel 3D). Per questo è opportuno eliminare triangoli con lati eccessivamente lunghi e quelli con angoli eccessivamente acuti.

2.2.1 Registrazione

I sensori range tipicamente non catturano la forma di un oggetto con una sola immagine, ne servono molte, ciascuna delle quali cattura una parte della superficie dell'oggetto. Le porzioni della superficie dell'oggetto che si ottengono da diverse immagini range sono espresse ciascuna nel proprio sistema di riferimento (legato alla posizione del sensore). Lo scopo della registrazione o allineamento è di portarle nello stesso sistema di riferimento tramite una opportuna trasformazione rigida (rotazione e traslazione nel 3D).

Se la posizione ed orientazione del sensore sono note (perchè vengono rilevate oppure perchè esso è controllato meccanicamente), il problema si risolve banalmente. Tuttavia in molti casi, le trasformazioni vanno calcolate usando i soli

dati, a tale proposito esistono vari algoritmi che risolvono questo problema, uno dei quali è chiamato ICP (Iterated Closest Point).

Nella figura 2.5 sono riportate a sinistra otto immagini range di un oggetto ognuna nel proprio sistema di riferimento mentre a destra, tutte le immagini sono state sovrapposte dopo la registrazione.

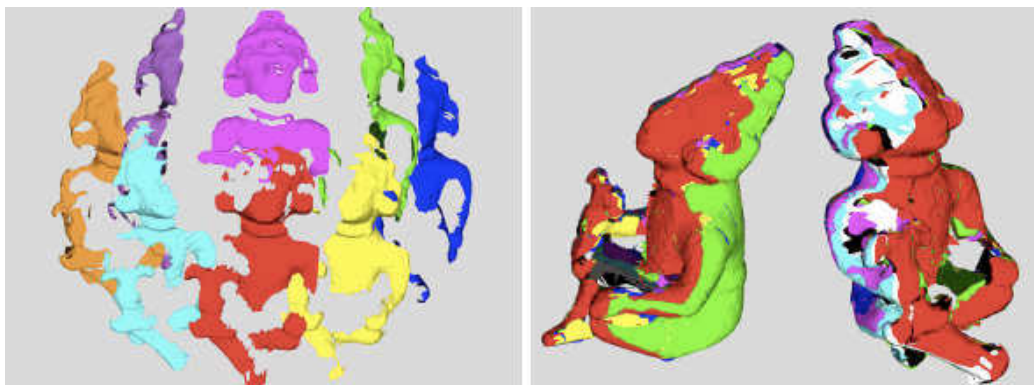


Fig. 2.5

2.2.2 Fusione geometrica

Una volta che tutti i dati range sono stati precisamente registrati in un sistema di coordinate comune, si devono fondere i dati in una singola forma, rappresentata, ad esempio, da una maglia triangolare. Questo è il problema della ricostruzione della superficie, che si può formulare come la stima di una varietà bidimensionale che approssimi la superficie dell'oggetto incognito da un insieme di punti 3D campionati. Si possono suddividere i metodi di fusione geometrica in due categorie:

- **Integrazione di maglie:** vengono unite le maglie triangolari corrispondenti alle singole superfici range.
- **Fusione volumetrica:** i dati vengono fusi in una rappresentazione volumetrica, dalla quale si estrae poi una maglia triangolare.

2.2.2.1 Integrazione di maglie

Le tecniche di integrazione di maglie mirano ad unire più maglie triangolari 3D sovrapposte in un'unica maglia triangolare (sfruttando la rappresentazione in termini di superficie range).

Il metodo di Turk e Levoy (1994) unisce maglie triangolari sovrapposte impiegando una tecnica detta di zippering. Le maglie sovrapposte vengono erose fino ad eliminare la sovrapposizione e poi si impiega una triangolazione nel 2D per ricucire i bordi. Questo richiede che i punti delle due superfici 3D prossimi ai bordi da ricucire vengano proiettati su di un piano 2D.

Nella figura 2.6 sono riportate a sinistra due superfici registrate e a destra il risultato dello zippering.

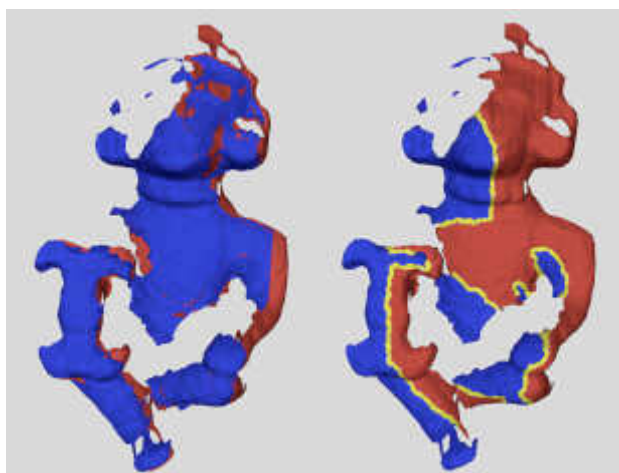


Fig. 2.6

Le tecniche di integrazione di maglie permettono la fusione di più immagini range senza perdere nell'accuratezza, poiché i vertici della maglia finale coincidono con i punti dei dati misurati. Per lo stesso motivo sono sensibili a misurazioni sbagliate, che possono provocare problemi nella ricostruzione della superficie.

2.2.2.2 Fusione volumetrica

La fusione volumetrica di misurazioni di superficie costruisce una superficie implicita intermedia che unisce le misurazioni sovrapposte in un'unica rappresentazione. La rappresentazione della superficie implicita è una iso-superficie di un campo scalare $f(x,y,z)$. Per esempio, se la funzione di campo è definita come la distanza del punto più vicino sulla superficie dell'oggetto, allora la superficie implicita è rappresentata da $f(x,y,z)=0$. Questa rappresentazione permette la modellazione di oggetti dalla forma sconosciuta con topologia e geometria arbitrari.

Per passare dalla rappresentazione implicita della superficie ad una maglia triangolare, si usa l'algoritmo Marching Cubes, sviluppato da Lorensen e Cline (1987) per la triangolazione di iso-superfici a partire dalla rappresentazione discreta di un campo scalare (come le immagini 3D nel campo medico). Lo stesso algoritmo torna utile per ottenere una superficie triangolata a partire da ricostruzioni volumetriche della scena (shape from silhouette e photo consistency).

Il metodo di Hoppe e altri (1992) trascura la struttura dei dati (cioè la superficie range) e calcola una superficie a partire dalla "nuvola" non strutturata dei punti. Curless e Levoy (1996) invece sfruttano l'informazione contenuta nelle immagini range per assegnare a "vuoto" i voxel che giacciono lungo la linea di vista che da un punto della superficie range arriva fino al sensore. Una limitazione evidente di tutti gli algoritmi di fusione geometrica basati su una struttura intermedia di dati volumetrici discreti è una riduzione nell'accuratezza, che provoca la perdita di dettagli della superficie.

Inoltre lo spazio necessario per la rappresentazione volumetrica cresce rapidamente al crescere della risoluzione.

2.3 Metodi ottici per l'acquisizione di forme

Le tecniche computazionali che saranno elencate di seguito, suddivise tra attive e passive, sfruttano degli indizi per recuperare la forma degli oggetti dalle immagini. Di seguito sono elencati i metodi suddivisi tra attivi e passivi.

Metodi ottici passivi:

- depth from focus/defocus
- shape from texture
- shape from shading
- stereo fotometrico
- stereopsi
- shape from silhouette
- shape from photo-consistency
- structure from motion

Metodi ottici attivi:

- active defocus
- stereo attivo
- triangolazione attiva
- interferometria
- tempo di volo

Tutti i metodi attivi tranne l'ultimo impiegano una o due telecamere ed una sorgente di luce speciale, e rientrano nella classe più ampia dei metodi ad **illuminazione strutturata**.

2.3.1 Depth from focus/defocus

Nella tecnica depth from focus/defocus si hanno due casi distinti dove si sfruttano le informazioni sulla messa a fuoco per la ricostruzione della profondità: la fochettatura e la misura dello sfocamento.

2.3.1.1 Fochettatura

Si vuole determinare la distanza da un punto prendendo diverse immagini con una messa a fuoco sempre migliore, finché il punto di interesse è perfettamente a fuoco. Nella letteratura in inglese la tecnica prende il nome di depth from focus. Sapere che un punto è a fuoco fornisce informazioni sulla sua profondità tramite la legge delle lenti sottili. Questo metodo richiede una ricerca in tutte le posizioni del fuoco per ogni punto nell'immagine (lento!). I problemi chiave nel recupero della distanza tramite fochettatura sono:

- la scelta del criterio di messa a fuoco
- la ricerca efficiente della posizione di fuoco ottimo (secondo il criterio fissato)

2.3.1.2 Sfocamento

Prendendo un piccolo numero di immagini (al minimo 2) con diversi parametri di fuoco, è possibile determinare la profondità di tutti i punti nella scena. Il metodo si basa sulla relazione diretta fra la profondità, i parametri dell'ottica e lo sfocamento

misurabile nell'immagine. Si chiama anche depth from defocus. I problemi nel recupero della distanza dallo sfocamento sono:

- la misurazione dello sfocamento (di solito stimato confrontando rappresentazioni frequenziali delle immagini)
- la calibrazione della relazione tra profondità e sfocamento.

Nella figura 2.7 sopra sono riportate due immagini di una scena prese con fuoco diverso, sotto a sinistra l'immagine range ottenuta con il metodo depth from defocus e a destra la superficie ricostruita.

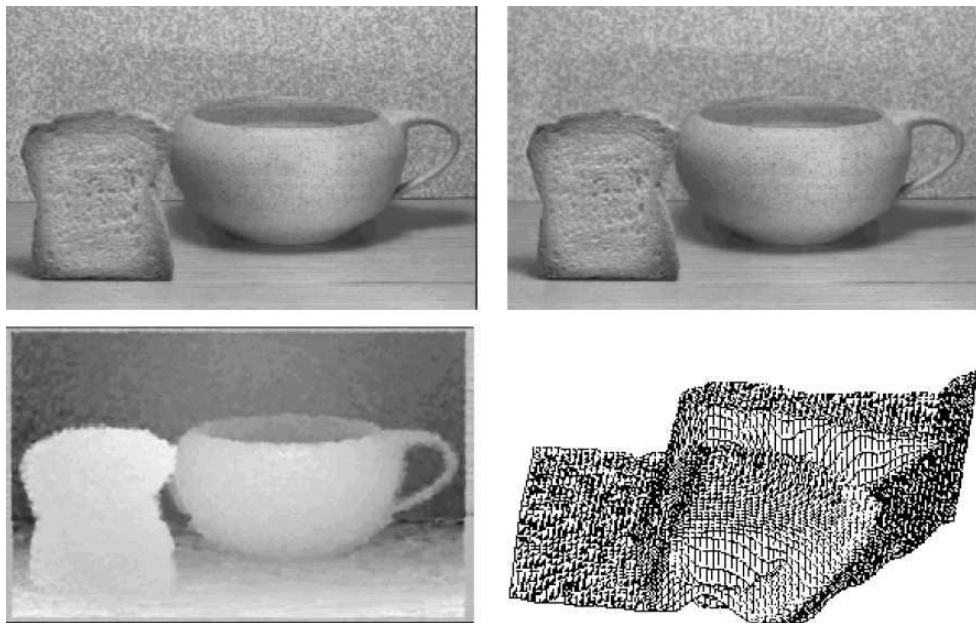


Fig. 2.7

2.3.2 Shape from texture

La tessitura, nel vocabolario italiano, si riferisce alle rocce ed indica il modo con cui si riuniscono i cristalli che la costituiscono (per esempio diversi tipi di marmo hanno diversa tessitura). In visione si riferisce ad un concetto simile, ovvero al modo in cui si dispongono pattern spaziali ripetuti su di una superficie. Alcuni esempi sono (figura 2.8): la disposizione dei mattoni sulla facciata di un edificio, le macchie sulla pelle del leopardo, i fili d'erba in un prato, i sassi su una spiaggia, le teste delle persone nella folla, i disegni della tappezzeria. Gli elementi che si ripetono (chiamati

anche texel) devono essere abbastanza piccoli da non poter essere distinti come oggetti separati (un sasso visto da vicino è un oggetto, quando fa parte della vista di una spiaggia è un texel). Alle volte la disposizione è regolare (o periodico), come nella tappezzeria o nella facciata a mattoni, a volte la regolarità è solo statistica, come nel prato o nella spiaggia. Si distinguono infatti tra tessiture deterministiche (figura 2.8 a) e tessiture statistiche (figura 2.8 b). Spesso le prime derivano dall'opera dell'uomo, mentre le seconde si trovano in natura.

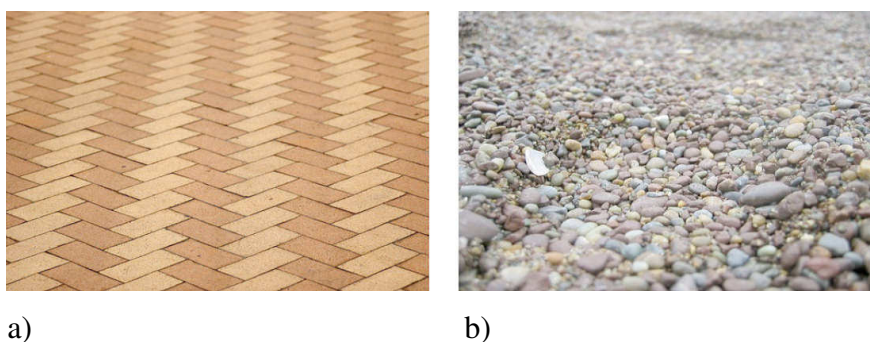


Fig. 2.8

Anche se i texel sono identici nella scena, in una immagine prospettica la loro dimensione apparente, forma, spaziatura etc. cambia, a causa della proiezione prospettica. Dunque l'immagine prospettica di una tessitura stesa (per esempio) su un piano subirà una distorsione dovuta alla scorciatoia, per cui per esempio:

- cambia la forma di ciascun texel, in ragione dell'angolo di inclinazione (slant) del piano (i cerchi diventano ellissi, tanto più eccentrici quanto più il piano è inclinato);
- cambia la dimensione apparente dei texel: più lontani sono, più piccoli appaiono (l'area delle ellissi decresce con la distanza).

Questi effetti sono legati alla orientazione ed alla posizione del piano e dunque ne costituiscono un indizio per la determinazione di informazioni spaziali della tessitura. Possiamo risalire alla sua orientazione misurando caratteristiche che quantificano la distorsione (per esempio eccentricità) subita dai texel, oppure la velocità di cambiamento (gradiente di tessitura) di certe caratteristiche del texel (per esempio variazione dell'area).

Parliamo di un piano, ma il discorso si può estendere ad una superficie generica che sarà approssimata come planare a tratti (maglia poligonale).

2.3.3 Shape from shading

Lo Shape from Shading (SFS), è quel processo che, data un'immagine di un oggetto (illuminato) consente di calcolarne la forma, sfruttando le informazioni connesse alla variazione della luminosità della superficie dell'oggetto. Il termine inglese shading si traduce con chiaroscuro o ombreggiatura o sfumatura. In altri termini, la determinazione della forma dal chiaroscuro può essere pensata come il problema di ricostruire una superficie nello spazio 3D, a partire da un'altra superficie, rappresentante la luminosità sul piano immagine della telecamera.

Come suggerisce la figura 2.9, la distribuzione dei livelli di grigio nell'immagine (il chiaroscuro) reca con sé un'informazione utile riguardante la forma della superficie e la direzione di illuminazione.

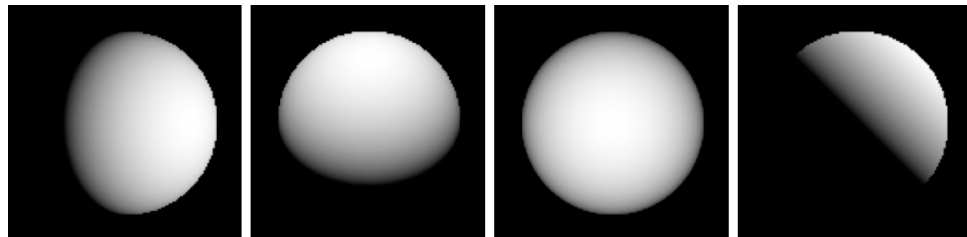


Fig. 2.9

2.3.4 Stereo fotometrico

Lo stereo fotometrico si basa sugli stessi principi dello SFS, cambiando la direzione di illuminazione. Con almeno tre immagini prese con diversa direzione di illuminazione, si risolvono alcuni problemi di indeterminazione che si hanno con il metodo SFS.

2.3.5 Stereopsi

La stereopsi (computazionale) è il processo che consente di ottenere informazioni sulla struttura tridimensionale da una coppia di immagini, provenienti da due telecamere che inquadrano una scena da differenti posizioni.

Si possono individuare due sottoproblemi:

- calcolo delle corrispondenze
- triangolazione.

Il primo consiste nell'accoppiamento tra punti nelle due immagini che sono proiezione dello stesso punto della scena (figura 2.10). Si definiscono tali punti coniugati.

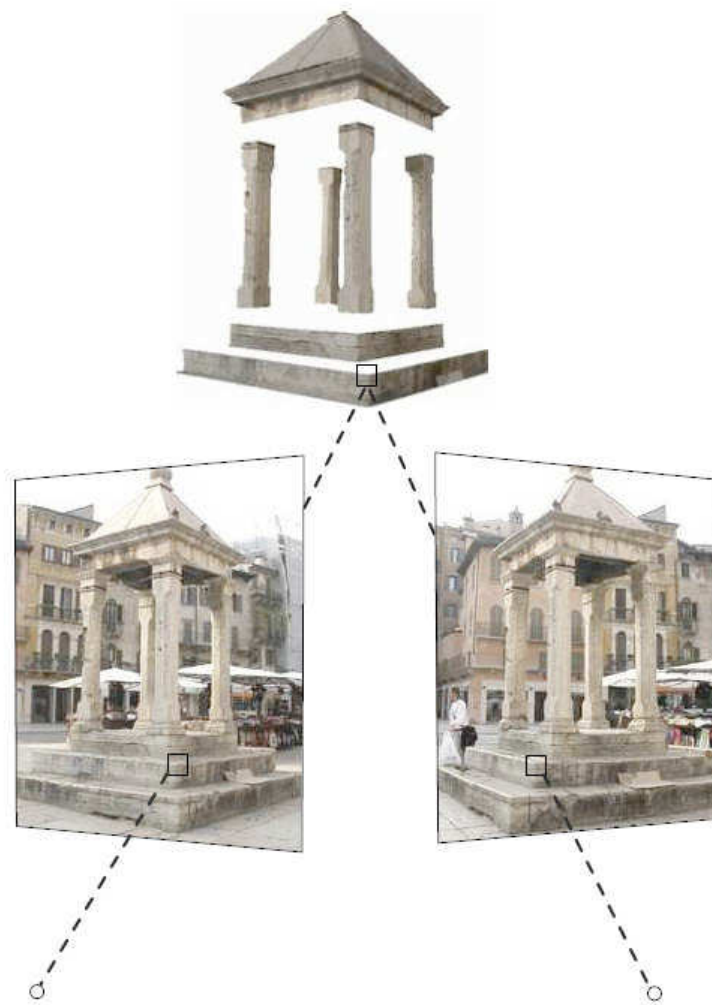


Fig. 2.10

Il calcolo dell'accoppiamento è possibile sfruttando il fatto che le due immagini differiscono solo lievemente, sicchè un particolare della scena appare simile nelle due immagini. Basandosi solo su questo vincolo, però, sono possibili molti falsi accoppiamenti. Si dimostra che è necessario introdurre altri vincoli che rendano il

calcolo delle corrispondenze trattabile. Il più importante di questi è il vincolo epipolare, il quale afferma che il corrispondente di un punto in una immagine può trovarsi solo su una retta (retta epipolare) nell'altra immagine. Grazie a questo la ricerca delle corrispondenze diventa unidimensionale, invece che bidimensionale.

Noti gli accoppiamenti tra i punti delle due immagini e nota la posizione reciproca delle telecamere ed i parametri interni del sensore è possibile ricostruire la posizione nella scena dei punti che sono proiettati sulle due immagini. Questo processo di triangolazione necessita della calibrazione dell'apparato stereo, ovvero del calcolo dei parametri interni (intrinseci) e della posizione reciproca (parametri estrinseci) delle telecamere.

2.3.6 Shape from silhouette

Intuitivamente una sagoma (silhouette) è il profilo di un oggetto, comprensivo della sua parte interna. Più precisamente si definisce sagoma una immagine binaria il cui valore in un certo punto (x,y) indica se il raggio ottico che passa per il pixel (x,y) interseca o meno la superficie di un oggetto nella scena. Un pixel della sagoma può essere un punto dell'oggetto (bianco) o un punto dello sfondo (nero).

Ogni punto bianco della sagoma identifica un raggio ottico che interseca l'oggetto ad una distanza ignota. L'unione di tutti i raggi ottici che attraversano pixel bianchi della sagoma definisce un cono generalizzato all'interno del quale è contenuto l'oggetto. L'intersezione dei coni generalizzati associati a tutte le telecamere identifica un volume all'interno del quale giace l'oggetto. Questo volume viene preso come una approssimazione dell'oggetto. La tecnica introdotta da Martin e Aggarwal (1983), basata sulla intersezione dei coni generati dalle sagome, prende il nome di shape from silhouette (figura 2.11).

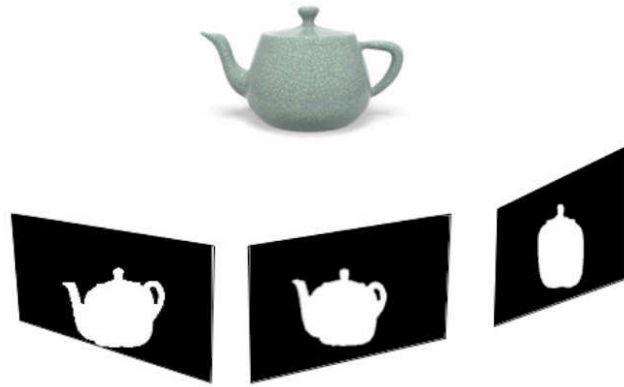


Fig. 2.11

Al fine di risparmiare memoria si sfrutta la struttura dati di octree per memorizzare le etichette dei voxel nello spazio campionato. Gli octree sono una rappresentazione più efficiente in termini di spazio quando la scena contiene ampie zone vuote. Si tratta di alberi a otto vie in cui ciascun nodo rappresenta un certo settore di spazio e i nodi figli sono le otto suddivisioni equispaziate di quel settore (ottanti).

Il visual hull, introdotto da Laurentini (1994), è la miglior approssimazione ottenibile da un numero infinito di sagome. Può essere definita come la forma massimale che restituisce le stesse sagome dell'oggetto reale per tutte le viste esterne al guscio convesso dell'oggetto (figura 2.12).

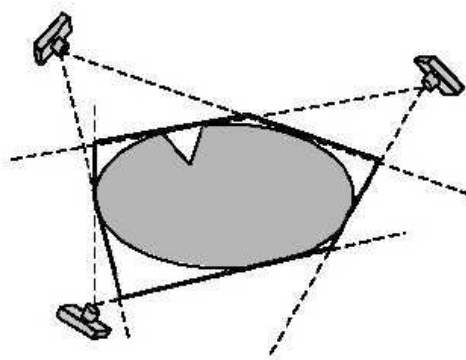


Fig. 2.12

Nella pratica è disponibile solo un numero finito di sagome e quello che si ottiene è una approssimazione del visual hull. Il volume dell'approssimazione decresce all'aumentare del numero di viste.

Nella figura 2.13 è riportato il risultato della ricostruzione volumetrica ottenuta con 12 sagome di una teiera:



Fig. 2.13

2.3.7 Shape from photo-consistency

Quando le immagini di input non sono delle semplici sagome binarie, ma sono foto a colori o a livelli di grigio, è possibile sfruttare l'informazione fotometrica aggiuntiva per migliorare il processo di ricostruzione 3D (shape from photo-consistency).

Si può interpretare l'insieme di immagini in input come un insieme di vincoli posti sulla scena 3D da ricostruire. La scena ricostruita dovrebbe restituire delle immagini molto simili a quelle di partenza. La definizione di somiglianza dipende dal grado di accuratezza che vogliamo raggiungere. Può esistere più di una ricostruzione che soddisfa i vincoli dati. Si dice che una riproduzione è coerente se tutti i punti delle superfici visibili nella scena sono foto-coerenti rispetto a ciascuna immagine. Un punto di una superficie della scena si definisce foto-coerente (photo-consistent) con un insieme di immagini se, per ogni immagine I_k in cui esso è visibile, la sua irradianza (vista dalla telecamera k) è uguale alla intensità del corrispondente pixel nell'immagine. L'uguaglianza può essere definita sulla base della deviazione standard o di una certa norma tra le coppie di colori.

Vi sono principalmente due metodi per determinare la foto-coerenza di un voxel. Nel primo si proietta il centroide del voxel in ciascuna immagine e si applica una soglia sulla varianza dei colori dei pixel così ottenuti. Un secondo metodo, meno

sensibile al rumore, sfrutta i colori di tutti i pixel intersecati dalla proiezione del voxel (e non soltanto il centroide).

Il vincolo di foto-coerenza offre alcuni vantaggi rispetto alla classica ricerca delle corrispondenze nelle immagini, seguita da un passo di triangolazione per ricostruire la struttura 3D:

- il test di foto-coerenza richiede soltanto delle semplici operazioni di proiezione e confronto tra pixel, dove invece la ricerca delle corrispondenze tra immagini è un compito complesso;
- i metodi basati sulle corrispondenze possono lasciare dei “buchi” (se non usano tutti i punti) oppure essere affetti da errori (se usano tutti i punti).

D'altra parte per verificare correttamente la foto-coerenza sono richieste informazioni sulla geometria degli oggetti, sulle proprietà di riflettanza delle superfici e sulla direzione dell'illuminazione. Per completezza rimangono ancora due problemi. In primo luogo, come determinare la ricostruzione più corretta all'interno della classe di tutte le possibili scene coerenti con i vincoli. In secondo luogo, sono necessari dei metodi efficienti per determinare la visibilità di un voxel (condizione necessaria al test di foto-coerenza).

Infatti, quando si applica il test di foto-coerenza ad un voxel, è necessario sapere su quali immagini quel voxel si proietta, altrimenti potrebbe essere considerato erroneamente non foto-coerente, come accadrebbe al voxel blu di figura 2.14 se si ignorasse che esso è occluso dal voxel rosso nella vista di destra. Nel caso specifico, la foto-coerenza del voxel rosso deve essere valutata prima di quello blu. In generale, la foto-coerenza di un voxel dovrebbe venire valutata solo dopo che tutti i voxel che ne potrebbero influenzare la visibilità sono stati valutati (ovvero hanno una etichetta opaco/trasparente).

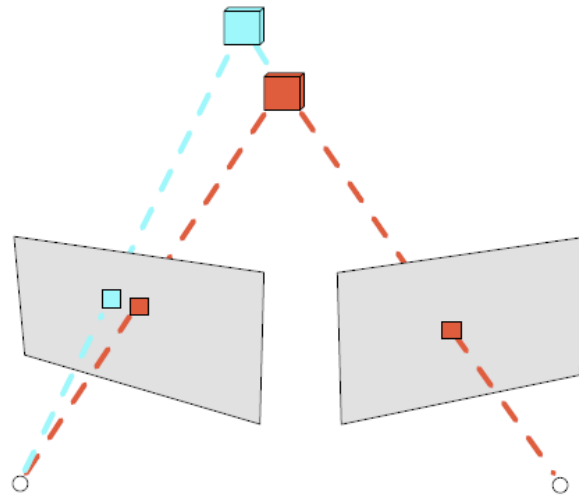


Fig. 2.14

L'implementazione può partire da uno spazio completamente vuoto come in Seitz e Dyer (1999) e marcare come opachi i voxel che superano il test di foto-coerenza (si parla allora di voxel coloring, equivalente a modellare con la creta), oppure può partire da uno spazio pieno [Kutulakos e Seitz (2000)] e rimuovere selettivamente i voxel che non superano il test (space carving, equivalente alla scultura). Si può osservare che con il primo approccio si ottiene in un insieme di voxel che definiscono soltanto le superfici della scena, mentre nel secondo caso si ottiene una ricostruzione del pieno volume degli oggetti.

2.3.8 Structure from motion

Il problema che riguarda la structure from motion può essere formulato come segue: date diverse viste di una scena prese da una telecamera in movimento con parametri intrinseci noti e dato un insieme di punti corrispondenti, dobbiamo ricostruire il moto della telecamera e la struttura della scena.

In fotogrammetria questo è conosciuto come problema dell'orientazione relativa. Si consideri un'unica telecamera in movimento; i parametri intrinseci sono noti, ma il moto della telecamera è incognito (cioè mancano i parametri estrinseci). Il problema di ricavare la struttura della scena con una telecamera in movimento (structure from motion) è stato ampiamente studiato in passato [Huang e Netravali (1994)]. Gli

approcci al problema si possono partizionare in metodi differenziali [Tian e altri(1996)], [Soatto e altri(1996)], [Soatto e Brockett(1998)] o discreti, a seconda che si prendano in ingresso le velocità dei punti nell'immagine (il campo di moto) o un insieme di punti corrispondenti. In [Marr e altri (1998)] viene esaminata la relazione tra i due tipi di approccio. Uno dei metodi più interessanti, che usa il modello prospettico della telecamera, fu proposto da Longuet-Higgins [Longuet-Higgins (1981)]. Questa tecnica si basa sulla matrice essenziale, che descrive la geometria epipolare di due immagini prospettiche (con i parametri intrinseci noti).

La matrice essenziale codifica il moto rigido della telecamera ed, infatti, un teorema [Huang e Faugeras(1989)] permette di fattorizzarla in una matrice di rotazione ed in una di traslazione. Poiché i parametri intrinseci sono noti, ciò è equivalente alla completa conoscenza delle matrici di proiezione prospettica delle telecamere. La struttura (cioè la distanza dei punti dalla telecamera) si trova facilmente con la triangolazione. Da notare che la componente traslazionale dello spostamento può essere calcolata solo a meno di un fattore di scala, perchè è impossibile determinare se il moto misurato nell'immagine è causato da un oggetto vicino che si sposta lentamente o da un oggetto distante che si muove rapidamente. Questo fatto è conosciuto come ambiguità profondità-velocità o depth-speed ambiguity.

2.3.9 Active defocus

Il metodo active defocus è come il metodo visto in precedenza: shape from defocus ma viene impiegata una illuminazione strutturata per creare tessitura ove non fosse presente.

2.3.10 Stereo attivo

Il processo di stereopsi visto precedentemente prende il nome di “passivo”, contrapposto ai metodi “attivi” che includono nel sistema un dispositivo di illuminazione strutturata (structured lighting). Questi sono maggiormente efficaci nel risolvere il problema delle corrispondenze. Vi sono varie tipologie di sistemi di illuminazione strutturata che rientrano nella tecnica di ricostruzione stereo attiva, di seguito sono elencate alcune di queste:

- viene proiettata una tessitura artificiale sulla scena (per esempio punti casuali “sale e pepe”), agevolando il calcolo delle corrispondenze (come in figura 2.15).
- un raggio laser scandisce la scena, proiettando sulle superfici un punto che viene facilmente rilevato e messo in corrispondenza nelle due immagini. E’ necessario prendere molte immagini, poichè per ciascuna coppia di immagini solo ad un punto viene assegnata la disparità.
- una lama di luce (laser) viene fatta passare sulla scena (figura 2.16). Le strisce determinate dalla lama nelle due immagini intersecate con le rispettive rette epipolari forniscono i punti corrispondenti. Per ciascuna coppia di immagini solo ai punti della striscia viene assegnata la disparità. E’ più veloce della soluzione precedente, ma servono sempre molte immagini.

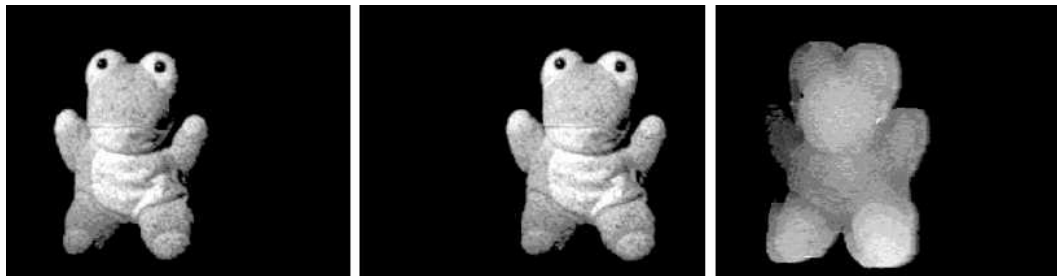


Fig. 2.15

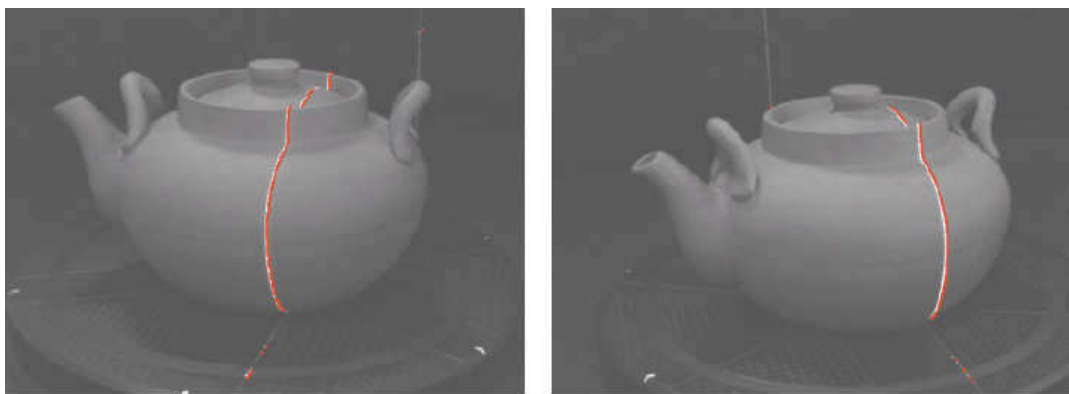


Fig. 2.16

2.3.11 Triangolazione attiva

Il metodo stereo attivo come si è visto nel paragrafo precedente si ha quando sono presenti due telecamere e l'illuminazione serve a rendere più agevole il calcolo delle corrispondenze, che viene però effettuato come per lo stereo passivo, così come pure la successiva triangolazione. Nella triangolazione attiva, invece, si ha una telecamera ed un proiettore (calibrato) ed il calcolo delle corrispondenze e la triangolazione vengono adeguati alla situazione.

Ci sono problemi di zone d'ombra se sorgente e rivelatore non sono allineati (analogamente alle occlusioni nello stereo). Di seguito sono elencate due metodologie che rientrano nella classe di metodo a triangolazione attiva:

- proiettando una lama di luce è possibile rimuovere una telecamera, e procedere alla triangolazione intersecando il piano della lama di luce con il raggio ottico. In questo caso è necessario fare un modello geometrico del proiettore laser e calibrarlo, per poter determinare il piano nello spazio (figura 2.17).
- Invece che proiettare un solo piano, se ne possono proiettare molti simultaneamente, usando un proiettore di bande di luce. In questo caso le bande devono essere codificate in qualche modo: questo è il principio della luce codificata (coded-light). A parità di numero di piani di triangolazione individuati, posso usare meno immagini.

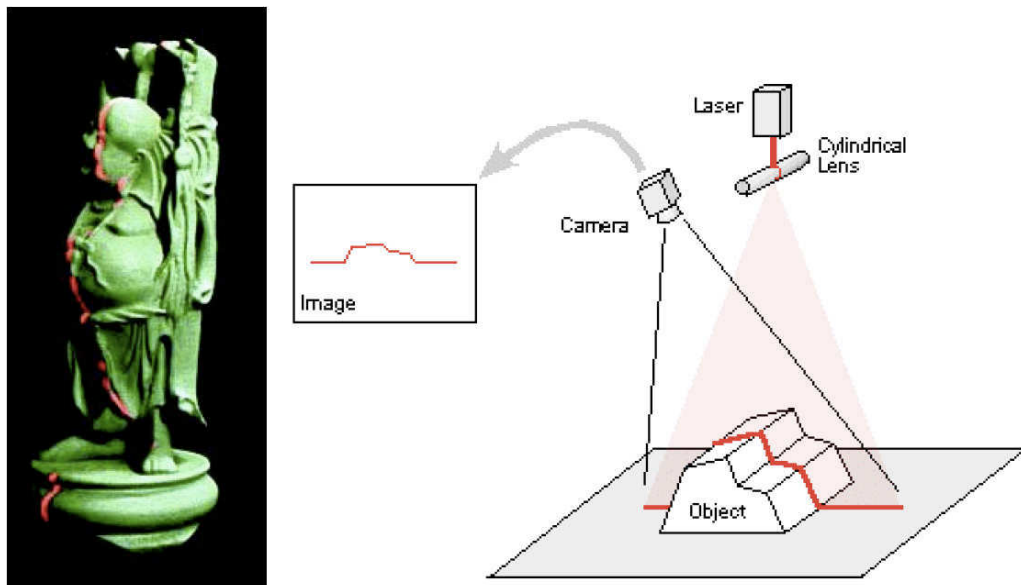


Fig. 2.17

2.3.11.1 Metodo a luce codificata

Nel metodo a luce codificata (coded-light) vengono proiettate, nello stesso istante, numerose strisce (o bande) di luce, che devono essere codificate in qualche modo per distinguerle. La codifica più semplice è realizzata assegnando una diversa luminosità ad ogni direzione di proiezione, per esempio proiettando una scala di intensità lineare. Un'ulteriore sviluppo di questa tecnica sfrutta il colore per la codifica delle direzioni di proiezione [Boyer e Kak(1987)]. Una tecnica molto robusta è la cosiddetta codifica spazio-temporale delle direzioni di proiezione.

Questo metodo si realizza proiettando, all'interno della scena, una sequenza temporale di n-pattern di luce a bande, opportunamente generati da un proiettore a cristalli liquidi (LCD) controllato da un calcolatore.

A ciascuna direzione di proiezione viene così associato un codice di n bit, in cui il bit i-esimo indica se la corrispondente banda era in luce o in ombra nel pattern i-esimo (vedi figura 2.18). Questo metodo permette di distinguere 2^n direzioni di proiezione differenti. Una telecamera, da una posizione diversa da quella del proiettore, acquisisce le n immagini a livelli di grigio dei pattern a strisce proiettati sulla superficie dell'oggetto. Queste vengono quindi convertite in forma binaria in modo da separare le aree illuminate dal proiettore da quelle non illuminate e per ogni pixel

si memorizza il codice di n bit, che, per quanto detto sopra, codifica la direzione di illuminazione della striscia più sottile che ha illuminato il pixel.

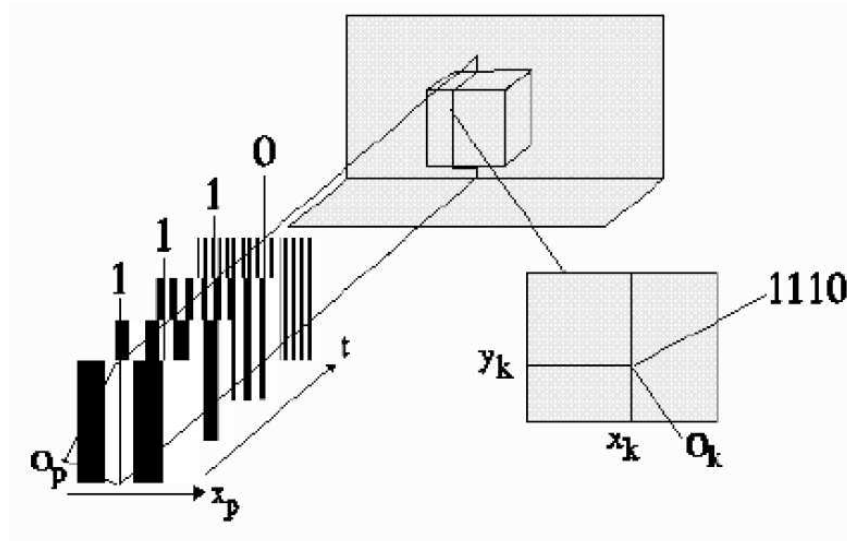


Fig. 2.18

Conoscendo la geometria del sistema, la direzione del raggio ottico della telecamera e l'equazione del piano della banda corrispondente, le coordinate 3D del punto della scena osservato possono essere calcolate intersecando il raggio ottico con il piano.

Per minimizzare l'effetto degli errori si impiega la codifica di Gray per le direzioni dell'illuminazione, in modo che strisce adiacenti differiscano di un solo bit. Il risultato dipende anche dalla corretta binarizzazione delle strisce nelle immagini. Per rendere questo passo indipendente dal cambio delle condizioni di illuminazione dell'ambiente e dalla variazione delle proprietà di riflessione della superficie, le immagini vengono binarizzate con una soglia che varia attraverso l'immagine. Questa soglia si calcola, per ciascun pixel, come la media tra il livello di grigio nell'immagine della scena completamente illuminata e di livello di grigio nell'immagine della scena non illuminata. Il calcolo della immagine di soglia è utile anche per identificare le aree d'ombra presenti nella scena, ovvero quelle zone osservate dalla telecamera ma non illuminate dal proiettore. Esse sono caratterizzate dal fatto che hanno variazioni minime di livello di grigio tra l'immagine illuminata e quella non illuminata. In queste particolari aree si evita di binarizzare le bande e di

conseguenza non si ha una valutazione dei valori di profondità. Nella figura 2.19 sono riportate un esempio di ricostruzione, a sinistra sono riportate le immagini delle frange proiettate mentre a destra è riportata l'immagine range ottenuta, in questo caso le bande sono approssimativamente parallele alle colonne delle immagini.

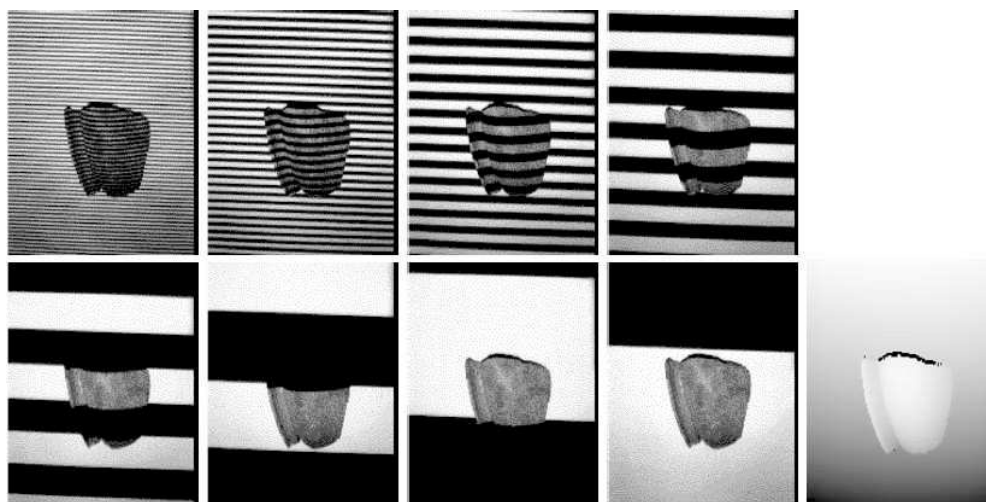


Fig. 2.19

2.3.12 Interferometria

Dei metodi basati sulla interferometria saranno riportati quelli che sfruttano le frange di interferenza di Moirè. L'idea di questo metodo è quella di proiettare una griglia su un oggetto e prenderne un'immagine attraverso una seconda griglia di riferimento. Questa immagine interferisce con la griglia di riferimento (fenomeno di battimento) e crea le cosiddette frange di interferenza di Moirè, che appaiono come bande di luce ed ombra, come viene mostrato in figura 2.20.

L'analisi di questi pattern fornisce informazioni sulla variazione della profondità. In particolare, ogni banda rappresenta un contorno della superficie dell'oggetto di iso-profondità. Un problema è che non è possibile determinare se i contorni adiacenti hanno una profondità più alta o più bassa. La soluzione consiste nel muovere una delle griglie e raccogliere più immagini. La griglia di riferimento si può anche simulare via software. Questi metodi sono in grado di fornire dati molto accurati sulla profondità (risoluzione fino a circa $10\ \mu\text{m}$), ma presentano alcuni inconvenienti. Sono tecniche piuttosto costose dal punto di vista computazionale.

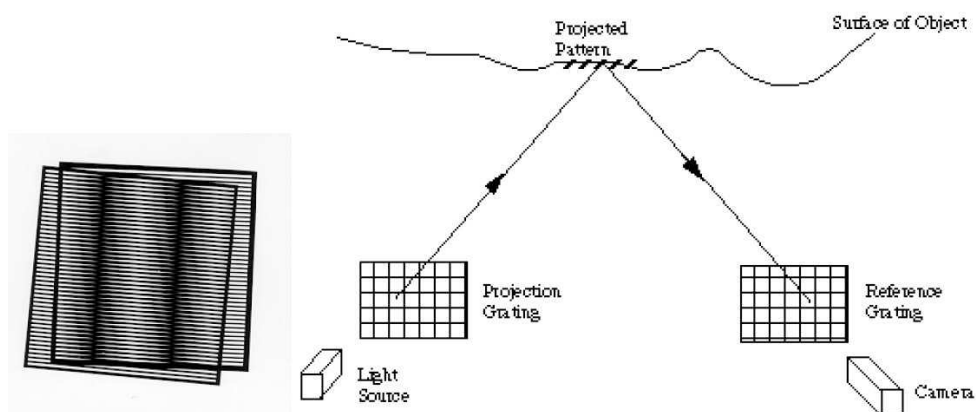


Fig. 2.20

Le superfici con un'ampia angolazione non sono sempre possibili da misurare, poiché la densità delle bande diventa troppo fitta. Nella figura 2.21 è riportato un'applicazione di questa metodologia con i risultati.

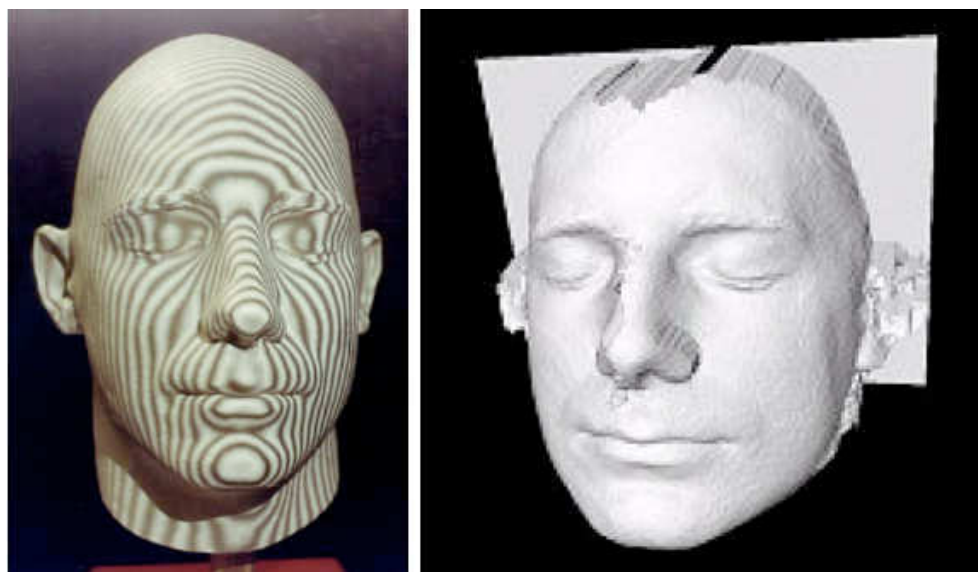


Fig. 2.21

2.3.13 Tempo di volo

Un Range finder a tempo di volo (TOF) sfrutta lo stesso principio del RADAR, ovvero il fatto che la superficie dell'oggetto riflette la luce laser indietro verso un ricevitore che misura quindi il tempo trascorso tra la trasmissione e la ricezione, in modo da calcolare la distanza. Uno scanning laser range finder – o scanner – usa un

fascio di luce laser per scandagliare la scena, in modo da ricavare la distanza di tutti i punti visibili (figura 2.22).

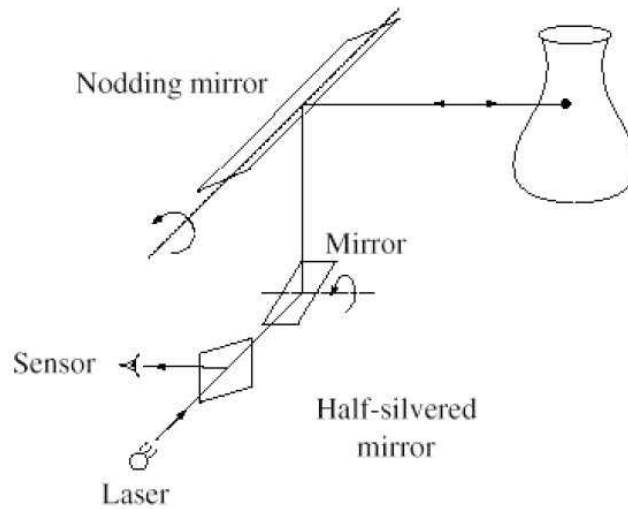


Fig. 2.22

La distanza (o range) r si ottiene dal tempo di propagazione t_r di un'onda luminosa riflessa dal bersaglio (tempo di volo). Se c indica la velocità della luce, allora t_r è dato da: $t_r = 2r/c$. Questo tempo di volo si può misurare con differenti metodi: direttamente (onda impulsiva) o con la conversione in un ritardo di fase mediante modulazione di ampiezza (AM).

CAPITOLO 3

Ricostruzione tridimensionale con tecnica: Shape from Silhouettes

3.1 Principio di ricostruzione volumetrica da sagome

L'obiettivo della ricostruzione volumetrica è creare una rappresentazione che descriva non soltanto la superficie di una regione, ma anche lo spazio che essa racchiude. L'ipotesi fondamentale è che esista un volume noto e limitato all'interno del quale giacciono gli oggetti d'interesse. Tale volume può essere visto come un cubo (box) che circonda la scena e viene rappresentato mediante una griglia discreta di "cubetti" chiamati voxel, equivalenti 3D dei pixel. La ricostruzione coincide con l'assegnazione di un'etichetta di occupazione (o colore) ad ogni elemento di volume. Il valore di occupazione è solitamente binario (trasparente o opaco).

Rispetto alle tradizionali tecniche di stereopsi, la ricostruzione volumetrica offre alcuni vantaggi:

evita il difficile problema della ricerca delle corrispondenze (funziona anche per superfici non tessitate), consente il trattamento esplicito delle occlusioni e consente di ottenere direttamente un modello tridimensionale dell'oggetto (non bisogna allineare porzioni del modello) integrando simultaneamente tutte le viste (che sono dell'ordine della decina). Come nella stereopsi, le fotocamere sono calibrate. Di seguito viene proposto uno dei tanti algoritmi sviluppati per ottenere la ricostruzione volumetrica dalle sagome di un oggetto di interesse. Nel capitolo due sono stati forniti gli elementi basilari riguardanti la definizione di sagoma all'interno della tecnica shape from silhouette.

3.1.1 Algoritmo di Szeliski [15]

L'octree (definito nel capitolo 2) viene costruito ricorsivamente suddividendo ogni cubo in otto cubetti (ottanti) a partire dal nodo radice che rappresenta il volume di lavoro. Ogni cubo ha associato un colore:

- nero: rappresenta un volume occupato;
- bianco: rappresenta un volume vuoto;
- grigio: nodo interno la cui classificazione è ancora incerta.

Per ciascun ottante, si verifica se la sua proiezione sull'immagine i -esima è interamente contenuta nella regione nera. Se ciò accade per tutte le N sagome, l'ottante viene marcato come nero. Se invece la proiezione dell'ottante è interamente contenuta nello sfondo (bianco), anche per una sola telecamera, l'ottante viene marcato come bianco. Se si verifica uno di questi due casi, l'ottante diventa una foglia dell'octree e non viene più processato, altrimenti, esso viene classificato come grigio e suddiviso a sua volta in otto figli. Per limitare la dimensione dell'albero, gli ottanti grigi di dimensione minima vengono marcati come neri. Al termine si ottiene un octree che rappresenta la struttura 3-D dell'oggetto.

Un ottante proiettato sull'immagine forma in generale un esagono irregolare. Eseguire un test accurato di intersezione tra una tale figura e la sagoma può essere complesso. Solitamente si adopera un test approssimato basato sul bounding box o sul bounding circle dell'esagono che, pur essendo approssimativo, classifica con certezza i cubi che stanno totalmente dentro o fuori dalla sagoma. Anche se qualche ottante viene erroneamente classificato come grigio, la decisione sul suo colore definitivo viene semplicemente rimandata alla successiva iterazione. Un esempio di costruzione del modello volumetrico con l'algoritmo di Szeliski è riportato in figura 3.1 [14].

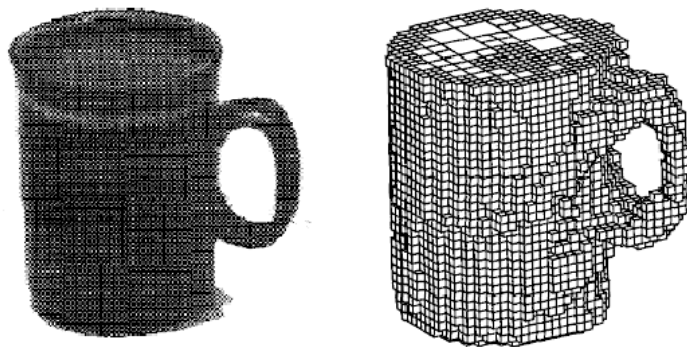


Fig. 3.1

3.1.2 Metodo proposto

Il metodo adottato consiste nel definire un box di voxel contenente l'oggetto nello spazio tridimensionale e scartare successivamente tutti quei punti del volume iniziale che hanno intersezione vuota con almeno uno dei coni delle sagome ottenute dalle immagini acquisite, si è scelto di assegnare ai voxel soltanto valori binari, cioè bianco o nero e non introdurre come viene fatto nell'algoritmo precedente un ulteriore valore assumibile individuato con il colore grigio. Questa scelta è stata dettata dal fatto che con la semplice metodologia implementata i risultati si sono mostrati accettabili.

3.2 Elaborazione delle immagini

L'elaborazione delle immagini è un processo che può essere suddiviso in più fasi ognuna delle quali è rilevante o meno a seconda dell'utilizzo cui le immagini sono dedicate. Nella nostra applicazione quello che si vuole fare è ottenere la silhouette a partire da una immagine in RGB (figura 3.2), a questo proposito è stata effettuata la segmentazione dell'immagine.



Fig. 3.2

La segmentazione consiste nel dividere un'immagine nelle regioni (o oggetti) che la compongono. La segmentazione ha come ingresso l'immagine e come uscita degli attributi (le regioni), essa è basata su forti discontinuità dell'intensità o su regioni che presentano intensità omogenea sulla base di un criterio stabilito. Esistono quattro tipi

principali di discontinuità: punti, linee, contorni e spigoli (corner) o, in generale, punti di interesse.

Per separare un oggetto dallo sfondo si utilizza il metodo della sogliatura con soglia s : ogni punto (u,v) con $f(u,v) > s$ ($f(u,v) < s$) viene identificato come oggetto altrimenti come sfondo. L'elaborazione proposta serve ad individuare e caratterizzare le varie regioni che sono presenti all'interno di ciascuna immagine; in particolare, si cerca di separare gli oggetti dallo sfondo.

Il primo passo è stato quello di rendere le immagini a colori delle immagini in scala di grigi (figura 3.3) in modo da poter avere come valore di intensità di soglia uno scalare.



Fig. 3.3

L'immagine che si ottiene è una matrice delle dimensioni pari al numero di pixel lungo le due direzioni del sensore con gli elementi aventi valori di intensità da 0 a 255. Successivamente è selezionato un insieme limitato di pixel che contiene la proiezione dell'oggetto nel piano immagine considerato in modo da facilitare l'operazione di segmentazione (figura 3.4).

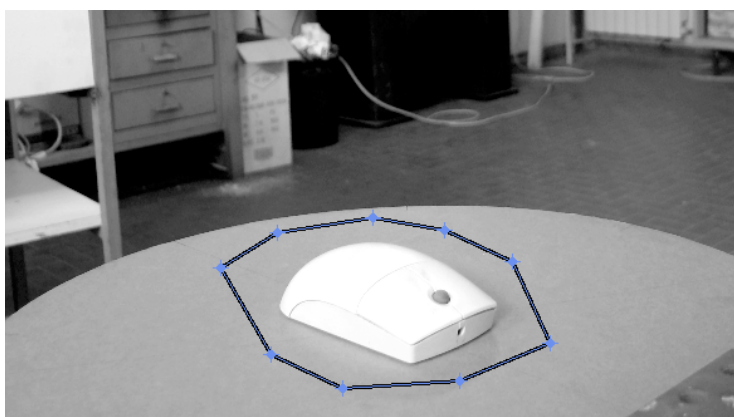


Fig. 3.4

Si individuano quei pixel che hanno intensità luminosa $f(u,v)$ maggiore di quella dello sfondo T , la quale è attribuibile a quei punti dell'insieme considerato non rappresentativi dell'oggetto. Questa tecnica è molto efficace, se come detto in precedenza, si riesce a distinguere bene lo sfondo dall'oggetto.

Ad ogni pixel (u,v) è stato assegnato rispettivamente valore unitario o nullo se rappresenta un raggio ottico passante per l'oggetto oppure no (figura 3.5).

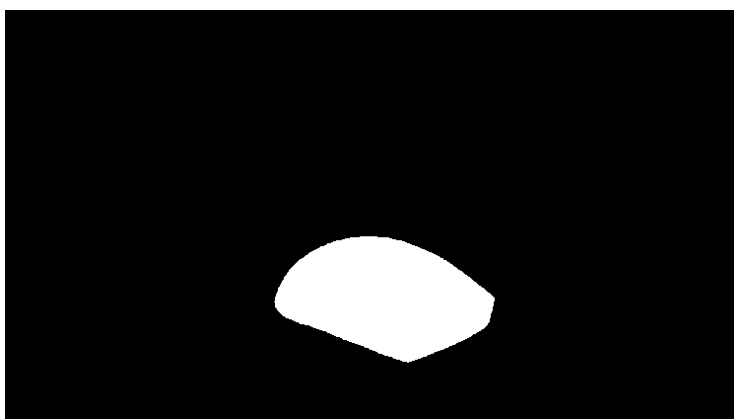


Fig. 3.5

3.3 Calibrazione della telecamera [1]

Affinchè la ricostruzione della forma tridimensionale sia precisa c'è bisogno di individuare con estrema esattezza i parametri intrinseci ed estrinseci della telecamera.

Nel primo capitolo è stato descritto il processo di formazione delle immagini, ovvero quella che si è definita visione di basso livello. In sostanza lo studio di tale processo

è utile alla comprensione del funzionamento dei sistemi di visione, ma non consente ancora di estrapolare delle metodologie per utilizzare attivamente una telecamera in applicazioni che vadano al di là di una semplice rappresentazione.

Sempre nel primo capitolo è stato dato un formalismo matematico alle descrizioni fatte, attraverso un modello della telecamera, concepito come un'applicazione lineare dallo spazio fisico tridimensionale allo spazio delle immagini bidimensionale. Pur costituendo ancora un'analisi della visione a basso livello, la modellistica offre degli strumenti e per poter condurre delle considerazioni quantitative, e per poter manipolare i dati con altri scopi. In altre parole la conoscenza del modello della telecamera introduce alla visione ad alto livello.

Tuttavia permane il problema della stima dei parametri della telecamera. Del modello se ne conosce la struttura mediante la (3.1) :

$$\{\tilde{\phi}\} = [M] \cdot \{\tilde{w}\} \quad (3.1)$$

Dove gli elementi della matrice $[M]$ rappresentano i parametri intrinseci ed estrinseci della telecamera.

Pur riferendosi alle specifiche tecniche fornite dal costruttore, risulta impossibile conoscere con un livello di precisione sufficiente i parametri intrinseci della telecamera, in quanto la distanza focale non è fissa ed noto solo l'intervallo entro il quale varia. Pertanto si dispone di immagini acquisite con un certo fuoco, ma non se ne conosce la posizione esatta.

Per quanto riguarda i parametri estrinseci la situazione è ancora più complicata: la misura di distanze ed angoli viene fatta in applicazioni edilizie, su superfici dai 100 m² in sù, a mezzo di tecniche di triangolazione, dotate di un livello di incertezza piuttosto alto. Volendo utilizzare le stesse tecniche in un laboratorio di visione, la cui superficie non supera i 30 m², si avrebbe un errore relativo molto maggiore. Se poi si aggiunge che la precisione richiesta ad un sistema di visione di alto livello deve essere comparabile con quella di strumenti di misura piuttosto raffinati, tali metodi risultano senza ombra di dubbio inaccettabili.

E' necessario, dunque, disporre di un metodo indiretto di identificazione dei suddetti parametri. La ricerca di tale metodo rappresenta il problema di calibrazione della telecamera.

3.3.1 Formalizzazione del problema di calibrazione

Si consideri il modello della telecamera così come è stato espresso nella (3.1). La matrice $[M]$ può essere scritta in modo da evidenziare la sua struttura :

$$[M] = \begin{bmatrix} \{m_1\}^T & m_{14} \\ \{m_2\}^T & m_{24} \\ \{m_3\}^T & m_{34} \\ \{m_4\}^T & m_{44} \end{bmatrix} \quad (3.2)$$

Che nel caso di $\hat{n} \equiv \zeta$ diventa :

$$[M] = \begin{bmatrix} \{m_1\}^T & m_{14} \\ \{m_2\}^T & m_{24} \\ \{0\}^T & 0 \\ \{\zeta\}_o^T & t_\zeta \end{bmatrix} \quad (3.2)$$

Considerato un punto di vettore posizione $\{w_p\}$, applicando il modello si ha che:

$$\begin{cases} u_p = \frac{\{m_1\}^T \{w_p\} + m_{14}}{\{\zeta\}^T \{w_p\} + t_\zeta} \\ v_p = \frac{\{m_2\}^T \{w_p\} + m_{24}}{\{\zeta\}^T \{w_p\} + t_\zeta} \end{cases} \quad (3.3)$$

Se nelle (3.3) si suppongono note le coordinate del punto nel piano immagine e nel riferimento $\{O, x, y, z\}$, le uniche incognite risultano i 12 parametri della telecamera, ovvero gli elementi della matrice $[M]$. Riunendo tali incognite nel vettore :

$$\{\pi_c\} = \begin{Bmatrix} \{m_1\} \\ m_{14} \\ \{m_2\} \\ m_{24} \\ \{\zeta\}_o \\ t_\zeta \end{Bmatrix} \quad (3.4)$$

Il sistema (3.3) si può scrivere come :

$$[\Omega_p] \cdot \{\pi_c\} = \{0\} \quad (3.5)$$

Dove la matrice $[\Omega_p]$ ha espressione :

$$[\Omega_p] = \begin{bmatrix} w_{Px} & w_{Py} & w_{Pz} & 1 & 0 & 0 & 0 & 0 & -u_p w_{Px} & -u_p w_{Py} & -u_p w_{Pz} & -u_p \\ 0 & 0 & 0 & 0 & w_{Px} & w_{Py} & w_{Pz} & 1 & -v_p w_{Px} & -v_p w_{Py} & -v_p w_{Pz} & -v_p \end{bmatrix} \quad (3.6)$$

Il sistema (3.5), come è evidente, è omogeneo, ed ha 2 equazioni in 12 incognite, quindi non è risolubile. Per rendere il sistema risolubile bisognerà aumentare il numero di equazioni. Poiché per ogni punto $\{w_i\}$ considerato si hanno due equazioni (una per u e una per v), scegliendo 6 punti si ha un sistema omogeneo di 12 equazioni in 12 incognite :

$$[\Omega] \cdot \{\pi_c\} = \{0\} \quad (3.7)$$

Con:

$$[\Omega] = \begin{bmatrix} [\Omega_1] \\ [\Omega_2] \\ [\Omega_3] \\ [\Omega_4] \\ [\Omega_5] \\ [\Omega_6] \end{bmatrix}$$

Le soluzioni della (3.7) dipendono dal determinante della matrice $[\Omega]$, ovvero da come sono stati scelti i punti nello spazio di lavoro.

Poiché non è facile conoscere con esattezza le coordinate di 6 punti indipendenti dello spazio geometrico, si corre il rischio di avere soluzione banale, ovvero parametri tutti nulli. Per aggirare quest'ostacolo si può evitare un approccio di algebra tradizionale e ricorrere a metodi di calcolo numerico. Scelti N punti non complanari dello spazio geometrico si può impostare un problema di ottimizzazione lineare vincolata del tipo:

$$\begin{aligned} \text{trovare } \{\pi_c\}: \min_{\{\pi_c\}} \|[\Omega] \cdot \{\pi_c\}\| \\ \text{con il vincolo di } \|\{\zeta\}\| = 1 \end{aligned} \quad (3.8)$$

Il problema così esposto può essere materialmente risolto utilizzando il *metodo dei moltiplicatori di Lagrange*. Avendo fatto ricorso al calcolo numerico, è intuitivo comprendere che quanto più è grande il numero N di punti noti, più accurato sarà il risultato.

Volendo invece evitare di risolvere un sistema di equazioni omogeneo (che comporta il rischio di soluzione banale) è possibile sfruttare una proprietà delle coordinate omogenee :

$$k \cdot [M] \{\tilde{w}\}_o = \begin{Bmatrix} k \cdot u \\ k \cdot v \\ 0 \\ k \end{Bmatrix} = \begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} \quad (3.9)$$

Dove k è una qualsiasi costante non nulla. Ciò che appare evidente nella (3.9) è che le coordinate omogenee sono sempre definite a meno di una costante di proporzionalità, e questo consente di avere un grado di libertà in più in un problema di stima parametrica. In altre parole, al posto di utilizzare come incognita il vettore $\{\pi_c\}$, si può lavorare con una sua normalizzazione rispetto ad una costante diversa da zero. Se si fa la scelta di $k = t_\zeta$, i nuovi parametri normalizzati saranno:

$$\{\hat{\pi}_c\} = \begin{Bmatrix} \{\hat{m}_1\} \\ \hat{m}_{14} \\ \{\hat{m}_2\} \\ \hat{m}_{24} \\ \{\hat{\xi}\} \end{Bmatrix} \quad (3.10)$$

Dove con l'apice angolato si sono intese le rispettive quantità normalizzate.

Considerati al solito N punti non complanari, di coordinate note sia nello spazio geometrico che nello spazio delle immagini, si può scrivere il sistema :

$$\begin{bmatrix} w_{1x} & w_{1y} & w_{1z} & 1 & 0 & 0 & 0 & 0 & -u_1 w_{1x} & -u_1 w_{1y} & -u_1 w_{1z} \\ 0 & 0 & 0 & 0 & w_{1x} & w_{1y} & w_{1z} & 1 & -v_1 w_{1x} & -v_1 w_{1y} & -v_1 w_{1z} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ w_{Nx} & w_{Ny} & w_{Nz} & 1 & 0 & 0 & 0 & 0 & -u_N w_{Nx} & -u_N w_{Ny} & -u_N w_{Nz} \\ 0 & 0 & 0 & 0 & w_{Nx} & w_{Ny} & w_{Nz} & 1 & -v_N w_{Nx} & -v_N w_{Ny} & -v_N w_{Nz} \end{bmatrix} \cdot \{\hat{\pi}_c\} = \begin{Bmatrix} u_1 \\ v_1 \\ \dots \\ u_N \\ v_N \end{Bmatrix} \quad (3.11)$$

Ovvero $[\hat{\Omega}]\{\hat{\pi}_c\} = \{\mu\}$.

Il sistema (3.11) è di N equazioni in 11 incognite. Essendo il sistema ancora lineare si può ricorrere ad una risoluzione che utilizzi la pseudoinversa della matrice $[\hat{\Omega}]$ e, in condizioni non singolari, riuscire ad ottenere :

$$\{\hat{\pi}_c\} = \left([\hat{\Omega}]^T [\hat{\Omega}] \right)^{-1} \cdot [\hat{\Omega}]^T \cdot \{\mu\} \quad (3.12)$$

Tale metodo, come già si è accennato, richiede che il determinante della matrice $[\hat{\Omega}]^T [\hat{\Omega}]$ sia diverso da zero e quindi, anche in questo caso, la scelta dei punti da utilizzare per la calibrazione può condizionarne l'efficienza.

I metodi risolutivi illustrati nella (3.8) e nella (3.12) si fondano tutti su strutture lineari. L'aver voluto preservare tale linearità con opportuni formalismi matriciali ha però l'handicap di fare una stima affetta da un errore algebrico. Questo significa che eventuali effetti di compensazione fra scostamenti di segno opposto possono dar luogo ad una calibrazione di cui non si può valutare l'affidabilità. Per questo si può

ricorrere ad un metodo non lineare, computazionalmente più oneroso, ma che valuti la media quadratica dell'errore.

Il problema può essere allora posto nei seguenti termini : trovare gli elementi della matrice $[M]$, espressa nella forma della (5.1), tali da minimizzare la funzione :

$$E = \sum_{i=1}^N \left(\left\| \frac{\{m_1\}^T \{w_i\} + m_{14}}{\{\zeta\}^T \{w_i\} + t_\zeta} - u_i \right\|^2 + \left\| \frac{\{m_2\}^T \{w_i\} + m_{24}}{\{\zeta\}^T \{w_i\} + t_\zeta} - v_i \right\|^2 \right) \quad (3.13)$$

Con il vincolo $\|\{\zeta\}\| = 1$

Il metodo opera attraverso i seguenti passi :

- 1) Si fa una stima approssimativa dei parametri, magari utilizzando un metodo lineare.
- 2) I valori stimati vengono utilizzati come parametri di inizio ottimizzazione
- 3) Si calcolano le coordinate immagine di tutti i punti.
- 4) Si valuta la media quadratica sugli N punti degli scostamenti tra le coordinate in pixel calcolate e quelle note .
- 5) Si itera il procedimento fino a quando la funzione E si stabilizza sul minimo valore ottenuto.

Tale metodo è detto *dei minimi quadrati* proprio perché attua una minimizzazione della media quadratica degli scostamenti.

3.3.2 Un algoritmo di calibrazione [17]

Lineari o meno che siano, i metodi di risoluzione del problema di calibrazione della telecamera dipendono dal numero N di punti di cui si conoscono le coordinate, e nello spazio geometrico, e nello spazio delle immagini. Naturalmente più questo numero è basso, più si degradano le prestazioni del metodo. Si pone dunque il problema di trovare un numero elevato di punti, di cui si possano conoscere, senza incappare nelle solite imprecisioni della misura diretta, le coordinate. Un metodo efficiente per fare ciò consiste nel realizzare una struttura di riferimento, sulla quale sia semplice individuare una famiglia di punti.

Si immagini di realizzare una griglia piana, a scacchi bianchi e neri, come quella rappresentata in figura 3.6:

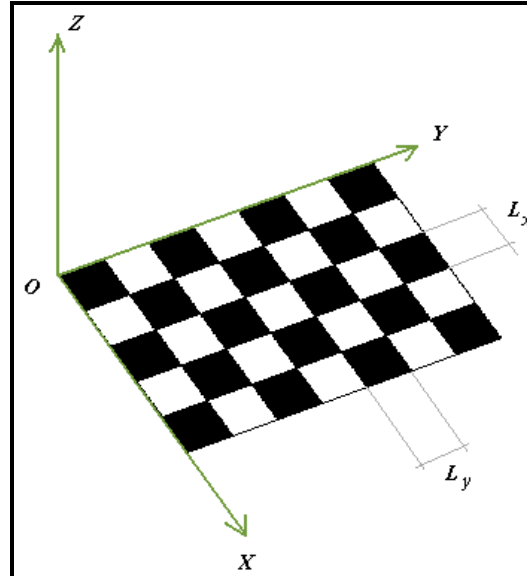


Fig. 3.6

Scelto un opportuno sistema di riferimento è possibile individuare con semplicità le posizioni di tutti i punti della griglia. Difatti, note le seguenti quantità :

- L_x : dimensione del singolo rettangolino lungo x
- L_y : dimensione del singolo rettangolino lungo y
- N_x : numero di rettangoli lungo la direzione x
- N_y : numero di rettangoli lungo la direzione y

E' possibile individuare la posizione di ogni vertice mediante una semplice convenzione. Se si pensa di ordinare i vertici come elementi di una matrice rettangolare, il punto P_{ij} avrà coordinate :

$$P_{ij} \equiv (i \cdot L_x \quad j \cdot L_y \quad 0)$$

Con $i, j \in \{1, \dots, N\}$ (3.14)

Una volta definiti tutti i punti della griglia nello spazio geometrico, occorre individuarli nello spazio delle immagini.

Nella realtà i vertici della griglia sono legati da relazioni di omografia planare, ma quando si passa nel piano immagine tali relazioni si alterano: ciò è dovuto al fatto che la matrice $[M]$ non è ortogonale, ovvero la trasformazione che essa attua non conserva né distanze né angoli. Pertanto risulta estremamente difficile estrarre le coordinate in pixel dei vertici. Per implementare comunque una tecnica è necessario catalogare le informazioni che si conservano nel passaggio dalla realtà all'immagine. Nel caso in cui si conoscano solo i quattro angoli esterni della griglia tali informazioni sono :

- 1) Nell'approssimazione che gli effetti di distorsione radiale e tangenziale siano trascurabili, i quattro punti estremi sono connessi da linee rette.
- 2) Anche se l'immagine catturata non è a colori, ma in scala di grigio, si suppone che le condizioni di illuminazione non alterino sensibilmente le differenze di tono dei vari rettangoli della griglia-scacchiera. In altre parole deve essere ancora nettamente rilevabile il motivo a scacchi chiari e scuri della griglia.
- 3) Tutti i vertici che si trovano su un lato esterno della scacchiera appartengono alla congiungente fra due angoli estremi.

Naturalmente l'algoritmo ha un suo grado di approssimazione, e per questo, ci sarà sempre un errore residuo già a valle dell'estrazione dei vertici.

Si cominciano dunque ad individuare i quattro angoli estremi della griglia. E' possibile farlo mediante un software di elaborazione delle immagini, in grado di fornire le coordinate in pixel di un punto quando esso viene "cliccato" dall'utente. Alla fine di questa operazione si possono congiungere i quattro angoli estremi e si ottiene così il perimetro esterno della scacchiera :

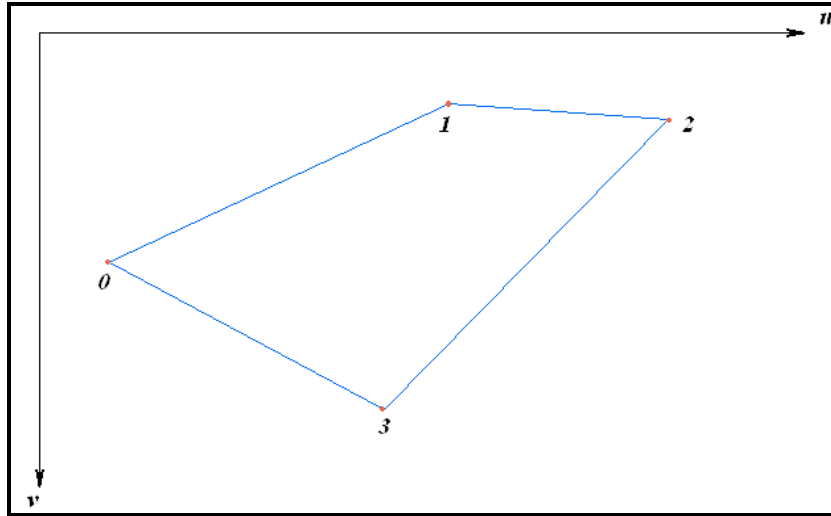


Fig. 3.7

I punti in figura 3.7 sono stati ordinati in senso orario. A questo punto si scrive l'equazione parametrica di ogni lato nel piano immagine. Per il lato $0\bar{1}$ si ha :

$$\begin{cases} u(s) = \frac{(u_1 - u_o)}{\ell_{01}} \cdot s + u_o \\ v(s) = \frac{(v_1 - v_o)}{\ell_{01}} \cdot s + v_o \end{cases}$$

$$s \in \{0, \varepsilon, 2\varepsilon, \dots, (n_c - 1)\varepsilon, \ell_{01}\}$$

$$\ell_{01} = \|\{m_1\} - \{m_o\}\|$$

$$\varepsilon \leq \min(\delta_u, \delta_v)$$

$$n_c = \frac{\ell_{01}}{\varepsilon}$$
(3.15)

Dove con un abuso di notazione si sono indicate con u_o e v_o le coordinate in pixel del punto 0 , s è l'ascissa curvilinea discretizzata, ed ε il campionamento spaziale.

Poiché lo scopo è quello di rintracciare tutti i vertici della griglia, si farà la scansione del primo lato per individuare i punti in cui la differenza di colore raggiunge un valore sufficientemente elevato. Per tradurre questa affermazione in termini quantitativi si consideri la funzione immagine :

$$\mathfrak{I} : (u, v) \rightarrow \tau$$
(3.16)

Dove τ è il valore associato al colore del pixel di coordinate (u, v) . Tale funzione rappresenta proprio il processo di rappresentazione delle immagini sullo schermo del calcolatore elettronico. Considerate le coppie di punti (u_c, v_c) e (u_s, v_s) che rappresentano rispettivamente un punto all'interno di un qualsiasi rettangolo chiaro e un punto all'interno di un qualsiasi rettangolo scuro per cui il valore di colore è pari a quello mediato sull'intero rettangolo, si definisce la soglia Δ :

$$\Delta = \min_{\text{ret tan goli}} \{|\mathfrak{S}(u_c, v_c) - \mathfrak{S}(u_s, v_s)|\} \quad (3.17)$$

Che rappresenta la minima differenza di colore che si ha passando da un rettangolo chiaro ad un rettangolo scuro e viceversa.

Indicato il generico valore di ascissa curvilinea discretizzata con $s_i = i \cdot \varepsilon$, e facendolo variare su tutto il primo lato, ogni qual volta si verifica la condizione :

$$|\mathfrak{S}(u(s_{k+1}), v(s_{k+1})) - \mathfrak{S}(u(s_k), v(s_k))| \geq \Delta \quad (3.18)$$

il punto $(u(s_k), v(s_k))$ è un vertice della griglia appartenete al primo lato.

Iterando la procedura su tutti e quattro i lati si otterrà una mappatura del perimetro della scacchiera, illustrata in figura 3.8.

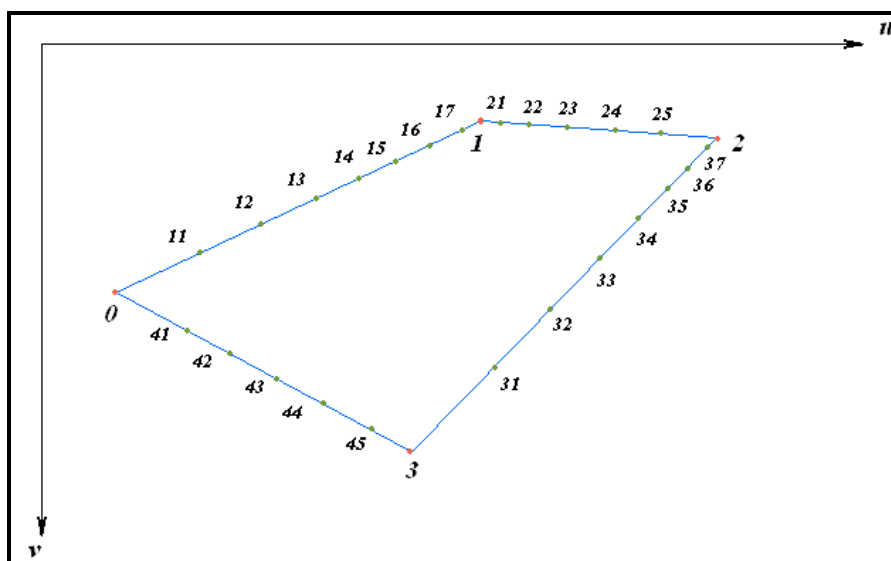


Fig. 3.8

A questo punto, previa istituzione di una convenzione per l'ordinamento dei punti ricavati, non resta che dedurre le posizioni di tutti i vertici interni al rettangolo. Se, con riferimento alla figura 3.8, si sceglie di congiungere i punti aventi il secondo indice uguale, si ha :

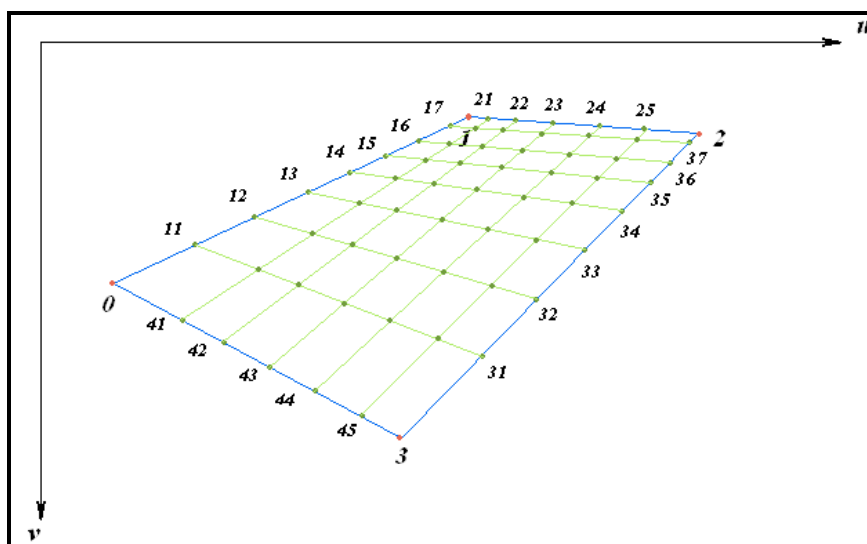


Fig. 3.9

Ricavate le coordinate dei vertici nello spazio geometrico mediante la (3.14) e quelle nel piano immagine mediante la procedura appena descritta, si dispone di un numero $N = N_x \cdot N_y$ di punti sui quali effettuare la calibrazione della telecamera.

Di norma i processi di ottimizzazione vincolata danno vita a risultati corrotti da un errore ancora troppo alto, soprattutto per quanto riguarda i parametri intrinseci. Questo è dovuto al fatto che l'estrazione dei vertici è fortemente condizionata dall'errore dell'utente, che pur agendo sui soli angoli estremi, finisce per inficiare l'intera calibrazione. Essendo gli errori umani incertezze accidentali, essi sono dei fenomeni aleatori, e in particolare hanno una distribuzione simmetrica intorno ad un valor medio. Per ridurre gli effetti complessivi dovuti a tali incertezze si può pensare di operare una calibrazione con più immagini, nelle quali, di volta in volta, la scacchiera assume una posizione ed un orientamento differenti. In questo modo si aumenta la popolazione primitiva del fenomeno stocastico, e si ha un effetto benefico anche sull'identificazione dei parametri intrinseci. Questo avviene per il fatto che, mentre i parametri estrinseci variano per ogni immagine (la posizione della scacchiera relativa alla telecamera cambia), i parametri intrinseci sono sempre gli stessi, e quindi in relazione ad essi si incrementa il numero di equazioni a parità di incognite.

Per avere un'idea sul numero di immagini da utilizzare per la calibrazione bisogna fare un bilancio fra il numero di equazioni ed il numero di incognite. Si definiscono le seguenti quantità :

- n : numero di immagini
- N_e : numero di equazioni
- N_i : numero di incognite
- $r = N_e / N_i$: grado di ridondanza

In base alle considerazioni fatte valgono le relazioni :

$$\begin{aligned}
N_e &= 2nN = 2nN_x N_y \\
N_i &= 6n + 4 \\
2nN &= 6nr + 4r \\
n &= \frac{2r}{N - 3r} \\
r &= \frac{2nN}{6n + 4}
\end{aligned}
\tag{3.19}$$

Dalle (5.20) si può ricavare o il numero delle immagini da utilizzare fissato un grado di ridondanza o viceversa.

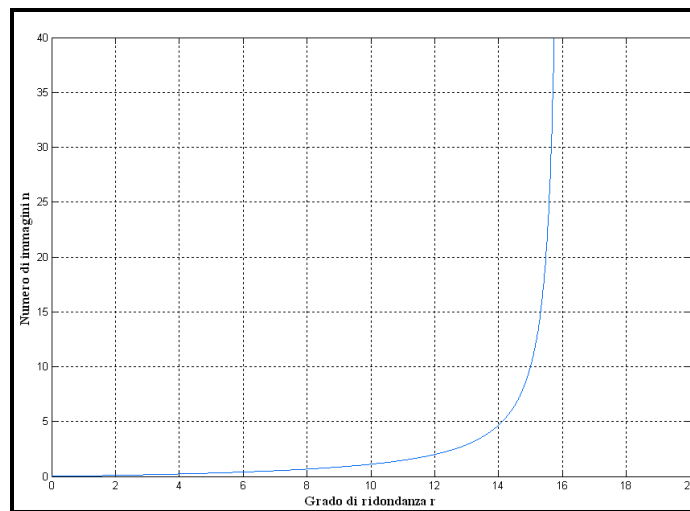


Fig. 3.10

Dal grafico di figura 3.10 si vede che, per una scacchiera a 48 rettangoli, aumentare il numero di immagini oltre 20 comporta un incremento del grado di ridondanza veramente poco percettibile. Inoltre il grado di ridondanza ottenibile è sempre limitato dall'asintoto di equazione $r = N/3$: le dimensioni della scacchiera limitano superiormente i gradi di libertà del metodo di ottimizzazione.

3.3.3 Software utilizzato per la calibrazione [18]

Di seguito viene proposta la procedura di calibrazione utilizzando il Camera Calibration Toolbox for Matlab[®].

L'intero processo è suddiviso in 3 fasi:

- caricamento delle immagini di calibrazione;
- estrazione degli angoli;
- calcolo dei parametri.

Le immagini di calibrazione utilizzate sono del tipo indicato in figura 3.11.

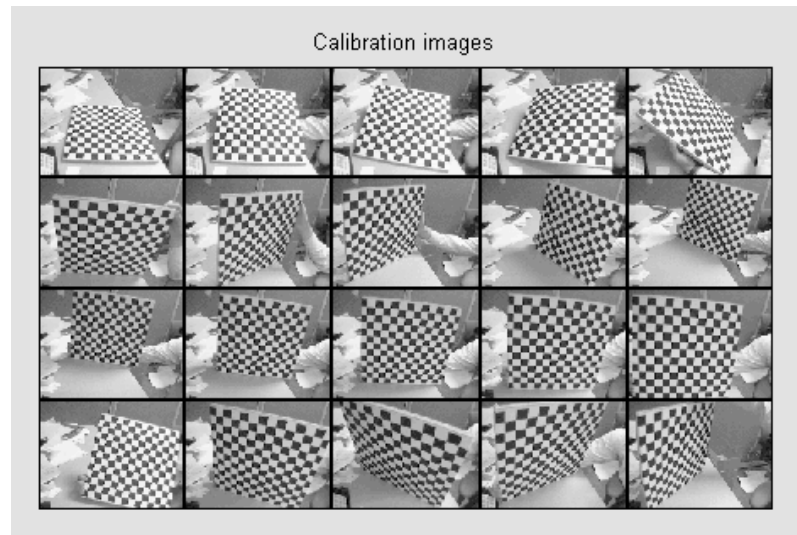


Fig. 3.11

Caricate le immagini si prosegue poi all'estrazione degli angoli della griglia. Si scelgono, per ogni immagine, gli angoli esterni della griglia, come indicato in figura 3.12.

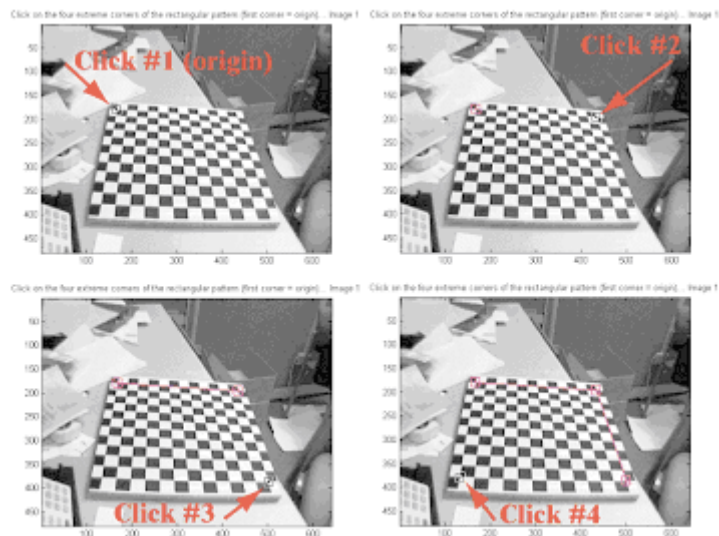


Fig. 3.12

Fissati i 4 angoli esterni, si procede con l'elaborazione per individuare tutti gli angoli della griglia; il risultato è riportato in figura 3.13.

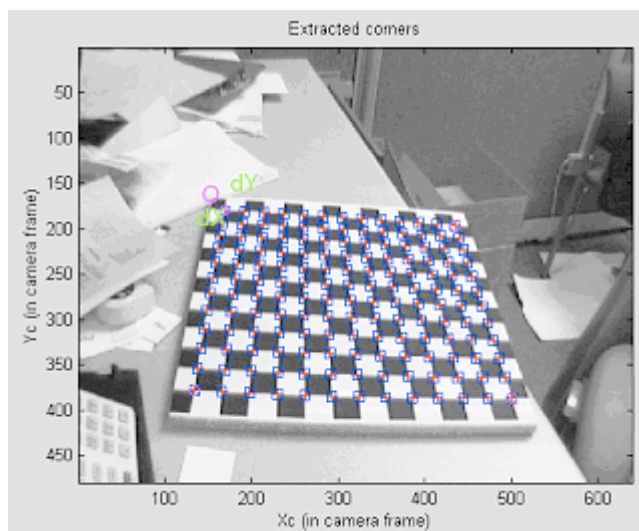


Fig. 3.13

Terminata la fase di estrazione degli angoli, si procede al calcolo dei parametri, che avviene in due passi: inizializzazione e poi ottimizzazione non lineare.

L'inizializzazione calcola una soluzione in forma chiusa che non include alcuna distorsione delle lenti.

L'ottimizzazione non lineare minimizza l'errore di riproiezione globale (in termini di minimi quadrati) su tutti i parametri.

Il processo può essere ripetuto, se i risultati non sono corretti; la verifica viene effettuata procedendo alla riproiezione degli angoli sulle immagini.

Infine il programma restituisce i parametri intrinseci nella versione più generale possibile:

- ↪ Il vettore distanza focale espresso in pixel $\{f\} = \{f_u \quad f_v\}^T$. Nel caso in cui il pixel ha forma quadrata si ha $f_u = f_v = f$
- ↪ L'angolo α di *skew*, definito come il complemento a 90° dell'angolo fra l'asse u e l'asse v .
- ↪ Le coordinate dell'origine della terna telecamera nel piano immagine, che nel capitolo 1 sono state denotate con (u_o, v_o) .

↯ I coefficienti di distorsione fino al sesto ordine, ovvero il vettore

$$\{k_c\} = \{k_{c1} \quad k_{c2} \quad k_{c3} \quad k_{c4} \quad k_{c5} \quad k_{c6}\}^T.$$

I parametri estrinseci, invece, si ottengono dalla fase di posizionamento della telecamera nello spazio ottenendo la matrice di trasformazione omogenea che permette di passare dalla terna telecamera a quella globale.

3.4 Descrizione dell'algorithmo proposto

In seguito alle elaborazioni delle immagini effettuate, si hanno, per ogni immagine, le coordinate in pixel dei punti appartenenti alle sagome (Fig. 3.23) e dei baricentri di ogni sagoma. Grazie alla procedura di calibrazione è possibile associare ad ogni sagoma un matrice di trasformazione omogenea che consente di esprimere le coordinate dei punti nel piano immagine in una terna globale.

Si procede alla stima di un volume iniziale contenente l'oggetto discretizzato in voxel (figura 3.14), oltre alle dimensione del volume di partenza è necessario conoscere anche la posizione del box rispetto alla terna globale, per fare ciò si prendono in considerazione due sagome e si calcola il punto di intersezione dei raggi ottici passanti per i baricentri delle due. Le coordinate del punto di intersezione sono espresse nel riferimento globale, passando allora alla definizione dell'insieme di voxel vengono individuate in tale riferimento le coordinate dei baricentri di tutti i voxel appartenenti al volume iniziale.

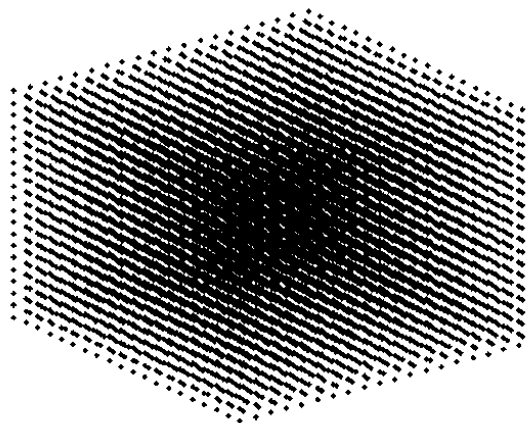


Fig. 3.14

Se si prendono in esame i centri dei voxel per ogni generica sagoma i , si può ottenere la proiezione di tali punti nel piano immagine contenente la sagoma i -esima come segue:

$$\{\tilde{\varphi}\}_j = [M]_i \cdot \{\tilde{w}\}_j \quad i = 1, \dots, p \quad j = 1, \dots, q \quad (3.20)$$

Dove si è indicato con p il numero di sagome ed q il numero dei centri dei voxel, $[M]_i$ è la matrice che permette di passare dalla terna globale al piano immagine contenente la sagoma.

L'algoritmo di ricostruzione parte con la proiezione di tutti i centri nel piano dell'immagine p -esima, si eliminano tutti i punti che risultano esterni all'area dell'immagine. Si può scrivere la seguente relazione:

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix}_j = [\tilde{\varphi}]_j = [M]_p \cdot \{\tilde{w}\}_j \quad j = 1, \dots, q \quad 1 \leq u \leq n \quad 1 \leq v \leq m \quad (3.21)$$

Dove n e m sono rispettivamente il numero di pixel lungo u e lungo v . Successivamente si crea una matrice sparsa di dimensione $n \times m$ avente elementi con valore unitario per ogni coppia di coordinate (u, v) ottenute dalla (3.21) per $p = 1$, cioè prendendo in considerazione la prima immagine:

$$[A] := a_{hk} = 1 \quad h = u_j, k = v_j \quad (3.22)$$

Si definisce l'insieme dei pixel di coordinate (\bar{u}, \bar{v}) appartenenti alla prima sagoma come segue:

$$\Omega_1 := (\bar{u}, \bar{v})_1 \quad (3.23)$$

Come detto in precedenza i punti dell'immagine contenente la sagoma sono così definiti:

$$I_1(u,v) = \begin{cases} 1 & (u,v) \in \Omega_1 \\ 0 & (u,v) \notin \Omega_1 \end{cases} \quad 1 \leq u \leq n \quad 1 \leq v \leq m \quad (3.24)$$

Sviluppando il prodotto della matrice definita in (3.22) con la matrice definita in (3.24) si ricava il seguente insieme di indici:

$$\bar{j} : [A] \cdot [I_1] = 1 \quad 1 \leq u \leq n \quad 1 \leq v \leq m \quad (3.25)$$

A questo punto si considerano le coordinate dei punti nello spazio globale aventi come indice un intero appartenente all'insieme: \bar{j} . Questi vengono usati come punti iniziali per fare l'operazione analoga a quella descritta in (3.21) per le immagini successive. Questo procedimento è compiuto per ogni sagoma prendendo come centri di voxel iniziali quelli ottenuti dalle proiezioni precedenti (figura 3.15).

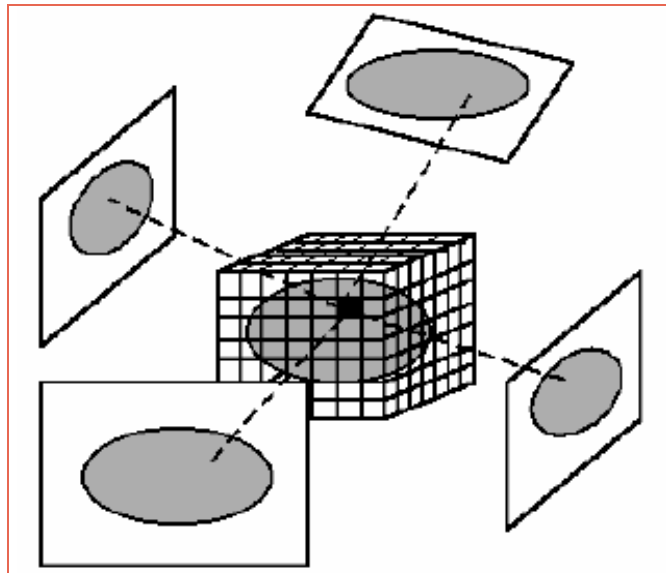


Fig. 3.15

3.4 Valutazione della risoluzione

Con la tecnica usata si definisce la risoluzione che si vuole raggiungere con la discretizzazione del box iniziale di voxel utilizzati per ricostruire la forma tridimensionale dell'oggetto preso in esame.

Si consideri un volume iniziale di dimensioni: l_x , l_y ed l_z ed ancora una discretizzazione lungo le tre direzioni: Δ_x , Δ_y e Δ_z allora si ha che la minima distanza lungo le varie direzioni che si può appurare sarà data dalla seguente relazione:

$$a_x = \frac{l_x}{\Delta_x}; a_y = \frac{l_y}{\Delta_y}; a_z = \frac{l_z}{\Delta_z} \quad (3.26)$$

Nella figura 3.27 sono riportati due esempi di ricostruzione per differenti scelte di definizione del box iniziale.

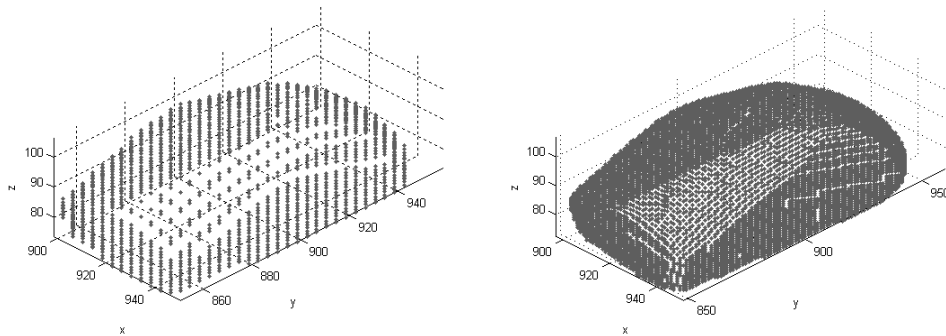


Fig. 3.27

E' importante notare che la scelta della risoluzione iniziale non è indipendente dal tipo di sensore ottico utilizzato. Ogni immagine per un'assegnata distanza dalla regione di spazio inquadrata sarà una discretizzazione della scena in base al numero di pixel che compone il sensore ottico, quindi è evidente che un oggetto più vicino al sensore sarà visualizzato più grande rispetto ad uno posizionato più lontano e quindi l'errore che si commette al variare di una unità di discretizzazione del sensore cresce al crescere della distanza per una assegnata lunghezza focale:

$$a_u^c = \delta_u \cdot \frac{W_\xi}{f}; a_v^c = \delta_v \cdot \frac{W_\xi}{f} \quad (3.26)$$

Ovviamente bisogna verificare che la risoluzione che si definisce in fase iniziale con la scelta della risoluzione non sia superiore a quella ottenibile con il sensore ottico utilizzato.

A tal fine una volta ricostruito l'oggetto si valuta mediante la (3.26) per ogni punto la minima variazione di lunghezza che può essere percepita con la telecamera e si valuta se è compatibile con la quella ipotizzata in fase iniziale.

La tecnica proposta è lenta perché pur essendo necessaria ai fini della ricostruzione bisogna proiettare ogni punto del box iniziale nei piani immagine. Inoltre c'è bisogno di un elevato numero di foto ed il procedimento non può essere compiuto in real-time. Bisogna notare comunque che tale metodologia può essere usata in modo molto semplice per avere una valutazione di massima del volume e della forma di un oggetto. Nel prossimo capitolo sarà trattata in maggiore dettaglio la tecnica di triangolazione laser che ha il vantaggio di essere molto più rapida pur mantenendo una qualità di ricostruzione accettabile.

CAPITOLO 4

Ricostruzione tridimensionale mediante Triangolazione Laser

4.1 Principio della triangolazione laser

Nel capitolo 2 è stata fatta una sintesi dei metodi a triangolazione attiva, in questo paragrafo viene descritto un metodo per la ricostruzione di una superficie che usa il principio di triangolazione attiva scegliendo come luce strutturata quella generata da un emettitore laser in posizione fissa rispetto ad una telecamera che inquadra un porzione di superficie illuminata dal laser stesso [19].

I punti della linea laser nel piano immagine rappresentano l'intersezione tra i raggi ottici e che passano per i punti di intersezione tra il laser e l'oggetto. Come visto nei precedenti capitoli grazie ad una matrice di trasformazione omogenea è possibile esprimere le coordinate dei punti appartenenti al piano immagine, espresse in pixel, in un sistema di riferimento locale.

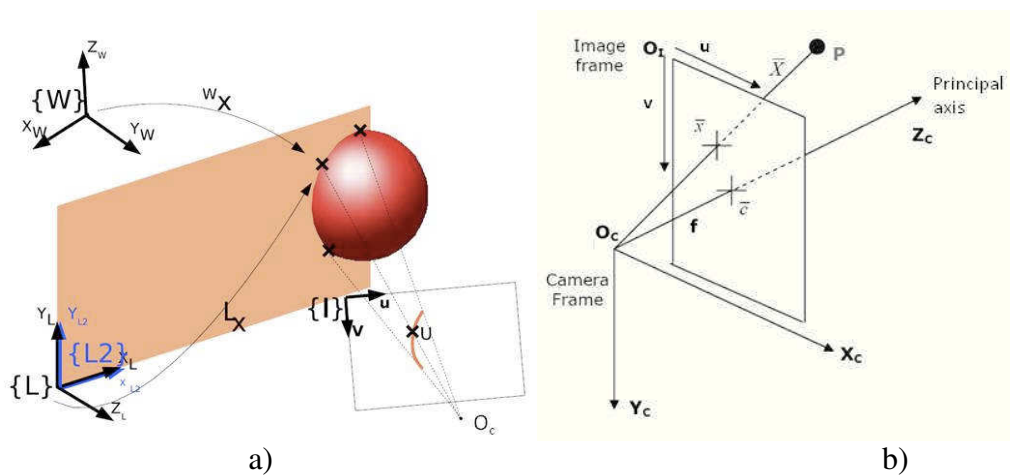


Fig. 4.1

Nella figura 4.1 è mostrato lo schema del sistema di scansione: $\{W\}$ è il sistema di riferimento globale, $\{I\}$ è il piano immagine con il sistema di coordinate $\{u, v\}$ e $\{L\}$ è la terna riferita all'emettitore laser. $\{L_2\}$ è il piano laser che contiene il fascio laser e la superficie scannerizzata appartenente all'oggetto in esame, esso coincide con il piano (x, y) della terna laser $\{L\}$ [20].

Partendo dalle coordinate in pixel (u, v) , nel piano immagine è possibile scrivere le coordinate dei punti appartenenti alla superficie scannerizzata nel sistema di riferimento della camera grazie alla (4.1). La terna camera è posizionata nel punto focale come si vede dalla fig. 4.2.

$$\begin{Bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{Bmatrix} = \begin{bmatrix} \delta_x & 0 & 0 & -\delta_x u_0 \\ 0 & \delta_y & 0 & -\delta_y v_0 \\ 0 & 0 & 0 & f \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} \quad (4.1)$$

Con:

(u_0, v_0) : coordinate del punto focale nel piano immagine;

(δ_x, δ_y) : dimensioni fisiche del sensore lungo le direzioni u e v ;

f : lunghezza focale

È possibile scrivere l'espressione del raggio ottico di un generico punto nel piano immagine in forma parametrica scegliendo come parametro t .

$$\begin{cases} x_c = (u - u_0)\delta_x t \\ y_c = (v - v_0)\delta_y t \\ z_c = ft \end{cases} \quad (4.2)$$

La terna $\{L\}$ è posizionata rispetto alla terna camera mediante successivi movimenti di traslazione e di rotazione ottenuti nella seguente sequenza:

- I. Traslazione Δx_{lc} lungo l'asse x_c ;
- II. Traslazione Δy_{lc} lungo l'asse y_c ;
- III. Traslazione Δz_{lc} lungo l'asse z_c ;
- IV. Rotazione φ_c intorno all'asse z_c ;
- V. Rotazione θ_c intorno all'asse y_c ;
- VI. Rotazione ψ_c intorno all'asse x_c ;

Nell'equazione (4.3) viene definite la matrice di trasformazione tra la terna laser e la terna camera:

$$\begin{aligned}
 {}^c T_l &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\psi_{lc}) & -\sin(\psi_{lc}) & 0 \\ 0 & \sin(\psi_{lc}) & \cos(\psi_{lc}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos(\theta_{lc}) & 0 & \sin(\theta_{lc}) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\theta_{lc}) & 0 & \cos(\theta_{lc}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \\
 &\begin{bmatrix} \cos(\varphi_{lc}) & -\sin(\varphi_{lc}) & 0 & 0 \\ \sin(\varphi_{lc}) & \cos(\varphi_{lc}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta z_{lc} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \Delta y_{lc} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & \Delta x_{lc} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
 \end{aligned} \tag{4.3}$$

Il piano laser $\{L2\}$ coincide con il piano (x,y) della terna laser $\{L\}$, così esso contiene sicuramente i seguenti tre punti:

$$\{p\}_l = \{0,0,0,1\}^T; \{q\}_l = \{1,0,0,1\}^T; \{r\}_l = \{0,0,1,1\}^T \tag{4.4}$$

Nella terna camera si ha:

$$\{p_x, p_y, p_z, 1\}_c^T = [{}^c T_l]^{-1} \{p\}_l; \{q_x, q_y, q_z, 1\}_c^T = [{}^c T_l]^{-1} \{q\}_l; \{r_x, r_y, r_z, 1\}_c^T = [{}^c T_l]^{-1} \{r\}_l \tag{4.5}$$

È possibile ottenere l'equazione del piano laser nella terna camera risolvendo l'equazione (4.6) in x_c, y_c e z_c .

$$\det \begin{bmatrix} x_c - p_x & y_c - p_y & z_c - p_z \\ q_x - p_x & q_y - p_y & q_z - p_z \\ r_x - p_x & r_y - p_y & r_z - p_z \end{bmatrix} = 0 \tag{4.6}$$

Se si definiscono:

$$M_x = \det \begin{bmatrix} q_y - p_y & q_z - p_z \\ r_y - p_y & r_z - p_z \end{bmatrix}; M_y = \det \begin{bmatrix} q_x - p_x & q_z - p_z \\ r_x - p_x & r_z - p_z \end{bmatrix}; M_z = \det \begin{bmatrix} q_x - p_x & q_y - p_y \\ r_x - p_x & r_y - p_y \end{bmatrix}$$

L'equazione di {L2} nella terna camera sarà:

$$(x_c - p_x)M_x - (y_c - p_y)M_y + (z_c - p_z)M_z = 0 \quad (4.7)$$

E' possibile poi valutare le coordinate x_c , y_c e z_c , nella terna camera, risolvendo il sistema (4.8) rispetto all'incognita t :

$$\begin{cases} x_c = (u - u_0)\delta_x t \\ y_c = (v - v_0)\delta_y t \\ z_c = ft \\ (x_c - p_x)M_x - (y_c - p_y)M_y + (z_c - p_z)M_z = 0 \end{cases} \quad (4.8)$$

Si ottiene la soluzione:

$$t = \frac{p_x M_x - p_y M_y + p_z M_z}{(u - u_0)\delta_x M_x - (v - v_0)\delta_y M_y + f M_z} \quad (4.9)$$

L'equazione (4.9) permette di calcolare nella terna camera, le coordinate dei punti della curva appartenente alla superficie dell'oggetto illuminata dal fascio laser a partire delle loro coordinate in pixel (u,v) nel piano immagine. In questo modo è possibile ottenere la superficie 3D di un oggetto mediante un fascio laser.

4.2 Determinazione della linea laser

Uno step molto importante per la ricostruzione tridimensionale con triangolazione laser è la procedura di segmentazione dell'immagine atta ad identificare solo i pixel identificativi della proiezione della superficie scansionata nel piano immagine [21, 22, 23], un esempio di immagine acquisita è mostrato nella figura 4.2.



Fig. 4.2

Il primo passo effettuato nella elaborazione dell'immagine è quello di individuare una regione dell'immagine contenete i punti di interesse in modo ad identificare l'intensità dei colori fondamentali (Rosso, Verde e Blu) dei pixel della linea laser, figura 4.3.

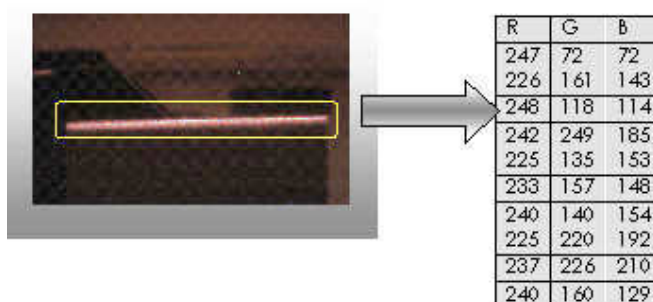


Fig. 4.3

Con il valore medio dei colori della linea scannerizzata è possibile definire un coefficiente di luminosità mediante la relazione (4.10):

$$s = \frac{\max(\text{mean}(R), \text{mean}(G), \text{mean}(B)) + \min(\text{mean}(R), \text{mean}(G), \text{mean}(B))}{2} \quad (4.10)$$

Grazie alla relazione (4.11) viene effettuata un'analisi dell'intensità di tutta l'immagine:

$$L(u, v) = \frac{\max(R, G, B) + \min(R, G, B)}{2} \quad (4.11)$$

Mediante la relazione (4.11), la matrice che contiene I valori dei tre colori principali RGB è trasformata in una matrice L che rappresenta l'intensità dell'immagine. Questa può essere considerata l'immagine in scala di grigi riferita a quella di partenza in formato RGB., figura 4.4 a) e b).

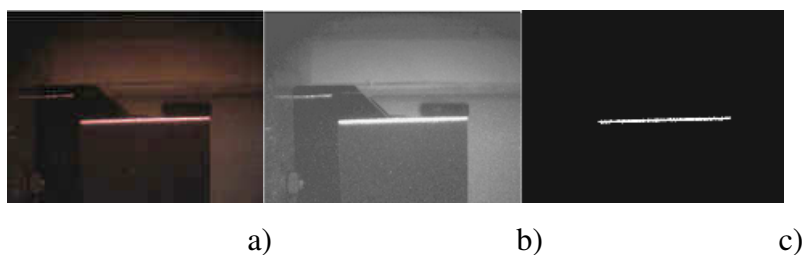


Fig. 4.4

Con la relazione (4.12), è possibile definire una matrice binaria I_b . Tale matrice identifica quegli elementi della matrice L aventi un'intensità nel range del 15% del coefficiente s .

$$I_b(u,v) = \begin{cases} 1 & \text{se } 0.85 \cdot s \leq L(u,v) \leq 1.15 \cdot s \\ 0 & \text{altrimenti} \end{cases} \quad (4.12)$$

Nella figura 4.4 c) è mostrata una rappresentazione della matrice I_b .

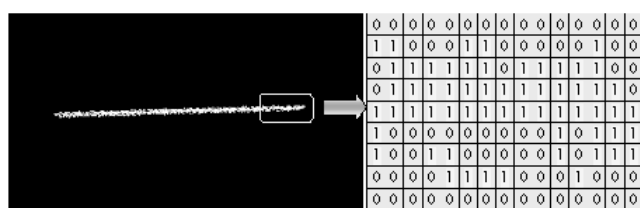


Fig. 4.5

Nella matrice I_b l'area scannerizzata è rappresentata dai pixel con valore unitario. L'area contenente tali pixel nel piano immagine dipende dalla reale dimensione dello spessore della linea laser a da altri fattori come riflessione e l'inclinazione della superficie dell'oggetto illuminato. La procedura di ricostruzione tridimensionale è basata sulla triangolazione, nella quale trattazione non viene considerato lo spessore

della curva scannerizzata, quindi è necessario che si associ a questa non un area ma una curva.

Per fare ciò si è proceduto a fare una media geometrica lungo le colonne della matrice I_b , in questo modo si seleziona lungo la direzione perpendicolare al fascio laser solo un pixel rappresentativo dei punti della superficie illuminata dal fascio laser. Questa procedura viene eseguita mediante la relazione (4.13).

$$\Omega := \{u, v\}; \bar{\Omega} := \{\bar{u}, \bar{v}\} \Rightarrow \left\{ \bar{u} = \frac{\sum_{u,v \in \Omega} u I_b(u, v)}{\sum_{u,v \in \Omega} I_b(u, v)} \neq \infty, \bar{v} = v : \sum_{u,v \in \Omega} I_b(u, v) \neq 0 \right\} \quad (4.13)$$

Successivamente viene definita una nuova matrice h_b come segue:

$$h_b(u, v) = \begin{cases} 1 & \text{se } \{u, v\} \in \bar{\Omega} \\ 0 & \text{altrimenti} \end{cases} \quad (4.14)$$

La matrice h_b è una matrice binaria delle stesse dimensioni di I_b che rappresenta la curva di scansione come si vede nella figura 4.6. Questa linea può essere considerata la curva media dell'area scannerizzata.

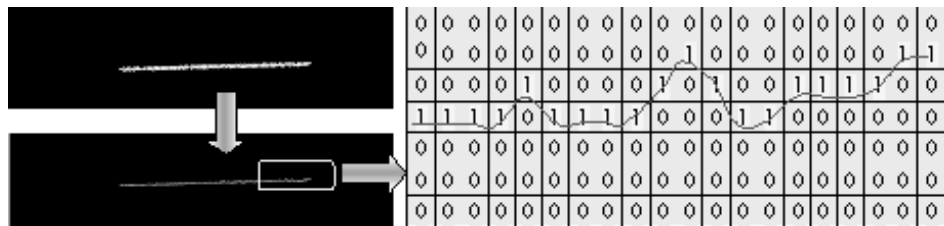


Fig. 4.6

Le relazioni (4.11), (4.12) e (4.13) rappresentano i passi da effettuare per l'elaborazione dell'immagine necessari per identificare la proiezione della superficie scannerizzata nell'immagine. I pixel così ottenuti sono quelli necessari per la ricostruzione 3D.

4.3 Calibrazione del sistema laser scanner

La procedura di calibrazione è necessaria per identificare tutti i parametri del modello in modo tale da ottenere una giusta rappresentazione tridimensionale [20, 22, 24]. Per questo motivo è stata sviluppata una metodologia per un modulo laser scanner standard.

Il sistema per la calibrazione è costituito da una guida sulla quale è fissato il modulo laser scanner costituito da una web cam, da un emettitore laser, e da un micrometro digitale sul quale è stato montato un bersaglio (figura 4.7).

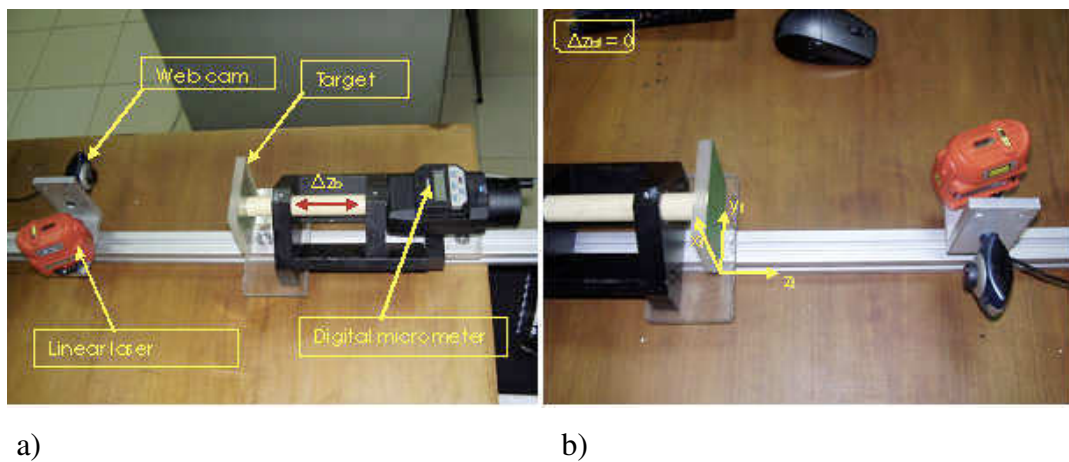


Fig. 4.7

Si fissa una terna $\{O_f, x_f, y_f, z_f\}$ nello spigolo basso del bersaglio come mostrato nella figura 4.7 b) e la movimentazione del micrometro è indicata con la grandezza Δz_{bf} . Nella figura 4.8 sono mostrati alcuni step della procedura di calibrazione.



Fig. 4.8

La procedura di calibrazione è basata su un set di immagini acquisite per un differente posizionamento del bersaglio rispetto al modulo laser scanner.

Nella figura 4.9 sono mostrate dieci delle immagini usate per la calibrazione per differenti valori di Δz_{bf} .



Fig. 4.9

Lo scopo della calibrazione è quello di identificare i parametri che definiscono la relazione tra la terna laser e la terna camera, figura 4.1.

Per ogni step della calibrazione è possibile definire la posizione relativa tra la terna fissa e la terna solidale al bersaglio così come mostrato dalla (4.15):

$$[{}^fT_b]_i = [1 \ 0 \ 0 \ 0; 0 \ 1 \ 0 \ 0; 0 \ 0 \ 1 \ -\Delta z_{fb}^i; 0 \ 0 \ 0 \ 1] \quad (4.15)$$

La matrice di trasformazione tra la terna fissa e la terna camera può essere ottenuta in modo analogo a quello che porta alla definizione di $[{}^cT_l]$ dalla relazione (4.3), e essa è funzione di sei parametri:

$$[{}^fT_c] = f(\Delta x_{cf}, \Delta y_{cf}, \Delta z_{cf}, \psi_{cf}, \theta_{cf}, \varphi_{cf}) = f(\pi_{cf}) \quad (4.16)$$

Avendo indicato con π_{cf} il set di parametri della matrice di trasformazione $[{}^fT_c]$.

L'equazione (4.3) allora può essere scritta come segue:

$$[{}^cT_l] = g(\Delta x_{lc}, \Delta y_{lc}, \Delta z_{lc}, \psi_{lc}, \theta_{lc}, \varphi_{lc}) = f(\pi_{lc}) \quad (4.17)$$

Avendo indicato con π_{lc} il set di parametri della matrice di trasformazione $[{}^cT_l]$.

Per ogni step è possibile valutare le coordinate dei punti nel piano immagine rispetto alla terna camera mediante le relazione (4.5) e (4.13):

$$\{\bar{x}_c, \bar{y}_c, \bar{z}_c, 1\}_i^T = f(\pi_{lc}, f) \cdot \{\bar{u}, \bar{v}, 0, 1\}_i^T \quad (4.18)$$

Grazie alla relazione (4.14), è possibile scrivere:

$$\begin{aligned} \{\bar{x}_f, \bar{y}_f, \bar{z}_f, 1\}_i^T &= [{}^f T_c]^{-1} \cdot \{\bar{x}_c, \bar{y}_c, \bar{z}_c, 1\}_i^T = \\ &[{}^f T_c]^{-1} \cdot f(\pi_{lc}, f) \cdot \{\bar{u}, \bar{v}, 0, 1\}_i^T = \\ &\omega(\pi_{cf}, \pi_{lc}, f) \cdot \{\bar{u}, \bar{v}, 0, 1\}_i^T \end{aligned} \quad (4.19)$$

Il problema di ottimizzazione può essere formulato nel modo seguente:

$$\min_{\rho \in R^{13}} F(\rho)^2 \quad (4.20)$$

$F(\rho)$ è una funzione vettoriale definita come segue:

$$\begin{aligned} F(\rho) = \{ &(\bar{z}_f^{i+1} - \bar{z}_f^i) - (\Delta \bar{z}_{bf}^{i+1} - \Delta \bar{z}_{bf}^i), \min(\bar{x}_f^i), \max(\bar{x}_f^i) \\ &- a, \max(\bar{z}_f^i) - \min(\bar{z}_f^i), \max(\bar{y}_f^i) - \min(\bar{y}_f^i), \min(\bar{y}_f^i), \min(\bar{z}_f^i) \}^T \end{aligned} \quad (4.21)$$

Con una procedura di ottimizzazione ai minimi quadrati è stata ottenuta l'identificazione dei parametri $[\pi_{cf}, \pi_{lc}, f]$. E' da notare che si è proceduto anche all'identificazione della lunghezza focale della telecamera, infatti pur avendo stimato i parametri intrinseci delle telecamera con una calibrazione apposita per i sensori ottici, per valutare l'effettiva in alterazione della lunghezza focale se ne è tenuto conto anche nella calibrazione dell'intero sistema laser scanner.

Nella figura 4.10 è mostrato il risultato della calibrazione, in particolar modo sono rappresentate le successive posizioni della terna solidale al bersaglio, la terna laser, la terna camera e il piano laser. Inoltre sono stati rappresentati tutti i punti delle superfici scannerizzate nel riferimento fisso.

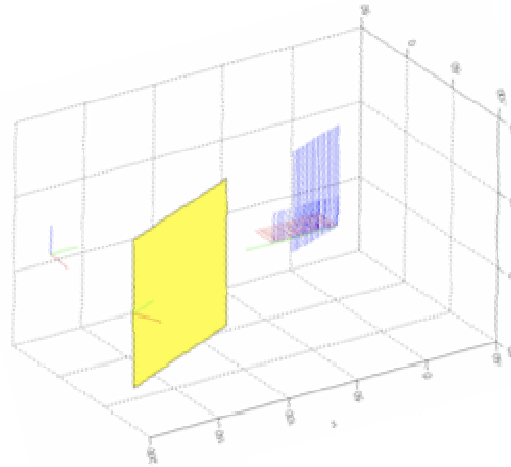


Fig. 4.10

4.4 Valutazione della risoluzione

Come si è già visto nel capitolo 3 un passo molto importante nella ricostruzione tridimensionale è la valutazione della risoluzione. Nei seguenti paragrafi saranno elencati alcuni dei principali contributi alla bontà dei risultati e alcune proprietà che riguarda la possibilità di scannerizzare o meno delle regioni di spazio.

4.4.1 Errore dovuto alla discretizzazione del sensore ottico

Come visto nel capitolo precedente una fonte di errore nella ricostruzione tridimensionale è dovuta sicuramente al sistema telecamera comprensivo dell'ottica e del sensore, anche a valle di una procedura di calibrazione. La risoluzione della camera è funzione della distanza dell'oggetto osservato e la legge di variazione è individuata dalla relazione (3.26).

4.4.2 Errore dovuto al modello di laser scanner

E' possibile definire la risoluzione del laser scanner in base alle relazioni (4.8) e (4.9).

Una variazione (α, β) delle coordinate (u, v), generano una variazione in t ottenuta dalla (4.9):

$$\Delta t = \frac{-(p_x M_x - p_y M_y + p_z M_z) \cdot (\alpha \delta_x M_x - \beta \delta_y M_y)}{(u \delta_x M_x + \alpha \delta_x M_x - v \delta_y M_y - \beta \delta_y M_y + f M_z) \cdot (u \delta_x M_x - v \delta_y M_y + f M_z)} \quad (4.22)$$

dove:

- α è la variazione espressa in pixel lungo la direzione u ;
- β è la variazione espressa in pixel lungo la direzione v ;

La variazione del parametro t , permette di definire l'espressione della risoluzione del sistema laser scanner. Infatti, grazie all'equazione (4.8) è possibile ottenere la variazione delle coordinate nella terna camera in funzione della variazione delle coordinate nel piano immagine:

$$\begin{cases} \Delta x_c = \alpha \delta_u \cdot \frac{p_x M_x - p_y M_y + p_z M_z}{(u - u_o + \alpha) \delta_x M_x - (v - v_o + \beta) \delta_y M_y + f M_z} + (u - u_o) \delta_x \cdot \Delta t \\ \Delta y_c = \beta \delta_v \cdot \frac{p_x M_x - p_y M_y + p_z M_z}{(u - u_o + \alpha) \delta_x M_x - (v - v_o + \beta) \delta_y M_y + f M_z} + (v - v_o) \delta_y \cdot \Delta t \\ \Delta z_c = \Delta t \cdot f \end{cases} \quad (4.23)$$

Una variazione unitaria di coordinate nel piano immagine ($\alpha=1$ e $\beta=0$, o $\alpha=0$ e $\beta=1$, o $\alpha=1$ e $\beta=1$), permette di definire tre parametri di risoluzione come segue:

$$\begin{cases} a^{SL}_x = \Delta x_c \\ a^{SL}_y = \Delta y_c \\ a^{SL}_z = \Delta z_c \end{cases} \Rightarrow \text{with} \Rightarrow \begin{cases} (\alpha, \beta) = (0,1) \\ \text{or} \\ (\alpha, \beta) = (1,0) \\ \text{or} \\ (\alpha, \beta) = (1,1) \end{cases} \quad (4.24)$$

Come è mostrato nella figura 4.11 la risoluzione peggiore del laser scanner per una variazione unitaria di pixel è quella lungo la direzione z_c . Si può notare che a_z ha un minimo per valori di $v = m$ e cresce fino a $v = l$ con una legge non lineare.

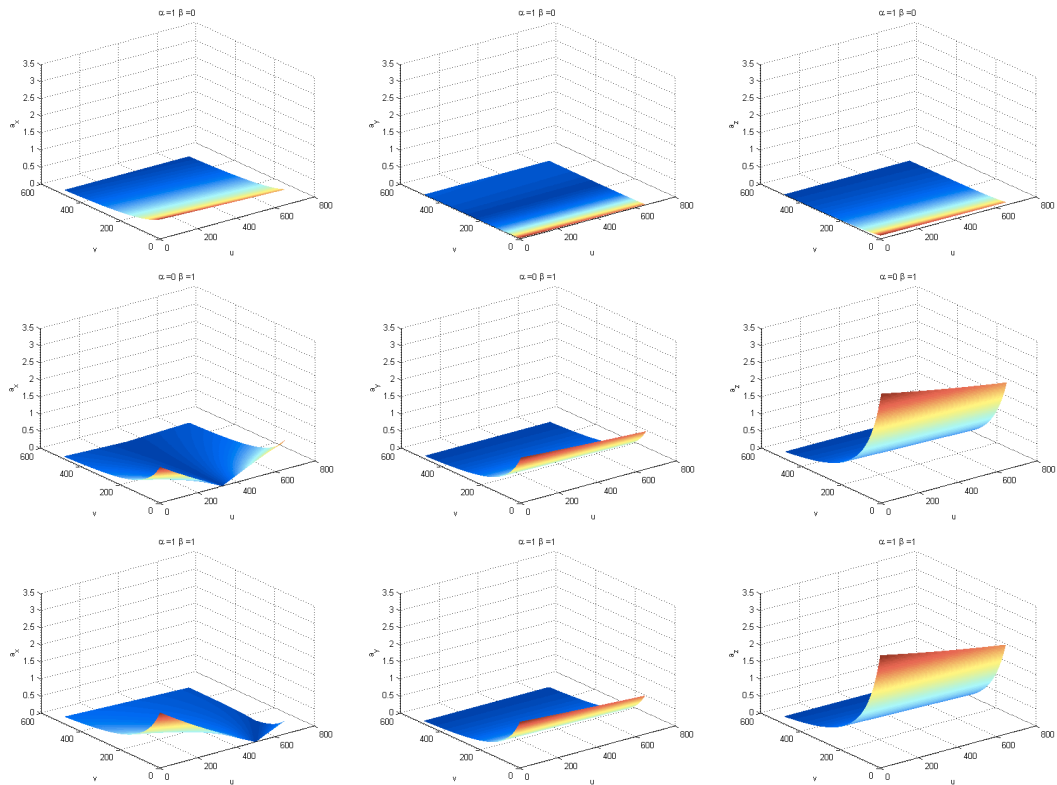


Fig. 4.11

4.4.3 Precisione dovuta al sistema laser scanner

Un fascio laser con una elevata energia viene riflesso da superfici differenti in modo completamente diverso, questo può comportare un problema per l'acquisizione dell'informazione della superficie scannerizzata. Inoltre, una elevata luminosità del fascio può dar luogo nel sensore al fenomeno conosciuto come "blooming", in questo caso parte dei raggi di luce riflessa è raccolta nei pixel circostanti l'area di effettiva identificazione del fascio laser. Tutti questi effetti sono causa di salti nell'immagine e di buchi. Uno degli errori più significati nella ricostruzione sicuramente è dovuto allo spessore del fascio luminoso. Un altro limite della ricostruzione è che la linea laser ha uno spessore è che quindi l'intersezione dei raggi ottici riferiti ai pixel non è più con un piano ma con un fascio di piani paralleli.

La risoluzione a^c diventa peggiore se si considera lo spessore della linea laser. Questo parametro si può misurare in termini di pixel nel piano immagine e può essere espresso con due componenti: t_{hu} che equivale alla dimensione dello spessore lungo la direzione u e t_{hv} quella lungo la direzione v . Si sono considerati due tipi di

dimensioni per il caso generale che il fascio laser sia orientato parallelamente a u oppure a v .

Un'espressione della risoluzione al variare dello spessore può essere ottenuta nella direzione z della terna camera grazie alla (4.26), nella quale (α, β) sono gli spessori generalizzati del fascio laser (t_{hu}, t_{hv}) :

$$\begin{cases} a_x = \Delta x_c \\ a_y = \Delta y_c \Rightarrow \text{with} \Rightarrow (\alpha, \beta) = (t_{hu} / 2, t_{hv} / 2) \\ a_z = \Delta z_c \end{cases} \quad (4.28)$$

L'equazione (4.28) definisce la risoluzione della ricostruzione laser scanner, e permettono di valutare la risoluzione di ogni punti della proiezione della superficie scansionata ottenuta dall'elaborazione dell'immagine.

4.4.4 Scanner range

Un'altra caratteristica del laser scanner è la minima e la massima distanza tra un generico punto di una superficie illuminata e il piano immagine. Questi parametri definiscono il range della procedura di scansione. Diminuendo l'angolo θ di inclinazione del piano laser rispetto al piano $x_c z_c$ della terna camera per una fissata distanza s tra le origini dei due lungo y_c , il range di scansione aumenta (figura 4.12).

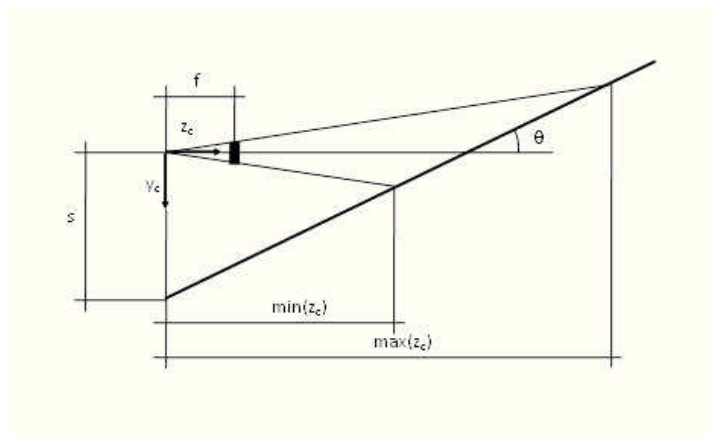


Fig. 4.12

In generale questa anche se sembrerebbe una scelta utile può causare una notevole diminuzione della risoluzione lungo z come mostrato in figura 4.13.

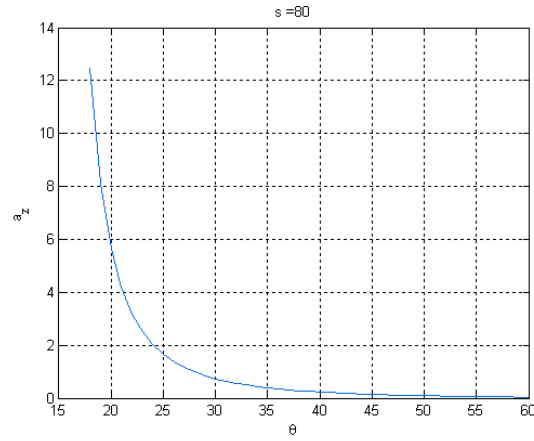


Fig. 4.13

Per un sistema avente $s = 90$ mm e $\theta = 23^\circ$ si ottiene: $max(z_c) = 525$ mm and $min(z_c) = 124$ mm;

CAPITOLO 5

Utilizzo di un robot antropomorfo per la digitalizzazione di superfici

5.1 Robot

Nel capitolo 3 e nel capitolo 4 sono stati presentati in dettaglio due metodi ottici per la ricostruzione tridimensionale; di seguito viene descritta una procedura di valutazione sperimentale delle due tecniche. Per fare ciò si è utilizzato un robot, in particolare saranno evidenziati quelli che sono i vantaggi dell'utilizzo di un tale sistema automatizzato nella procedura di acquisizione della forma 3D.

Il robot è costituito da:

- 1) una struttura meccanica o manipolatore che consiste in una serie di corpi rigidi (link) interconnessi tra di loro per mezzo di articolazioni (giunti); è bene ricordare che esistono due tipologie di giunti che prendono il nome di rotoidale o traslazionale a seconda se il moto relativo tra i due bracci è rispettivamente di tipo rotatorio o traslatorio; nel manipolatore si individuano una struttura portante che ne assicura la mobilità, un polso che conferisce destrezza e un organo terminale che esegue il compito per cui il robot è utilizzato,
- 2) attuatori che imprimono il movimento al manipolatore attraverso l'azionamento dei giunti; si impiegano usualmente motori elettrici, idraulici e talvolta pneumatici,
- 3) sensori che misurano lo stato del manipolatore (sensori propriocettivi) ed eventualmente lo stato dell'ambiente (sensori esteroceettivi),
- 4) una unità di governo (calcolatore) con funzioni di controllo e supervisione dei movimenti del manipolatore.

5.1.1 Struttura del robot [1]

Il modello del Robot Revoluto (vedi Figura 5.1) risulta sostanzialmente essere costituito da:

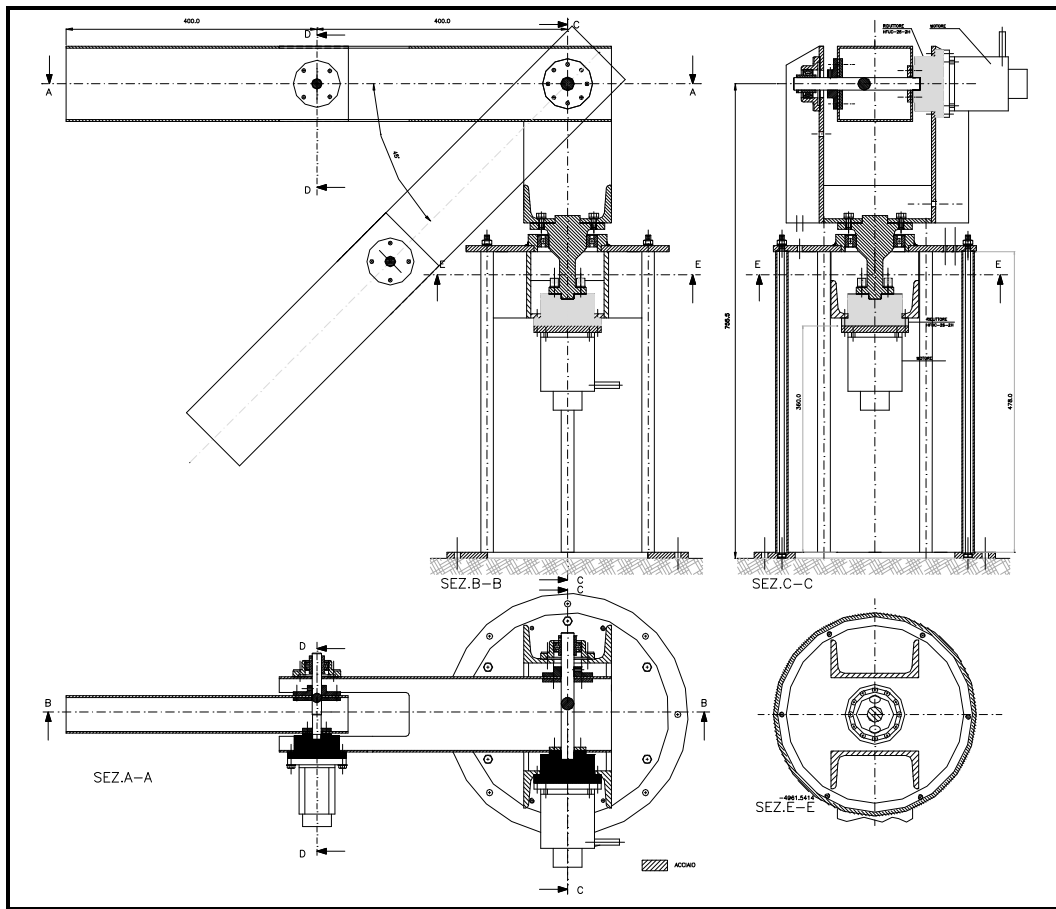


Fig. 5.1

a) Una base di appoggio cilindrica di acciaio su cui sono fissate sei colonnine, bullonate all'altra estremità, dotata all'interno di una struttura di rinforzo fissa, costituita da tre profilati a C saldati (vedi Figura 5.2).

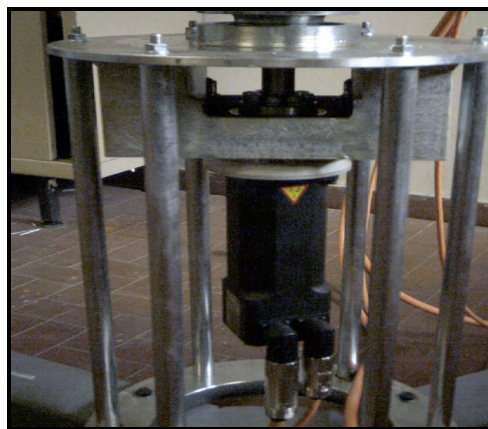


Fig. 5.2

b) Tre profilati a C esterni, anch'essi di acciaio saldati in modo da formare un giunto a "forchetta", del tutto simili ai precedenti, ma mobili rispetto alla struttura (vedi figura 5.3).

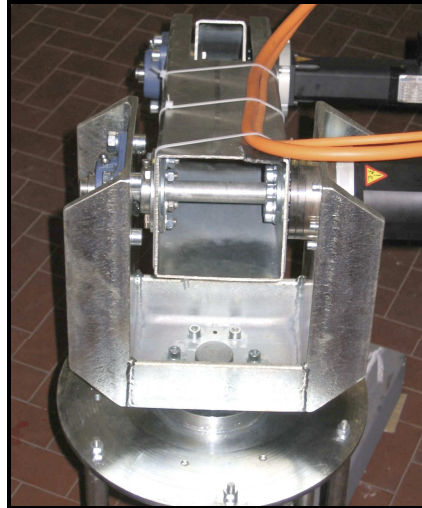


Fig. 5.3

Questo giunto a forchetta, che costituisce il primo segmento link 1 del Robot, è vincolato alla base stessa tramite un perno di forma particolare (visibile parzialmente in figura 5.3). Infine, i due segmenti di acciaio posti in successione costituiscono i link 2 e 3. Il primo è vincolato tramite un perno, attorno a cui può ruotare, (vedi figura 5.3), al giunto a C, costituente il primo link; il secondo è, a sua volta, vincolato al primo, nel medesimo modo (vedi figura 5.4).

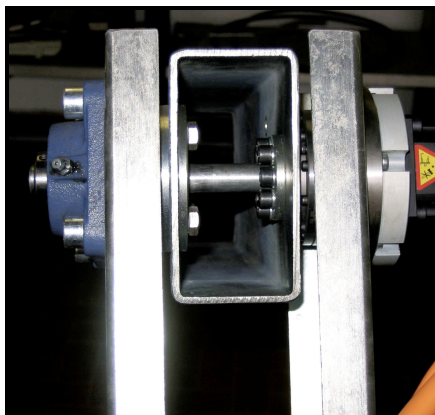


Fig. 5.4

Tutti gli elementi di cui è costituito il robot sono in acciaio. La geometria scelta per gli elementi costruttivi risulta abbastanza semplice e ciò perché si è pensato di ricorrere, ove possibile, a componenti e profili di tipo unificato. Lo scopo è di ottenere una soluzione sufficientemente economica e facile da realizzare. L'impiego di elementi costruiti mediante lavorazioni dal pieno è stato limitato al massimo. Per quanto concerne le proprie capacità di movimentazione, il Robot possiede tre assi, tutti di tipo rotazionale. Ciò gli conferisce una buona flessibilità nei movimenti ed un sufficientemente ampio volume di lavoro, in considerazione dei suoi più comuni campi applicativi. La flessibilità del Robot consente di soddisfare differenti e disparate esigenze applicative, a seconda dell'utensile montabile alla parte terminale dell'ultimo link.

Gli organi meccanici possiedono una elevata rigidità grazie alla scelta costruttiva, che prevede gli assi di simmetria dei segmenti tutti contenuti in uno stesso piano. Ciò ha implicato l'adozione dei già menzionati giunti "a forchetta" (gli alberi poggiano su due cuscinetti montati da parti opposte rispetto al carico), mentre l'impiego di un giunto di tipo "a sbalzo" è stato limitato al solo moto di cintura.

Per la realizzazione dei segmenti che costituiscono il braccio e l'avambraccio, si è scelto di impiegare semilavorati di tipo scatolare, poiché risultano leggeri e permettono di essere eventualmente attraversati da cavi elettrici.

La movimentazione dei link 2 e 3 viene ottenuta tramite alberini sostenuti, da una parte dal cuscinetto del riduttore, dall'altra, da un cuscinetto a sfere orientabili; ciò consente di poter tollerare imperfezioni di lavorazione, senza alcuna influenza sul buon funzionamento dei giunti stessi. Il rispetto della posizione degli assi è, infatti, affidato al montaggio dei riduttori, effettuato il quale, l'asse dell'albero del giunto si centra automaticamente sull'asse del riduttore.

5.1.1.1 Azionamenti

I motori scelti per l'azionamento sono del tipo brushless, sincroni, della DANAHER MOTION®, serie DBL (due DBL3-H03250 e uno DBL2-M0080), per i quali si sono adottati i servoazionamenti SERVOSTAR™ 603.

I motori sono stati concepiti in modo particolare come azionamento per dispositivi di movimentazione, pertanto possono essere messi in funzione solo come componenti integranti dell'impianto e sono stati acquisiti assieme ai relativi servoazionamenti consigliati dalla casa produttrice.

I motori sono trifase, particolarmente adatti per applicazioni di alta precisione. La loro scelta è, quindi, motivata dalla volontà di garantire un azionamento, e relativo controllo, quanto migliore, preciso e affidabile possibile. Il loro montaggio è, sicuramente, una delle poche operazioni che richiedano accortezza ed accuratezza nelle lavorazioni. In particolare flangia e albero possono trovarsi in una condizione critica per il montaggio.

I servoazionamenti sono esclusivamente destinati ad essere comandati con una regolazione della velocità o della coppia, nel nostro caso è in velocità, mentre il controllo è in posizione. Essi sono digitali, particolarmente idonei per compiti di posizionamento su robot industriali, con requisiti elevati in termini di dinamica e stabilità.

Attenzione particolare è stata, ovviamente, dedicata al corretto allineamento in sede di lavorazione dei link, poiché eventuali spostamenti avrebbero comportato vibrazioni non ammesse, eventualmente, determinato la rottura dei cuscinetti a sfere e dell'accoppiamento stesso.

I servoazionamenti della serie SERVOSTAR™ 603, valutano la posizione del resolver del rotore ed alimentano i motori con correnti sinusoidali.

I freni di arresto sono predisposti come freni di stazionamento, cioè realizzano il blocco del rotore nello stato privo di tensione, e non sono adatti per frenare in modo continuo durante il funzionamento. Se il freno è rilasciato, il rotore può muoversi. I freni di arresto sono comandati direttamente dal servoazionamento; il rilascio dello avvolgimento del freno, in questo caso, avviene nel servoamplificatore, pertanto non è necessario alcun cablaggio supplementare.

I primi due motori, a cui è affidato il moto di cintura e quello di spalla, sono uguali, con una velocità massima di 3000 giri/min. L'ultimo, più piccolo, ha una velocità massima di 4500 giri/min.

Si è scelto di impiegare motori capaci di fornire valori di coppia massima superiori a quella che è stato calcolato sia necessaria. Questo in base alle seguenti considerazioni:

- ↪ Motori più grandi garantiscono un assorbimento di corrente alquanto inferiore a quella massima, a beneficio dell'affidabilità
- ↪ Hanno un costo non molto maggiore di altri di minor potenza
- ↪ Il peso non è eccessivamente superiore a motori di potenza inferiore

5.1.1.2 I riduttori

L' utilizzo dei riduttori Harmonic Drive è stato necessario, oltre che opportuno, per poter garantire un prototipo affidabile e preciso, nonostante il maggiore costo di tali riduttori, rispetto ad altri disponibili in commercio. Inoltre, essi assicurano il minimo ingombro ed evitano di appesantire la struttura, essendo piccoli e costituiti da tre soli elementi, e questo pur realizzando rapporti di riduzione molto alti: nel nostro caso è di 100:1.

I tre elementi del riduttore sono, sinteticamente, descritti nella figura 5.5.

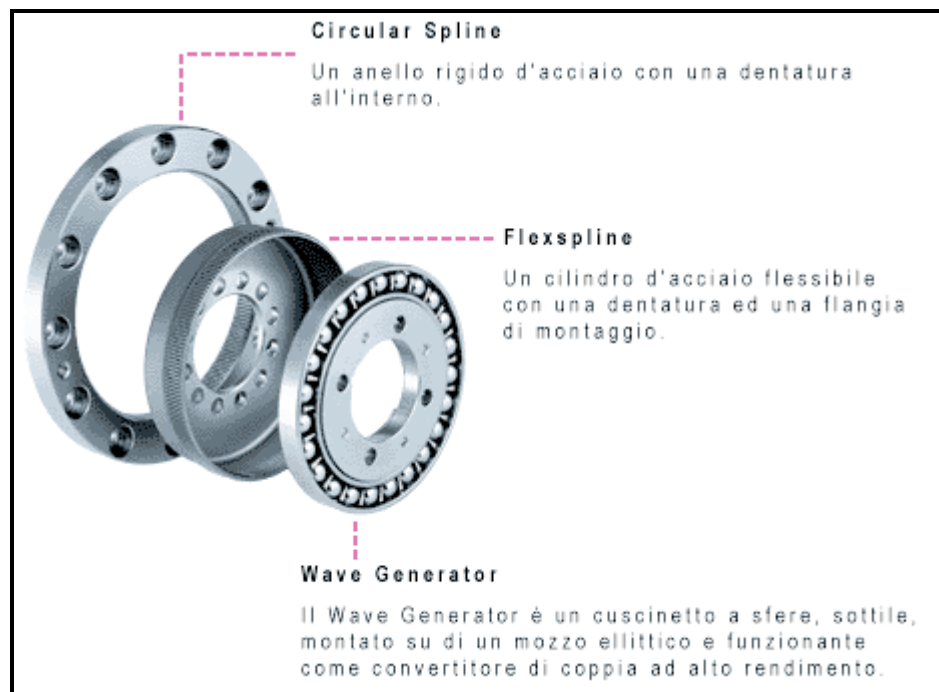


Fig. 5.5

I riduttori lavorano con gioco zero tra i denti in presa, grazie al precarico naturale della dentatura. Si possono ottenere rendimenti fino al 85%. Nelle dentature Harmonic Drive, i denti vengono in contatto con un movimento puramente radiale ed hanno velocità di scorrimento pari a zero, anche alle alte velocità. Le perdite, dovute all'attrito tra i denti e l'usura, sono praticamente trascurabili. I riduttori mostrano un'altissima rigidità torsionale, lungo tutta la gamma di velocità di lavoro. La curva caratteristica di rigidità, praticamente lineare, garantisce un ottimo comportamento operativo. Tra gli innegabili vantaggi, vi sono anche una maggiore facilità di montaggio e minori esigenze di manutenzione. Questo giustifica ampiamente la maggiore spesa sostenuta.

I riduttori appartengono alla serie Servo-Attuatori, sviluppati specificamente per garantire una precisione maggiore nelle applicazioni industriali, dove il controllo del moto è fondamentale ed è associato alla capacità di trasferire anche elevate coppie. Questa capacità è, in particolar modo, legata alla maggiore percentuale del numero dei denti in presa durante il moto.

Nelle figure 5.6 e 5.7 sono visibili, rispettivamente un riduttore già accoppiato con un motore elettrico e la sua vista in sezione.



Fig. 5.6

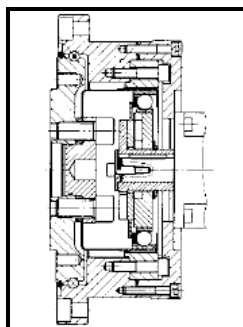


Fig.5.7

Essendo stati scelti motori capaci di fornire i valori di coppia riportati in precedenza, è stato necessario adottare i riduttori aventi le seguenti caratteristiche :

ASSE	Riduttore	Coppia media	Coppia max. picco
1 e 2	HFUC-25-2UH	108 Nm	284 Nm
3	HFUC-20-2UH	49 Nm	147 Nm

Tab. 5.1

La lubrificazione, richiesta dai riduttori, riguarda le cinque aree indicate nella figura 5.8 ed è realizzata con grasso.

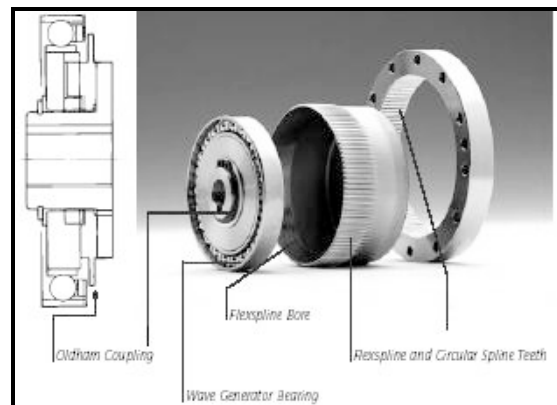


Fig. 5.8

Si è previsto di adottare per i riduttori un rapporto di trasmissione pari a 100; in tal modo la velocità massima di rotazione, a valle dei riduttori, sarà pari a 3.16 rad/s (≈ 30 giri/min) per i segmenti del Robot mossi dai motori più grandi ed a 4.7 rad/s (≈ 45 giri/min) per il segmento mosso dall' altro. Con tale rapporto di trasmissione, è lecito prevedere (vedi figura 5.9) che il rendimento dei riduttori non sia inferiore a: $\eta_r = 0,55$.

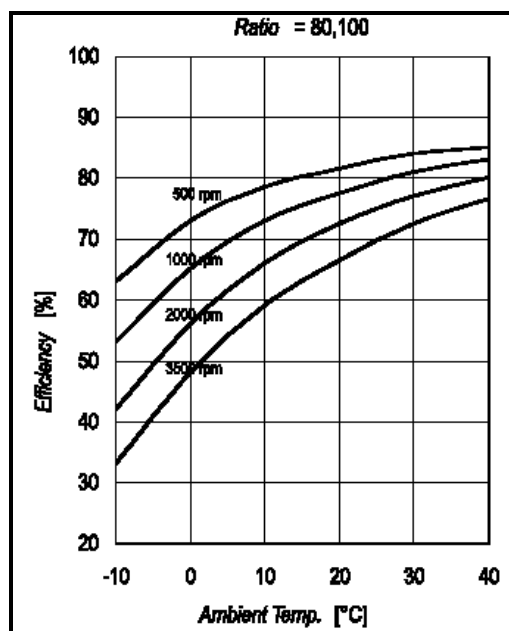


Fig. 5.9

5.1.1.3 I cuscinetti

I cuscinetti scelti sono prodotti dalla ditta SKF e precisamente si sono adottati i modelli: SKF-YAR-203/15-2F per il giunto 3 ed SKF-YAR-204-2F per il giunto 2 che uniti ai supporti Y-SKF-FY-15-TF (per il giunto 3) ed Y-SKF-FY-20-TF (per il giunto 2), realizzano sistemi semplici ed economici.

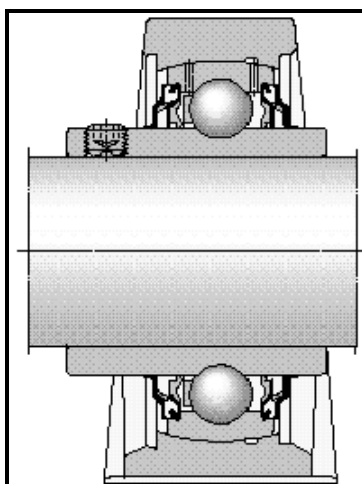


Fig. 5.10

I cuscinetti ad Y sono cuscinetti radiali ad una corona di sfere, muniti di guarnizioni striscianti, nelle quali la superficie diametrale esterna è sferica, tale da compensare errori di allineamento tra il cuscinetto ed il supporto. Sono fissati all'albero mediante viti.

Per quanto riguarda i supporti (vedi figura 5.11), essi sono a sezione quadrata, in ghisa, presentano delle tenute in feltro come protezione e sono lubrificati con grasso, di cui ne sono riempiti.

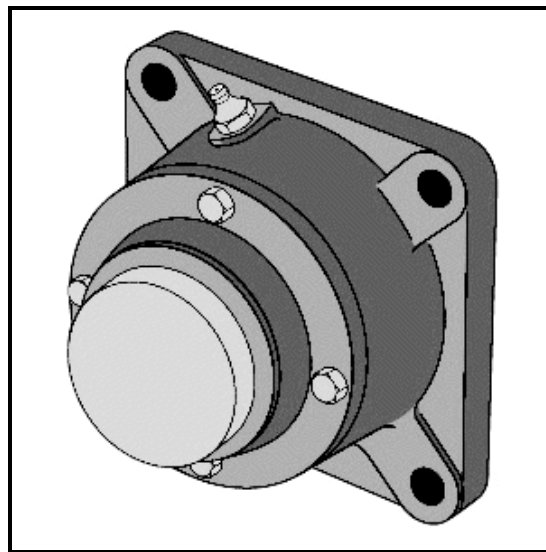


Fig. 5.11

Infine si osserva che questi cuscinetti consentono di alternare il verso di rotazione e quindi risultano idonei alle applicazioni previste.

La durata dei cuscinetti a sfere scelti è stata calcolata con una base di 20.000 ore d'esercizio. Inoltre i cuscinetti sono lubrificati con grasso che, in condizioni normali, è sufficiente per la durata di esercizio di base.

5.1.2 Cinematica [25, 26]

Lo scopo della cinematica diretta è quello di individuare la posizione dell'organo terminale nello spazio operativo, noti ed assegnati le coordinate del robot nello spazio dei giunti.

Assegnato un riferimento nello spazio operativo, sia esso $\{o, x_0, y_0, z_0\}$, conoscere le coordinate dell'organo terminale del robot significa conoscere le sue coordinate quando il robot ha assunto una determinata configurazione, mentre le coordinate nello spazio dei giunti sono dei valori che ci permettono di individuare la posizione relativa tra i vari link.

Affinché si possa passare dalle coordinate nello spazio

$$\{S(t)\} = \begin{Bmatrix} s_1(t) \\ s_1(t) \\ \bullet \\ \bullet \\ s_n(t) \end{Bmatrix} \quad (5.1)$$

operativo alle coordinate nello spazio dei giunti,

$$\{Q(t)\} = \begin{Bmatrix} q_1(t) \\ q_1(t) \\ \bullet \\ \bullet \\ q_m(t) \end{Bmatrix} \quad (5.2)$$

occorre conoscere semplicemente le relazioni geometriche che legano questi due spazi, tali relazioni le indicheremo nel seguente modo:

$$S = f(Q) \quad (5.3)$$

note le coordinate nello spazio dei giunti, il passaggio da queste alle coordinate nello spazio di lavoro prende il nome di problema cinematica diretto, il quale è caratterizzato dalla unicità della soluzione, mentre se si vogliono le coordinate nello spazio dei giunti note quelle nello spazio operativo, si ha il problema cinematica inverso e tale passaggio è caratterizzato dal fatto che la soluzione non è univocamente determinata.

Naturalmente esse sono funzioni del tempo perché a noi interessa conoscere queste relazioni durante l'intero moto del robot, per sapere istante per istante la posizione dell'organo terminale del robot.

5.1.2.1 Matrice di Denavit-Hartenberg

Una matrice di trasformazione omogenea, consente di ottenere la posizione relativa tra due corpi, potendo associare due terne di riferimento a due corpi rigidi ed utilizzandolo per la conversione delle coordinate.

Se si schematizza un robot seriale con un insieme di elementi rigidi "link", collegati tra di loro mediante giunti, si può rappresentare la posizione relativa tra due qualunque di essi, mediante una matrice di trasformazione.

A tale scopo occorre stabilire un sistema di riferimento solidale a ciascuno dei link del robot, alla base ed alla pinza eventualmente presente.

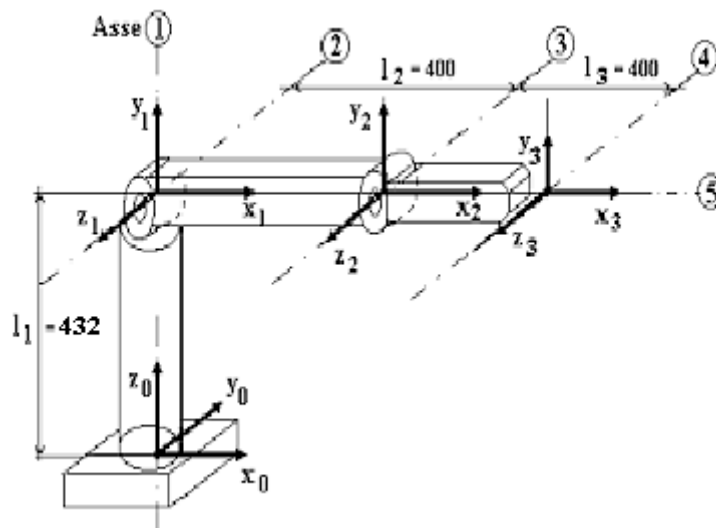


Fig. 5.12

Detta ${}^i\mathbf{T}_j$ la matrice di trasformazione omogenea tra il link i -esimo ed il link j -esimo del robot la matrice che consente di ottenere la posizione dell'organo terminale rispetto alla base fissa del robot è data da :

$${}^0\mathbf{T}_n = {}^0\mathbf{T}_1(q_1) * {}^1\mathbf{T}_2(q_2) * \dots * {}^{n-1}\mathbf{T}_n(q_n) \quad (5.4)$$

dove q_i rappresenta la variabile del giunto che indica la posizione relativa tra i link i -esimo e j -esimo; naturalmente rappresenta una lunghezza se il giunto è di tipo traslazionale, un angolo se il giunto è di tipo rotazionale.

Il modo di disporre le terne solidali a ciascun dei link rende più o meno semplice la scrittura delle matrici di trasformazioni omogenee. Denavit ed Hartenberg (D-H) hanno stabilito un criterio di posizionamento delle terne, seguendo il quale le matrici che esprimono la posizione relativa tra due link adiacenti risultano molto semplici. Le regole da seguire per posizionare una generica terna i possono così sintetizzarsi:

- ⊗ si fissa il primo asse della terna, in genere l'asse z_i , lungo l'asse del giunto $(i+1)$ -esimo, cioè lungo l'asse del moto relativo fra link i -esimo e $(i+1)$ -esimo,
- ⊗ si fissa il secondo asse, in genere l'asse x_i , normale all'asse z_i ed orientato nel verso che va dal giunto $i-1$ al giunto i . Se z_i e z_{i-1} , si intersecano, l'asse x_i dovrà essere disposto normale al piano individuato da questi due assi,
- ⊗ la terna i -esima viene completata disponendo l'asse y_i in modo tale da formare una terna levogira con i due assi precedenti.

In figura 2.1 sono posizionate le terne seguendo il criterio di D-H, inoltre il robot di figura è quello utilizzato in laboratorio per realizzare le traiettorie di cui si dirà nei capitoli successivi.

La rappresentazione di D-H consente di visualizzare facilmente le grandezze che individuano la struttura dei link ed i parametri dei giunti, essi sono :

- ⊗ a_i è la distanza tra gli assi z_{i-1} e z_i , oppure la distanza tra le origini delle terne, misurato lungo x_i ,
- ⊗ α_i è l'angolo tra gli assi z_{i-1} e z_i , misurato sul piano individuato dagli assi z_i ed x_{i-1} ,
- ⊗ θ_i è l'angolo compreso tra gli assi x_{i-1} e x_i , misurato sul piano individuato dagli assi x_{i-1} ed y_{i-1} ,

☞ d_i è la distanza tra gli assi x_{i-1} e x_i , oppure la distanza tra le origini delle terne, misurata lungo z_{i-1} .

Possiamo così ottenere i parametri di D-H per il robot in figura riportati nella tabella 5.2

Giunto	a_i	α_i	θ_i	d_i
1	0	-90°	$0^\circ \div 360^\circ$	432
2	400	0°	$-45^\circ \div 135^\circ$	0
3	400	0°	$-90^\circ \div 90^\circ$	0

Tabella 5.2

Disposte le terne secondo la convenzione di D-H, e noti i parametri dei giunti di D-H, si può calcolare la matrice omogenea di trasformazione composta che consente di rappresentare lo spostamento della terna i -esima rispetto alla terna $(i-1)$ -esima. Tale matrice è nota come matrice di D-H ed è indicata ${}^{i-1}\underline{\mathbf{A}}_i$, deve essere calcolata come prodotto di matrici fondamentali compiendo nell'ordine i seguenti "spostamenti":

- ☞ rotazione attorno all'asse z_{i-1} , di un angolo θ_i in modo da portare x_{i-1} parallelo ad x_i ,
- ☞ traslazione pari a d_i lungo z_{i-1} in modo da portare x_{i-1} a coincidere con x_i ,
- ☞ traslazione pari ad a_i lungo x_i in modo da portare l'origine della terna $i-1$ a coincidere con l'origine della terna i ,
- ☞ rotazione, pari ad α_i attorno all'asse x_i , in modo da portare gli assi y e z delle due terne a coincidere.

$${}^{i-1}\underline{\mathbf{A}}_i = \underline{\mathbf{T}}_{zd} \cdot \underline{\mathbf{T}}_{z\theta} \cdot \underline{\mathbf{T}}_{xa} \cdot \underline{\mathbf{T}}_{x\alpha} = \begin{bmatrix} C\theta_i & -C\alpha_i \cdot S\theta_i & S\alpha_i \cdot S\theta_i & a_i \cdot C\vartheta_i \\ S\theta_i & C\alpha_i \cdot C\theta_i & -S\alpha_i \cdot C\theta_i & a_i \cdot S\vartheta_i \\ 0 & S\alpha_i & C\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.5)$$

La matrice di D-H consente di ottenere la posizione relativa dei due link contigui, vedi figura 2.2. Note le coordinate di un punto del segmento i -esimo e si desidera conoscere le sue coordinate riferite alla terna solidale al basamento del robot e detto \mathbf{P}_0 il vettore che esprime tali coordinate si ha:

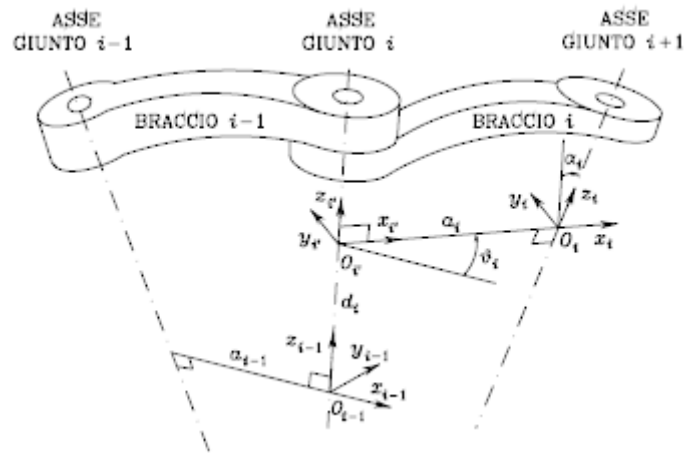


Fig. 5.12

$$\mathbf{P}_0 = {}^0\underline{\mathbf{T}}_i \cdot \mathbf{P}_i \quad (5.6)$$

con

$${}^0\underline{\mathbf{T}}_i = {}^0\underline{\mathbf{A}}_1 \cdot {}^1\underline{\mathbf{A}}_2 \cdot \dots \cdot {}^{i-1}\underline{\mathbf{A}}_i \quad (5.7)$$

per il robot di figura (5.11) la matrice è la seguente:

$$[{}^0T_3] = \begin{bmatrix} C\vartheta_1 C(\vartheta_2 + \vartheta_3) - C\vartheta_1 S(\vartheta_2 + \vartheta_3) & S\vartheta_1 & C\vartheta_1 [l_2 C\vartheta_2 + l_3 C(\vartheta_2 + \vartheta_3)] \\ S\vartheta_1 C(\vartheta_2 + \vartheta_3) - S\vartheta_1 S(\vartheta_2 + \vartheta_3) & -C\vartheta_1 & S\vartheta_1 [l_2 C\vartheta_2 + l_3 C(\vartheta_2 + \vartheta_3)] \\ S(\vartheta_1 + \vartheta_2) & C(\vartheta_2 + \vartheta_3) & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} l_1 + l_2 S\vartheta_2 + l_3 S(\vartheta_2 + \vartheta_3) \\ 1 \\ 1 \end{matrix} \quad (5.8)$$

avendo posto $i=3$.

5.2 Descrizione del sistema robot – sistema di acquisizione 3D

Di seguito saranno descritte le procedure utilizzate per acquisire la forma 3D di un oggetto, accoppiando vari sistemi di visione al robot [27]. Il sistema di visione è solidale al robot, in modo tale da sfruttare, all'interno degli algoritmi di ricostruzione tridimensionale, le informazioni relative alla posizione nello spazio di lavoro, messe a disposizione dal sistema di controllo del robot stesso. Inoltre è possibile impostare dei cicli di acquisizione con percorsi stabiliti, così come movimentare direttamente il robot, al fine di ispezionare dettagliatamente la superficie da analizzare. A valle di una specifica tecnica di calibrazione, saranno individuate tutte le informazioni necessarie per fissare il sistema di visione nello spazio di lavoro del robot. Tale procedura consente di utilizzare le informazioni di posizionamento fornite dal robot, con le loro caratteristiche di ripetibilità e di precisione, all'interno degli algoritmi di ricostruzione tridimensionale.

5.2.1 Sistema robot – fotocamera – tavolo girevole

Il sistema utilizzato è formato dal robot, una fotocamera digitale ed un tavolo girevole come mostrato in figura 5.1.



Fig. 5.1

Come si è visto nel capitolo 3 la tecnica shape from silhouette consiste nell'acquisire immagini dell'oggetto da diverse angolazioni, ottenere le varie silhouette, definire un box iniziale di punti e, grazie alla posizione relativa tra i vari piani immagine, proiettare tali punti nei sistemi di riferimento telecamera e quindi selezionare soli i punti appartenenti all'oggetto.

L'utilizzo del robot accoppiato ad un tavolo girevole è stato necessario per inquadrare l'oggetto da diverse angolazioni, ovviamente per mettere in relazione le silhouette ottenute con la reale posizione nello spazio dell'oggetto è stato opportuno definire il legame tra il sistema di riferimento telecamera, la terna base del robot ed il sistema di riferimento del tavolo girevole.

5.2.1.1 Posizionamento relativo tra i vari sistemi di riferimento

Come visto nel paragrafo precedente grazie alla matrice D-H è possibile conoscere la posizione relativa tra un punto definito nel sistema di riferimento dell'ultimo link del robot e la terna base dello stesso:

$$\{P\}_0 = [{}^0T_3] \cdot \{P\}_3 \quad (5.9)$$

inoltre è possibile stabilire la relazione tra la terna telecamera e la terna 3 del robot (figura 5.2):

$$\{P\}_{c0} = [T_{c0-3}] \cdot \{P\}_3 \quad (5.10)$$

e quindi la relazione tra le coordinate di un punto nella terna telecamera e quelle dello stesso nel riferimento base del robot:

$$\{P\}_{c0} = [T_{c0-3}] \cdot [{}^0T_3]^{-1} \cdot \{P\}_0 \quad (5.11)$$

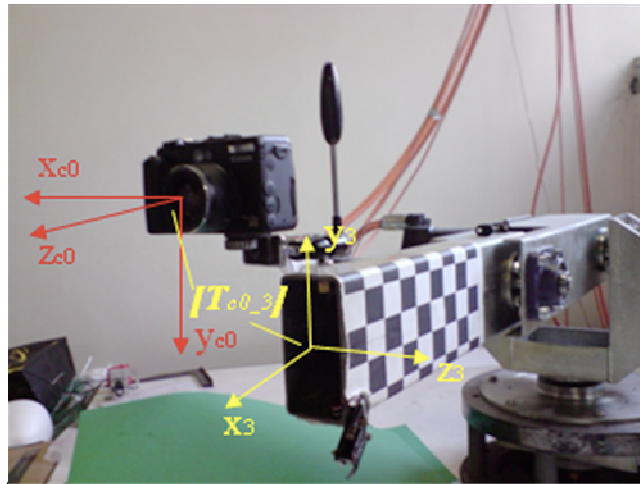


Fig. 5.2

Quando il tavolo girevole ruota, conoscendo la sua posizione nello spazio di lavoro del robot è possibile stabilire una relazione tra la terna telecamera fissa e la terna telecamera quando ruota il tavolo con sopra l'oggetto (figura 5.3):

$$\{P\}_c = [T_{c_c0}] \cdot \{P\}_{c0} \quad (5.12)$$

$$[T_{c_c0}] = f(\delta) \quad (5.13)$$

dove δ è l'angolo di rotazione del tavolo.

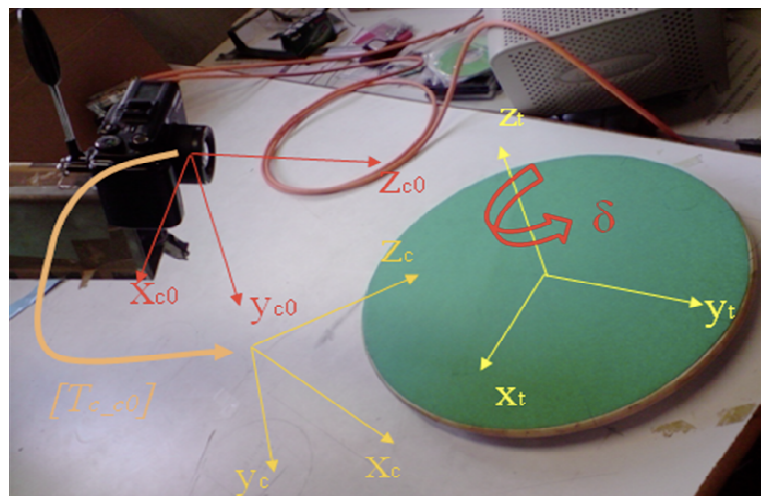


Fig. 5.3

Mediante la (5.11) si può scrivere la matrice di trasformazione omogenea:

$$[T] = [T_{c_c0}] \cdot [T_{c0_3}] \cdot [{}^0T_3]^{-1} \quad (5.14)$$

che mette in relazione il sistema di riferimento telecamera con la terna base del robot. Come visto nel capitolo 2 è possibile determinare la relazione tra il piano immagine e la terna telecamera e quindi in fine è possibile stabilire la relazione tra un punto nel riferimento base del robot e lo stesso nel piano immagine:

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} = \frac{1}{d} \cdot [K] \cdot [T] \cdot \begin{Bmatrix} x^0 \\ y^0 \\ z^0 \\ 1 \end{Bmatrix} = [M] \cdot \{P\}_0 \quad (5.15)$$

dove $[K]$ è la matrice dei parametri intrinseci della telecamera.

5.2.1.2 Calibrazione del sistema

La calibrazione del sistema consiste nel determinare le grandezze che regolano la relazione matriciale (5.15). I parametri strutturali del robot che insieme alle rotazioni dei giunti definiscono la matrice D-H sono stati ottenuti mediante una procedura di calibrazione cinematica del robot, la posizione del tavolo girevole è stata ottenuta posizionando l'end effector del robot al centro dello stesso e nell'ipotesi che l'asse di rotazione sia diretto come l'asse del primo link del robot; la matrice $[T_{c_c0}]$ è ottenuta traslando la terna base, ruotandola dell'angolo δ . Per quanto riguarda invece la posizione relativa tra la terna telecamera e quella dell'ultimo link del robot è stato opportuno stabilire una procedura di calibrazione di seguito descritta:

Utilizzando una griglia nota e inquadrandola con la telecamera da diverse angolazioni si ottengono i parametri intrinseci ed estrinseci così come mostrato nel capitolo 2 quando si è descritta la tecnica di calibrazione della telecamera. I parametri estrinseci descrivono il posizionamento della telecamera rispetto alla griglia, figura 5.4.

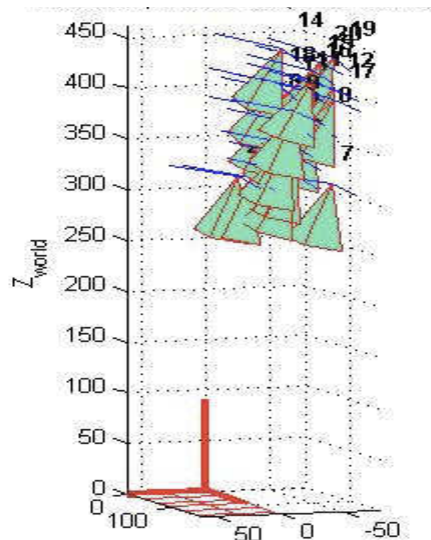


Fig. 5.4

Visto che la telecamera si sposta nelle varie configurazioni con una legge nota definita dalla cinematica del robot e nota la posizione della griglia nel riferimento base del robot è possibile stabilire la posizione della terna telecamera rispetto all'ultimo link, figura 5.5.

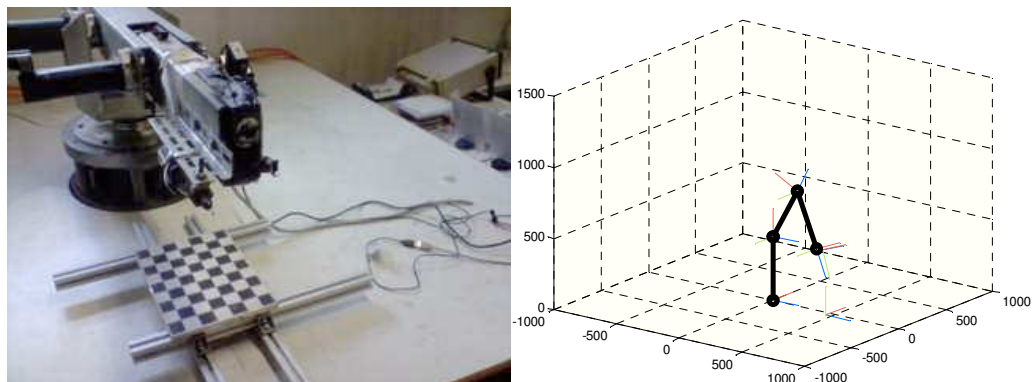


Fig. 5.5

5.2.1.3 Descrizione della procedura di ricostruzione tridimensionale

Il primo step da fare consiste nell'acquistare le immagini dell'oggetto da differenti angolazioni usando la movimentazione del robot e del tavolo girevole come mostrato in figura 5.6 e quindi conoscendo per l'immagine i -esima la corrispondente matrice di trasformazione $[M]$.

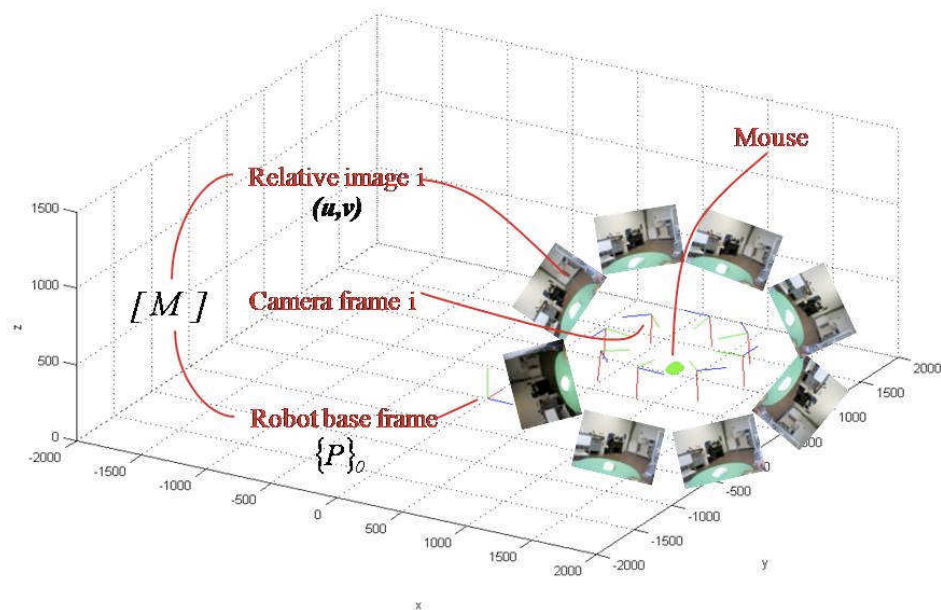


Fig. 5.6

E' importante notare che la matrice $[M]$ una volta effettuata la procedura di calibrazione dipenderà soltanto dalla rotazione del tavolo girevole e dalle variabili di giunto del robot, queste ultime sono ottenute acquisendo gli encoder degli attuatori e quindi l'errore dovuto alle rotazioni dei link è praticamente nullo.

Mediante la procedura riportata nel capitolo 3 si ottengono le silhouette per ogni immagine e intersecando i raggi ottici passanti per i baricentri dei profili si viene a stabilire la posizione del box iniziale di dimensioni dipendenti dalle dimensioni delle silhouette.

Una volta stabilito il posizionamento del box iniziale e la risoluzione si procede alla ricostruzione tridimensionale proiettando i punti nei vari piani immagine, figura 5.7.

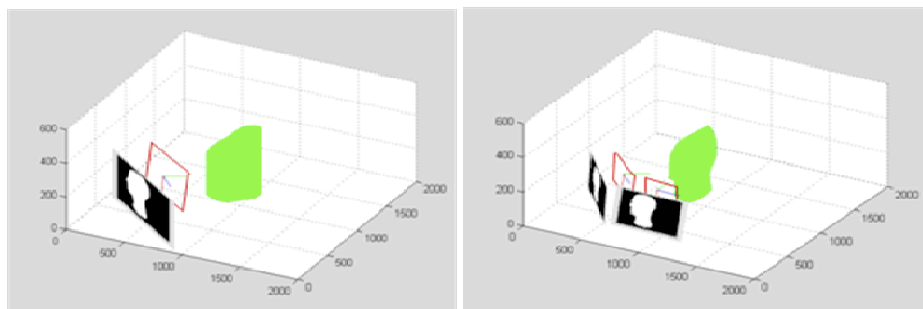


Fig. 5.7

Alla fine della procedura si ottiene la forma tridimensionale nel sistema di riferimento del robot come mostrato nella figura 5.8.

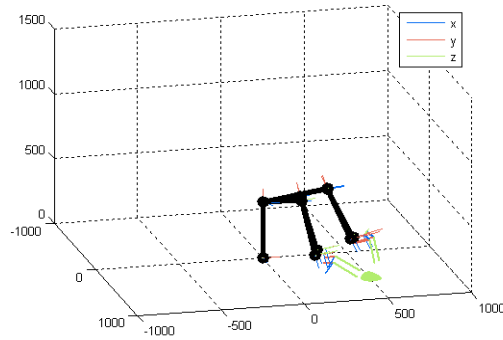


Fig. 5.8

5.2.1.4 Risultati

La procedura è stata validata su di un mouse da computer, sono state scelte varie risoluzioni in modo tale da valutare l'accuratezza e la velocità di ricostruzione.

In figura 5.9 sono mostrate le nuvole di punti per differenti risoluzioni: a) res = 50, b) res = 80, c) res = 100, d) res = 150.

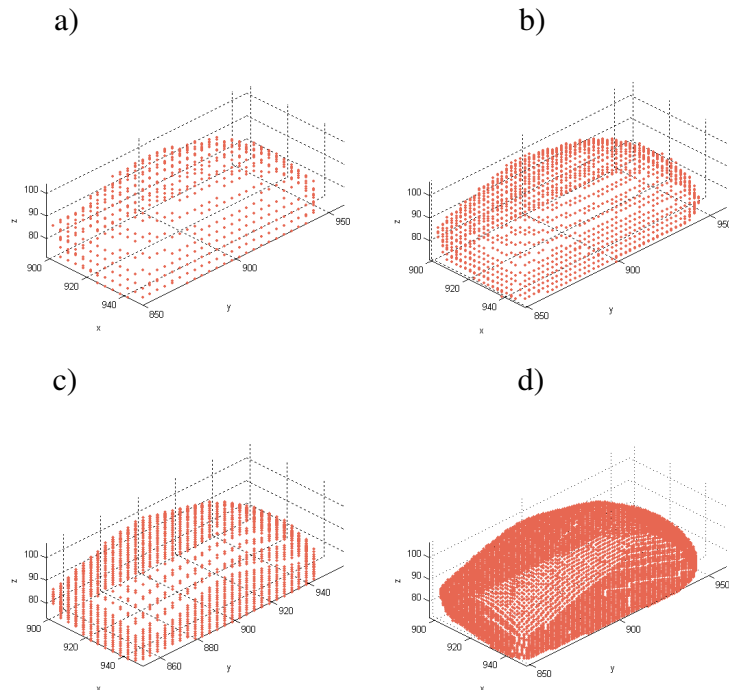


Fig. 5.9

Per una risoluzione $res = 150$ si hanno le seguenti accuratèzze: 1.33 mm lungo la direzione x, 1.33 mm lungo y, 0.66 mm lungo z.

La ricostruzione del mouse con 8 foto e 150^3 punti iniziali, è ottenuta in circa 6 minuti.

5.2.2 Sistema robot – laser scanner

L'altro sistema di acquisizione 3D implementato è costituito da un sistema laser scanner, formato da una web-cam ed un emettitore laser, fissato sull'ultimo link del robot, figura 5.10.



Fig. 5.10

Quando il laser scanner è installato sul robot è possibile usare, come visto in precedenza, le informazioni di posizione dal robot in modo da determinare la posizione e l'orientamento della finestra di scansione nel riferimento base del robot. A tal fine è importante conoscere la relazione tra la terna telecamera e la terna di riferimento dell'ultimo link (figura 5.11); successivamente mediante la matrice $[DH]$ è possibile determinare le coordinate di punti ottenuti mediante la triangolazione tra la telecamera e il modulo laser nel sistema di riferimento base del robot.

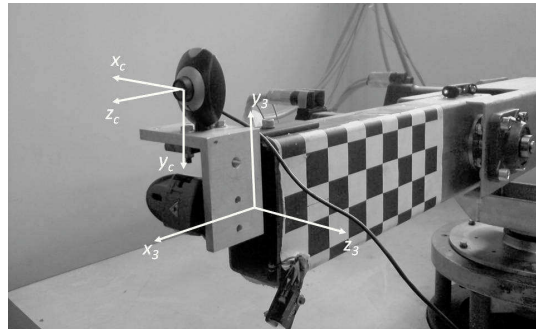


Fig. 5.11

5.2.2.1 Calibrazione del sistema

La calibrazione del sistema robot-laser scanner consiste nel determinare le posizioni relative tra modulo laser e telecamera e tra telecamera e robot. La prima relazione è ottenuta mediante la procedura di calibrazione vista nel capitolo 4 e che può essere implementata mantenendo la posizione del robot fissa, figura 5.12.

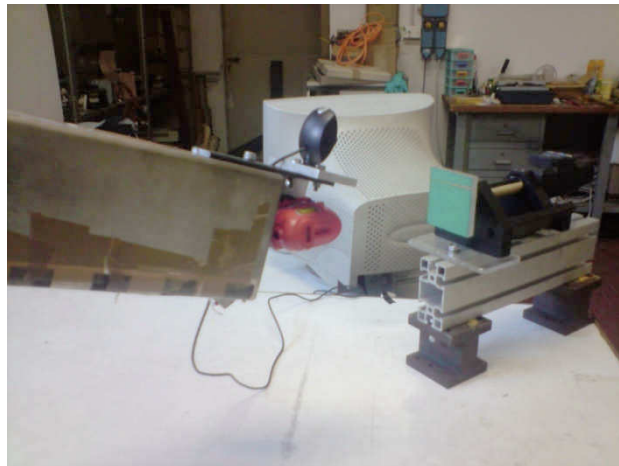


Fig. 5.12

La posizione relativa tra la telecamera e la terna dell'ultimo link del robot è ottenuta mediante la procedura di calibrazione descritta nel paragrafo 5.2.1.2. In fine si stabilisce una relazione tra le coordinate (u,v) della forma laser nel piano immagine alle sue coordinate nella terna base del robot. Grazie a queste relazioni è possibile ricostruire le superfici di interesse in real time acquisendo la posizione angolare dei link dagli encoder e estrapolando la linea laser dal resto dell'immagine.

5.2.2.2 Descrizione della procedura di ricostruzione tridimensionale

Il sistema laser scanner cattura i profili della superficie dell'oggetto, durante una procedura di scansione, il laser scanner viene mosso dal robot in modo da registrare immagini del profilo dell'oggetto da differenti posizioni ed orientamenti in base alla forma dell'oggetto.

In ambiente Matlab è stata sviluppata una GUI che permette all'utente di utilizzare il sistema in modo semplice, figura 5.13.

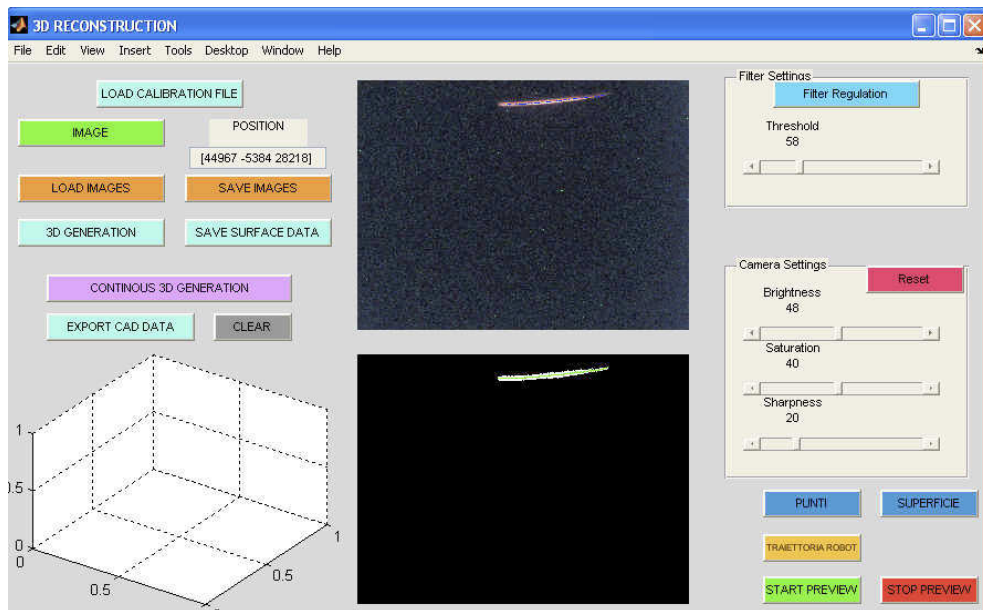


Fig. 5.13

Il primo step da fare consiste nel caricare i dati provenienti dalla calibrazione, fatto ciò è possibile regolare alcune proprietà della telecamera come la luminosità, la saturazione, ecc. ed i parametri del filtro software utile ad estrapolare la forma del laser dal resto dell'immagine. Nella finestra vengono visualizzate l'immagine originale e quella elaborata del profilo della linea laser e le variabili di giunto del robot; fatte le opportune regolazioni cliccando sul pulsante "3D generation" oppure "Continuous 3D generation" è possibile rispettivamente acquisire singoli punti 3D oppure pianificare una traiettoria di scansione e quindi acquisire in continuo la superficie 3D di interesse.

Finita l'operazione di scansione è possibile salvare i dati in vari formati, è previsto l'esportazione della nuvola di punti acquisita in un formato standard utilizzato dai software CAD più comunemente usati.

Per migliorare l'accuratezza della ricostruzione è stato utile sostituire il modulo laser iniziale con un modulo laser che emette luce su una prefissata lunghezza d'onda pari a 635 nm ed inoltre si è adoperato un filtro ottico passa banda centrato sulla lunghezza d'onda del modulo laser in modo tale da selezionare soltanto il fascio laser e rendere l'operazione di filtraggio software più semplice e robusta (vedi figura 5.14).

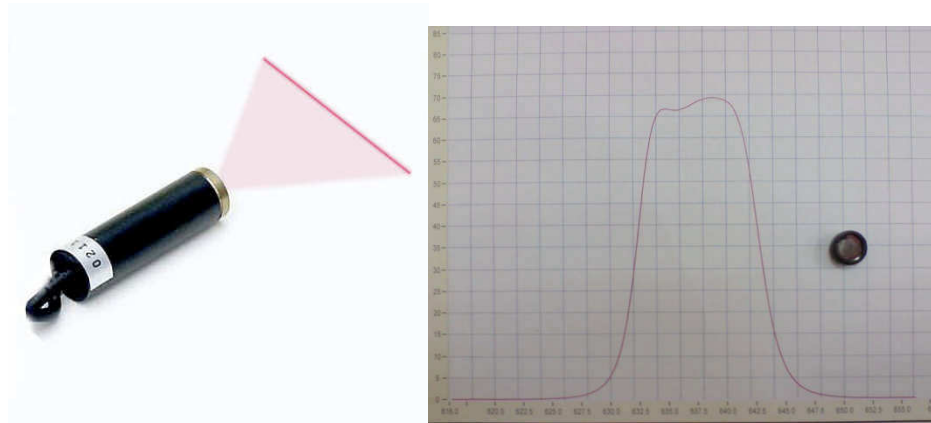


Fig. 5.14

In figura 5.15 è mostrato il dispositivo nella sua configurazione finale.



Fig. 5.15

5.2.2.3 Risultati

Il sistema è stato provato prima in una posizione fissa del robot, per verificare la taratura e le procedure di ricostruzione, successivamente è stato utilizzato considerando anche la movimentazione del robot. Gli oggetti in prova sono indicati nella figura 5.16.

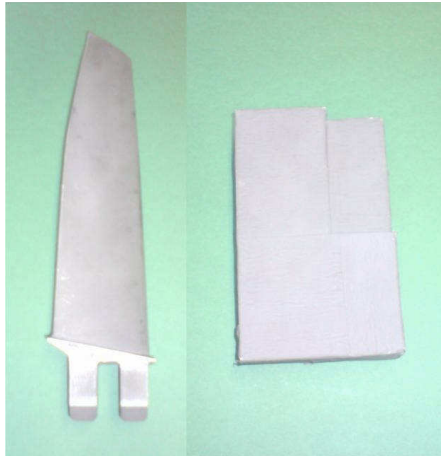


Fig. 5.16

Nelle figure 5.17 e 5.18 è possibile vedere un passo della procedura con i due risultati dei campioni in prova. Con l'uso del software CATIA è stato possibile costruire la superficie dei due oggetti, questo passo del metodo di ricostruzione 3D può essere visto come una vera e propria procedura di Reverse Engineering. La routine "Digitized shape editor" di "CATIA" permette di rielaborare la nuvola di punti in modo tale da eliminare delle imperfezioni che potrebbero rendere difficile l'operazione di ricostruzione della superficie.

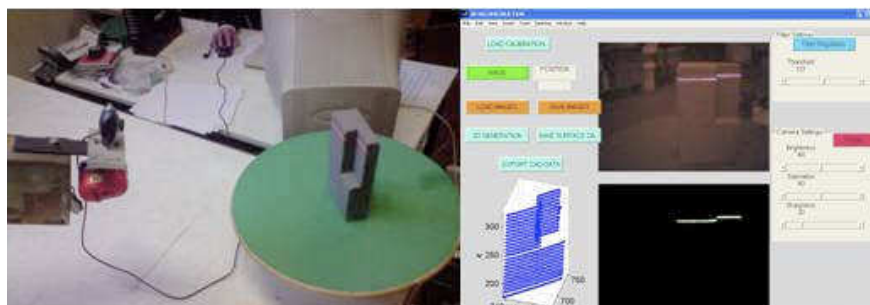


Fig. 5.17

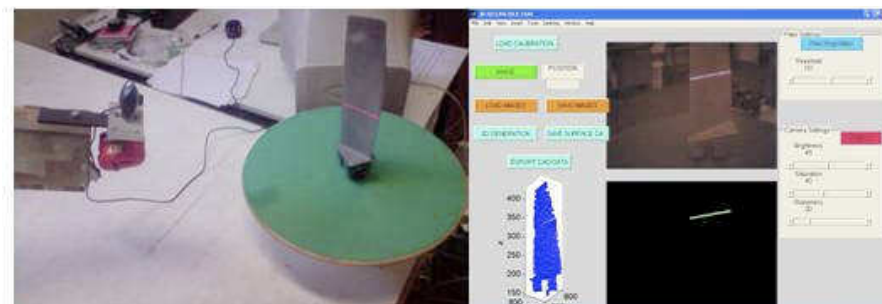


Fig. 5.18

Nelle figure 5.19 e 5.20 sono riportati i raffronti tra le nuvole di punti e le rispettive superfici per ciascun oggetto.

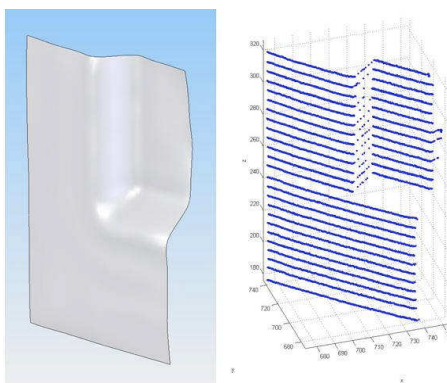


Fig. 5.19

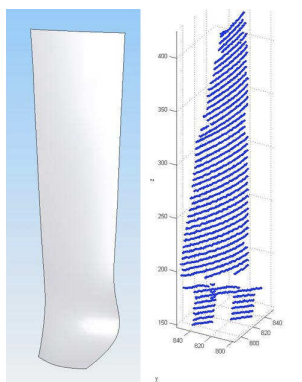


Fig. 5.20

Nelle figure 5.21 e 5.22, è indicata una valutazione della precisione di ricostruzione 3D per i due oggetti in prova. E' possibile osservare, che l'accuratezza peggiore si ha

lungo la direzione z dell'asse della fotocamera, questo risultato è in linea con le osservazioni riportate nel capitolo 4.

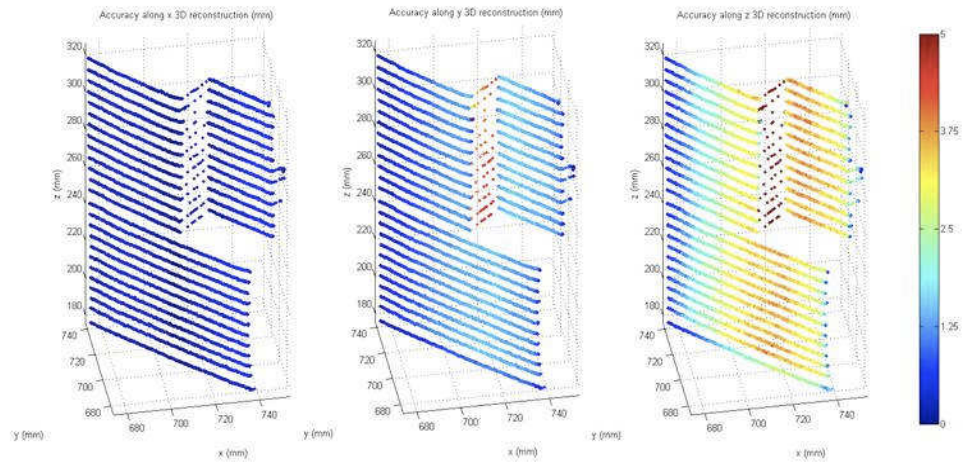


Fig. 5.21

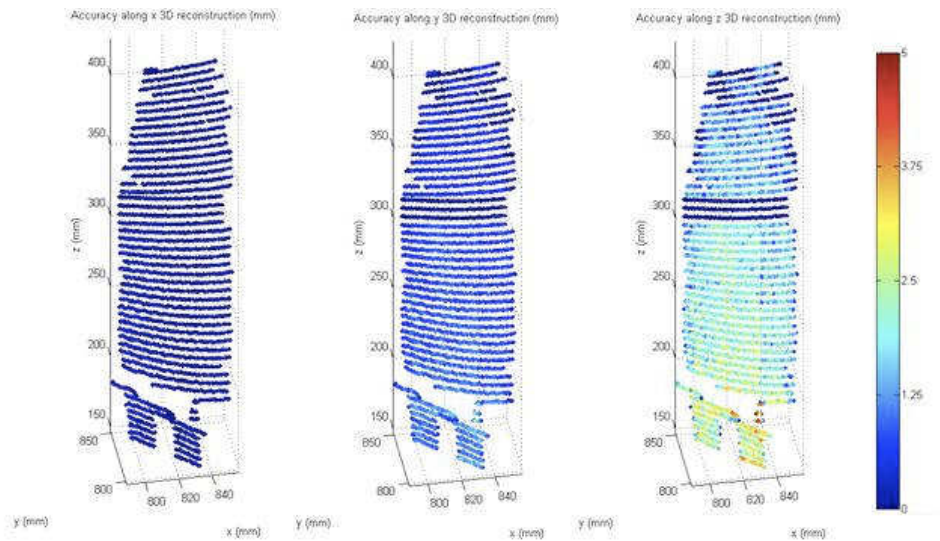


Fig. 5.22

Per validare l'accuratezza del prototipo robot-laser-scanner sviluppato si è deciso di comparare la ricostruzione 3D di un oggetto con quella ottenuta mediante uno scanner in commercio della Konica Minolta, figura 5.23.



Fig. 5.23

L'oggetto scelto è una testa d un manichino, figura 5.24.



Fig. 5.24

Nelle figure 5.25 e 5.26 sono riportate rispettivamente le nuvole di punti ottenute con il nostro sistema e con lo scanner 3D commerciale.

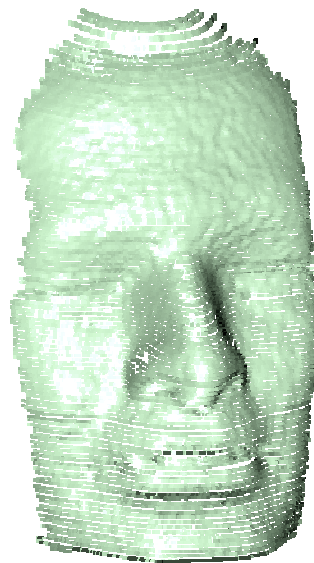


Fig. 5.25

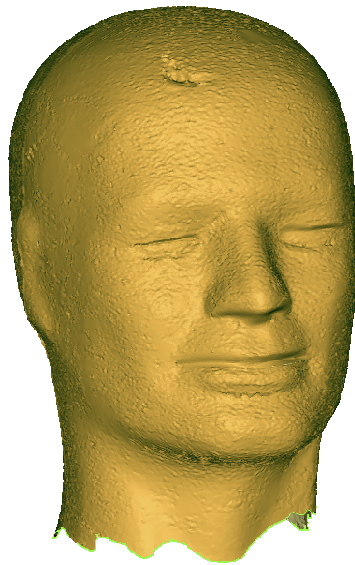


Fig. 5.26

In figura 5.27 è riportata la comparazione tra le due nuvole di punti; è interessante notare che la deviazione 3D standard media è minore di ± 1.5 mm.

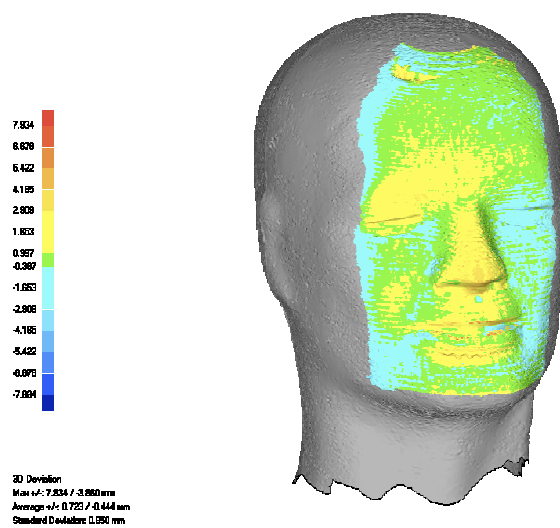


Fig. 5.27

Da un'analisi più approfondita è emerso che tali differenze riguardano singoli punti, quindi è possibile ipotizzare che una elaborazione preliminare della nuvola di punti potrebbe permettere un aumento generale della precisione di ricostruzione.

5.3 Confronto fra le due tecniche

In fine è utile fare un confronto tra le due tecniche, in particolare si ha che la Shape from silhouette non necessita di un ambiente a luce strutturata, necessita della calibrazione della telecamera ha dei tempi di elaborazione lunghi e in post-processing riuscendo e l'accuratezza è legata al numero di punti definiti nel box iniziale.

La triangolazione laser_invece necessita di una luce strutturata, c'è bisogno della calibrazione della telecamera e del piano laser, è possibile la ricostruzione real-time e l'accuratezza è legata principalmente all'angolo di incisione tra il fascio laser e la superficie, all'elaborazione del profilo laser e alla precisione della posizione tra il piano laser e la telecamera. Per la nostra applicazione che, come detto nell'introduzione, consiste nel costruire una macchina che replica superfici, si è scelto di adottare la tecnica laser scanner vista la velocità di esecuzione della ricostruzione 3D a parità di accuratezza richiesta.

CAPITOLO 6

Utilizzo di un robot antropomorfo per la realizzazione di superfici digitalizzate

6.1 Descrizione della procedura

In questa fase è stata sviluppata una procedura tale per cui si realizza la superficie scannerizzata a partire da un blocco di materiale grezzo. L'obiettivo preposto sostanzialmente è quello di utilizzare lo stesso robot adottato per la scannerizzazione, come è stato mostrato nel capitolo precedente, per realizzare mediante lavorazione meccanica le superfici digitalizzate.

Per fare ciò è stata sviluppata quella che è nota come programmazione CAM che è l'acronimo per l'espressione inglese "Computer-Aided Manufacturing", che significa fabbricazione assistita da computer.

Tale espressione indica una categoria di prodotti software che analizzano un modello geometrico bidimensionale o tridimensionale, e generano le istruzioni per una macchina utensile atte a produrre un manufatto avente la forma specificata nel modello.

Un programma di CAM è un programma interattivo, che permette all'utente di eseguire le seguenti operazioni:

- Caricamento di un modello geometrico da un file in un formato standard, normalmente generato da un programma di CAD.
- Impostazione del sistema di coordinate usato dalla macchina utensile
- Selezione della parte di modello da lavorare.
- Impostazione dei parametri di lavorazione (raggio dell'utensile, velocità di spostamento) per ogni porzione del modello.
- Generazione delle istruzioni per la macchina utensile.
- Visualizzazione o modifica della sequenza di istruzioni generate.
- Visualizzazione grafica del percorso utensile generato.
- Invio dei dati al controllo numerico della macchina utensile.

La funzione più complessa e più onerosa dal punto di vista computazionale è la generazione delle istruzioni, in quanto per determinare il percorso utensile ottimale sono necessari complessi calcoli geometrici.

Esistono diversi tipi di lavorazioni che in generale vengono sviluppati con un processo CAM, i principali tipi sono i seguenti:

- Incisione.
- Taglio.
- Svuotamento di percorsi chiusi.
- Svuotamento di bitmap.
- Modellazione di superfici.

Incisione

L'incisione è una lavorazione in cui l'utensile penetra solo di pochi decimi di millimetro nel pezzo, seguendo un percorso definito da linee. Ne è un tipico esempio l'incisione di targhe o di medaglie. Il modello geometrico è costituito da spezzate aperte o chiuse, e forature. Una spezzata è una sequenza di linee (tipicamente segmenti di retta, archi, b-spline, o bezier) in cui ogni linea inizia dove la linea precedente finisce. Le forature sono da ritenersi spezzate particolarmente semplici, in quanto hanno un solo punto.

Per effettuare la lavorazione, l'utensile deve, per ogni spezzata, posizionarsi rapidamente sopra il primo punto, scendere nel pezzo, seguire la spezzata, risalire dal pezzo. Alcune varianti sono la compensazione (o offset) e i vincoli sull'angolo di incidenza. La compensazione consiste nel fatto che, dato che l'utensile ha un raggio non nullo, se si vuole che l'utensile passi rasente alla spezzata, il reale percorso utensile deve avere uno scostamento rispetto al percorso originale pari al raggio dell'utensile. Alcuni CNC hanno internamente la capacità di applicare la compensazione del raggio utensile, altri no. Per questi ultimi è necessario che la compensazione sia calcolata dal software di CAM. I vincoli sull'angolo di incidenza si hanno quando la superficie da lavorare non è a quota uniforme. Per fare in modo che l'utensile sia sempre ortogonale alla superficie da incidere, o comunque che l'angolo di incidenza non superi una soglia definita, si deve applicare una rotazione

all'utensile oppure al pezzo da lavorare. Questo è possibile solo con macchine aventi più di tre gradi di libertà.

Taglio

Il taglio è una lavorazione in cui l'utensile penetra nel pezzo per una profondità superiore allo spessore del pezzo, seguendo un percorso definito da linee. Ne è un tipico esempio il taglio di lenti o mascherine. Dal punto di vista del software CAM, incisione e taglio sono molto simili. Tuttavia, per il taglio si possono avere due problemi:

Se una figura chiusa contiene un'isola (per esempio la lettera "a"), si deve tagliare prima la parte interna, e poi la parte esterna. Questo perché una figura chiusa, dopo che è stata tagliata, non è più solidale con la macchina, e quindi non può più essere lavorata.

Quando il taglio di una figura chiusa è quasi completato, la figura è connessa al resto del pezzo per una porzione infinitesima. In tale frangente, le vibrazioni della lavorazione possono produrre delle imperfezioni.

Svuotamento di percorsi chiusi

Lo svuotamento di percorsi chiusi consiste nel raggiungere con l'utensile tutti i punti interni a uno o più percorsi chiusi. Ne è un tipico esempio la produzione di timbri da disegni eseguiti con programmi di CAD. Il problema è lo stesso che ha un giardiniere che deve tagliare l'erba di un prato recintato. Il diametro dell'utensile corrisponde alla larghezza del tagliaerba.

Svuotamento di bitmap

Lo svuotamento di bitmap consiste nel raggiungere con l'utensile tutti i punti che in una matrice rettangolare di bit (detta "bitmap") hanno valore uno (oppure zero). Ne è un tipico esempio la produzione di timbri da disegni acquisiti tramite scanner. Il problema è analogo a quello dello svuotamento di percorsi chiusi, con la differenza che in questo caso la superficie da svuotare è definita da una matrice di bit invece che da entità vettoriali.

Modellazione di superfici

La modellazione di superfici consiste nello scolpire forme arbitrarie definite come superfici parametriche. Ne è un tipico esempio la produzione di stampi. Una superficie parametrica è una funzione reale di due variabili reali definite in intervalli. Una possibile rappresentazione in un computer di una superficie parametrica è una matrice rettangolare di numeri. Il numero alla colonna x della riga y rappresenta la quota che la superficie ha al punto di coordinate (x,y) . Tale istogramma bidimensionale viene comunemente chiamato DEM (Digital Elevation Model) o modello digitale di elevazione. Partendo da un DEM, e da un pezzo avente tipicamente la forma di un parallelepipedo, l'utensile deve scolpire la superficie desiderata. A tal fine, la lavorazione è divisa in due fasi: sgrossatura e rifinitura. Nella sgrossatura, si effettua un terrazzamento a più livelli con un utensile di grosso diametro, ottenendo un pezzo che somiglia a quello desiderato, ma è scalettato.

Nel nostro caso siamo interessati alla modellazione di superfici, i punti che abbiamo a disposizione provengono dalla ricostruzione 3D mediante il sistema robot-laser scanner, tali punti devono essere prima elaborati per eliminare eventuali imperfezioni di ricostruzione, successivamente viene generata la traiettoria da far eseguire al robot ed in fine saranno assegnate allo stesso le istruzioni per la lavorazione [28]. Le informazioni registrate con il sistema di acquisizione sono utilizzate, per pianificare una traiettoria da assegnare al robot, affinché, tramite un utensile installato su di esso, sia possibile riprodurre la superficie acquisita lavorandola dal pieno, oppure ricavando lo stampo necessario ad ottenerla. I dati possono essere utilizzati anche per riprodurre la superficie con un fattore di scala.

6.2 Pianificazione della traiettoria

Di seguito sono descritti i vari passi per ottenere una traiettoria da assegnare al robot a partire dai punti ricostruiti. Come detto in precedenza la prima operazione da fare una volta ottenuta la nuvola di punti mediante la scansione laser è quella di eliminare eventuali imperfezioni. Anche in questo caso per l'intera procedura di ottenimento della traiettoria è stata sviluppata una GUI indicata nella figura 6.1.

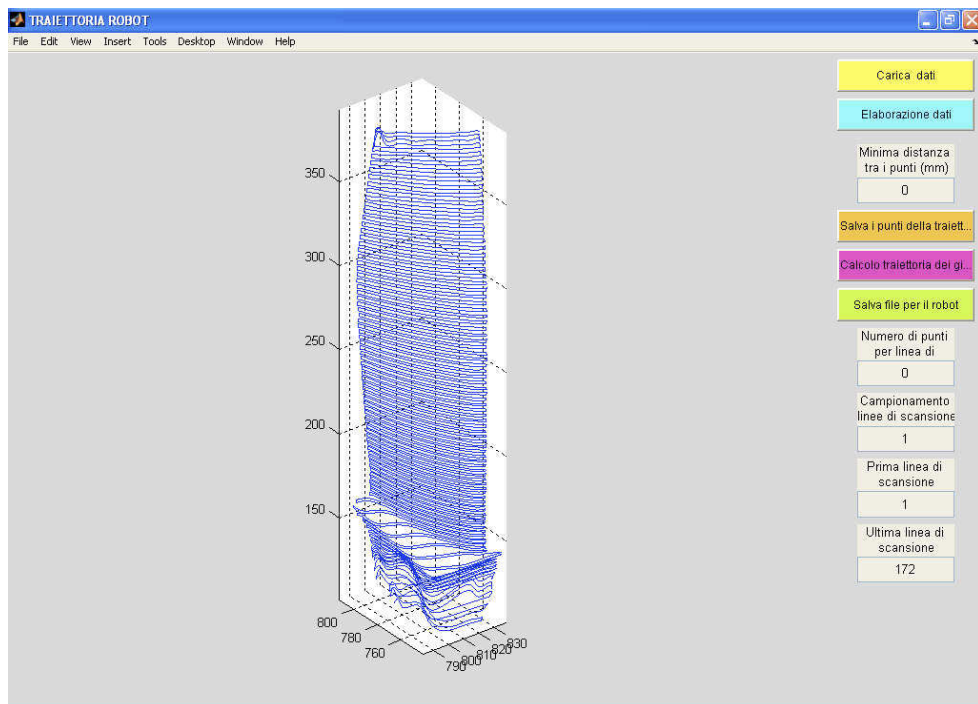


Fig. 6.1

Si prendano per esempio in considerazione i punti indicati nella figura 6.2, dopo l'elaborazione che permette di sotto campionare i punti acquisiti e di eliminare punti dovuti ad errori di ricostruzione si ottiene una nuova nuvola di punti riportata in figura 6.3.

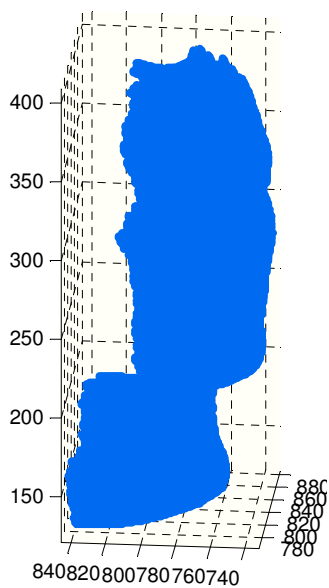


Fig. 6.2

A questo punto è possibile selezionare la parte della superficie che si vuole realizzare; per esempio nella figura 6.4 è riportata una porzione di faccia.

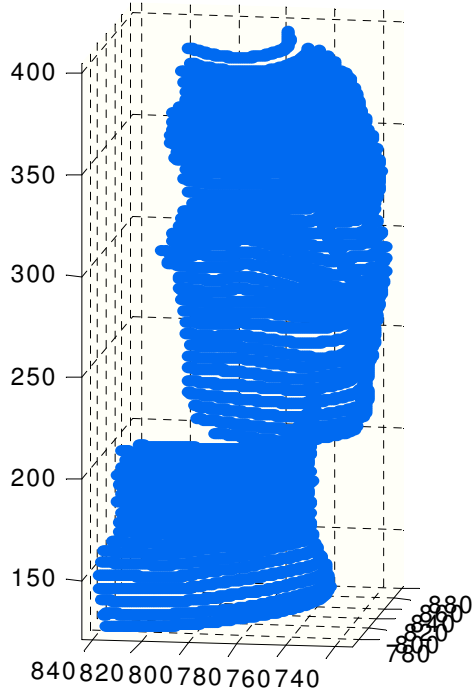


Fig. 6.3

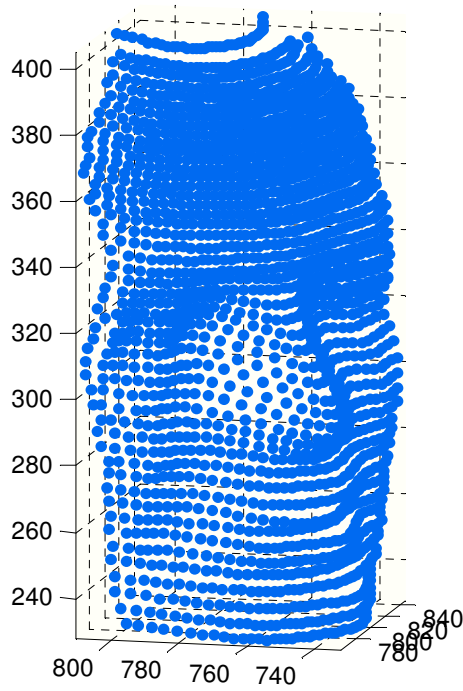


Fig. 6.4

Fatto ciò con un altro programma è possibile spostare ed orientare i punti ottenuti nello spazio di lavoro del robot ed eventualmente scalarli come descritto nella figura 6.5.

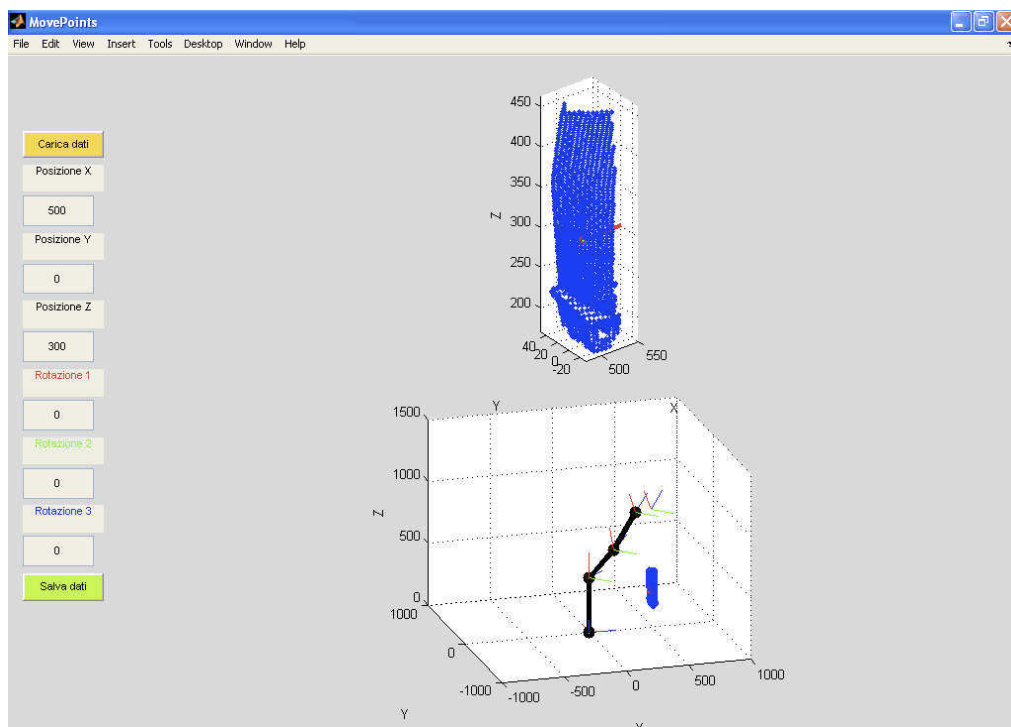


Fig. 6.5

A partire dalla forma del materiale da lavorare e dalla sua posizione nello spazio di lavoro del robot è necessario implementare un ciclo di lavorazione.

L'algoritmo di pianificazione della traiettoria prevede la realizzazione della superficie per mezzo di passate successive. La traiettoria è pianificata assegnando il reale percorso geometrico che l'utensile deve seguire, in termini di posizione, di velocità ed accelerazione.

Si prenda per esempio in considerazione un pezzo iniziale avente la forma di un parallelepipedo, e si fissi su di esso un sistema di riferimento $\{O_b, x_b, y_b, z_b\}$ così come mostrato in fig. 6.6.

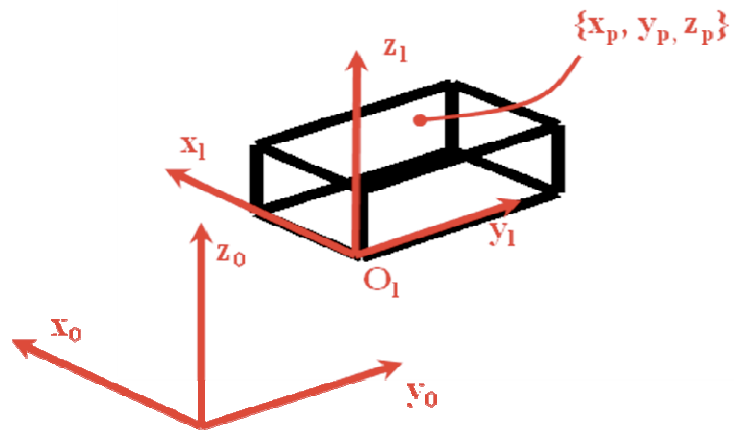


Fig. 6.6

Siano $\Theta = \{x_p, y_p, z_p\}$ i punti appartenenti al volume racchiuso dal parallelepipedo, per tali punti deve essere verificata la condizione tale per cui questi siano interni allo spazio di lavoro del robot.

Sia $\Gamma = \{x, y, z\}$ l'insieme costituito dalle coordinate dei punti della superficie da realizzare, si orienti Γ in modo tale che risulti:

$$\begin{cases} \max(x) - \min(x) \leq a_x \\ \max(y) - \min(y) \leq a_y \\ \max(z) - \min(z) \leq a_z \\ \max(z) = \max(z_p) \end{cases}$$

Dove a_x , a_y e a_z sono le dimensioni del pezzo da lavorare rispettivamente lungo gli assi x, y, z .

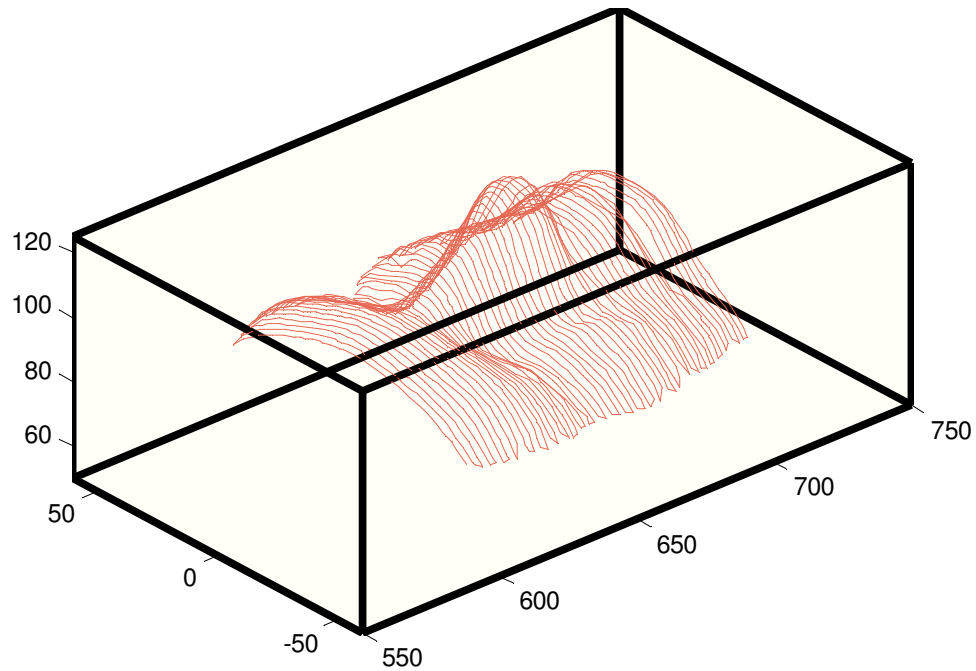


Fig. 6.7

Si definisco le curve di livello ottenute da Γ ed appartenenti a piani paralleli al piano x_1z_1 ed equidistanziati di una quantità δ_y dipendente dalla geometria dell'utensile utilizzato. Indicando con n_y il numero di curve si ha:

$$\gamma_i = \{ x_i, \bar{y}_i, z_i \} \quad i = 1, \dots, n_y$$

Si definisce la grandezza $\delta_z = \delta_y$ ed il numero n_z pari all'approssimazione per eccesso ad un numero intero del seguente rapporto:

$$\frac{\max(z_i) - \min(z_i)}{\delta_z}$$

allora, per ogni curva di livello i punti che definiscono il percorso da assegnare all'utensile sono definiti come segue:

$$\gamma_i^t = \{ x_i^t, \bar{y}_i^t, z_i^t \} = \begin{cases} \{ x_i, \bar{y}_i, z_i \}_{j=1, \dots, n_z} & \text{se } \max(z_i) - \delta_z j \leq z_i \\ \{ x_i, \bar{y}_i, \max(z_i) - \delta_z j \}_{j=1, \dots, n_z} & \text{se } \max(z_i) - \delta_z j \geq z_i \end{cases}$$

Tali punti saranno ordinati dal valore massimo al valore minimo di z_i^t ed alternando un ordinamento crescente – decrescente alle x_i^t al variare di j , come mostrato in figura 6.8.

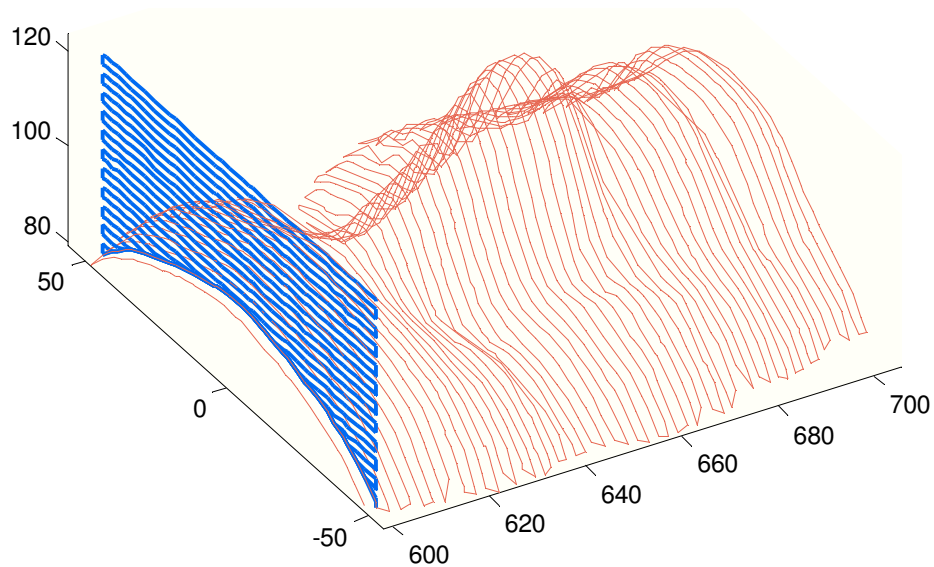


Fig. 6.8

Agli insiemi dei punti delle curve di livello viene aggiunto un punto avente coordinate x_l e y_l pari a quelle dell'ultimo punto e $z_l = \max(z_p) + 5 \delta_z$ in modo tale da assegnare all'utensile un punto che lo riporti una volta eseguita la curva di livello al di sopra del pezzo da lavorare e quindi in posizione tale da poter passare alle curve successive.

Il percorso totale da assegnare all'utensile per effettuare la lavorazione è rappresentato dall'unione delle curve di livello come indicato di seguito:

$$\Gamma^t = \bigcup_{i=1}^{n_y} \gamma_i^t$$

In figura 6.9 è raffigurato il percorso totale che deve seguire l'utensile per realizzare la superficie desiderata rappresentata in figura con il colore rosso.

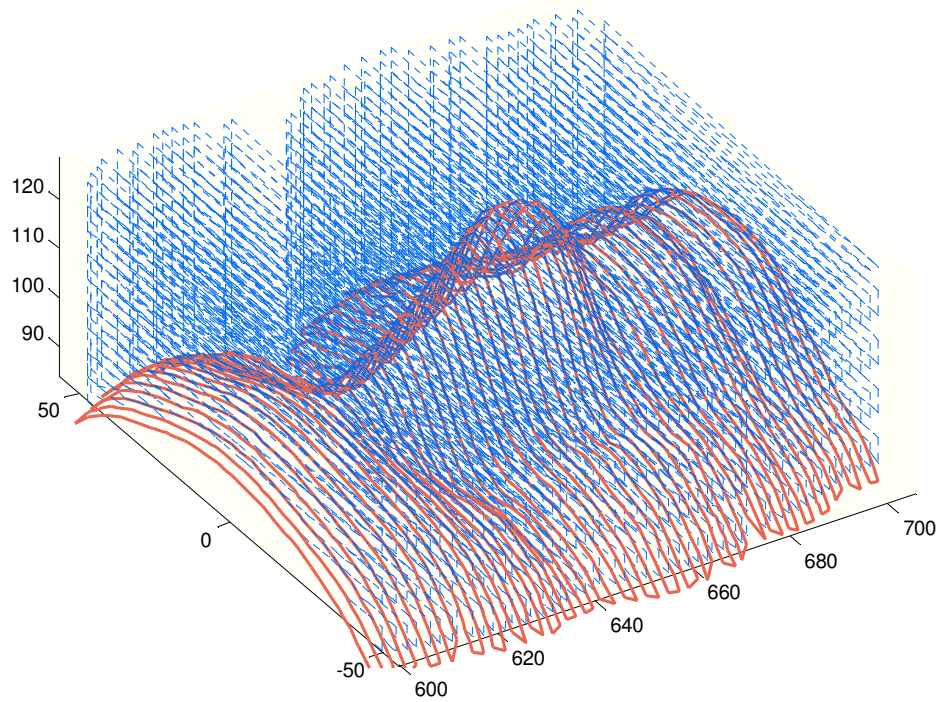


Fig. 6.9

L'obiettivo successo è stato quello di definire una curva parametrica che interpolasse i punti del percorso.

Una curva parametrica nello spazio è un'applicazione continua da un intervallo $I \subseteq \mathbb{R}$ in \mathbb{R}^3 :

$$\lambda \in I \rightarrow \{x(\lambda), y(\lambda), z(\lambda)\} \in \mathbb{R}^3$$

Per definire le varie grandezze associate alla curva si richiede per lo meno la regolarità di classe C^2 (vale a dire di ciascuna componente come funzione da I in \mathbb{R}).

Una curva inoltre è regolare in un punto se il vettore:

$$\{x'(\lambda), y'(\lambda), z'(\lambda)\} \tag{6.1}$$

non è il vettore nullo. Com'è noto, la (6.1) fornisce in ogni punto della curva la direzione tangente alla curva stessa (con l'apice si intende la derivazione rispetto

all'argomento). La mancanza di regolarità in un punto della curva è associata all'esistenza di punti singolari.

Per ottenere una curva parametrica a partire da Γ^t si è scelto di adottare la funzione interpolante spline cubica naturale. Si indichi con N_t il numero di punti del percorso, ed il parametro $k = 1, 2, \dots, N_t$, per ogni vettore componente di Γ^t si definisce una funzione interpolante.

A titolo di esempio si consideri il vettore colonna $x'_k = f(k)$, la spline cubica di tale funzione nei nodi x'_k sarà una funzione $s_x(\lambda)$ tale che:

$$s_x(\lambda) = \begin{cases} s_{x1}(\lambda) & \text{se } 1 \leq \lambda \leq 2 \\ s_{x2}(\lambda) & \text{se } 2 \leq \lambda \leq 3 \\ s_{x3}(\lambda) & \text{se } 3 \leq \lambda \leq 4 \\ \dots & \dots \\ s_{xN_t}(\lambda) & \text{se } N_t - 1 \leq \lambda \leq N_t \end{cases}$$

Dove ogni $s_{xk}(\lambda)$ è un polinomio di grado ≤ 3 , e tale inoltre che

$$(I) \quad s_{xk}(\lambda) = x'_k \quad \text{per } k = 1, 2, \dots, N_t$$

$$(II) \quad s_x(\lambda) \in C^2[1, N_t]$$

Così facendo si ottiene una curva parametrica nello spazio $\Psi(\lambda)$:

$$\Psi(\lambda) = \{s_x(\lambda), s_y(\lambda), s_z(\lambda)\} \quad (6.2)$$

Per ogni punto della funzione (6.2) è possibile definire il versore tangente:

$$t(\lambda) = \{t_x(\lambda), t_y(\lambda), t_z(\lambda)\} = \frac{\Psi'(\lambda)}{|\Psi'(\lambda)|}$$

Una volta definito il percorso geometrico si passa alla definizione della traiettoria.

Con la pianificazione della traiettoria si intende stabilire la modalità con cui si vuole che evolva il movimento del manipolatore, da una postura iniziale ad una postura finale. Si tratta di definire sia il percorso geometrico che la sua traiettoria. A titolo di chiarezza occorre precisare le differenze concettuali tra i due termini. Con il termine percorso si indica il luogo dei punti nello spazio dei giunti, o dello spazio operativo, che il manipolatore deve descrivere nell'esecuzione del movimento assegnato; il percorso riguarda solo una descrizione di tipo geometrico. Si indica invece con traiettoria un percorso in cui sia specificata la legge oraria di moto, ossia la dipendenza temporale di posizioni, velocità ed accelerazioni.

In linea di principio si può ritenere che un algoritmo di pianificazione della traiettorie debba avere come ingressi la definizione del percorso, i vincoli al percorso, i vincoli dovuti alla dinamica del manipolatore, e come uscita le traiettorie nello spazio dei giunti espresse come sequenze temporale dei valori assunti da posizione velocità ed accelerazione.

La definizione del percorso può essere fatta sia nello spazio dei giunti che nello spazio operativo; di solito si preferisce quella nello spazio operativo, perché consente in maniera più naturale di descrivere il compito che il manipolatore deve eseguire. Inoltre consente di affrontare problemi legati alla presenza di vincoli sul percorso: questi sono infatti dovuti alla necessità di evitare regioni dello spazio di lavoro in cui vi è presenza di ostacoli o che comunque si vuole impedire al manipolatore di raggiungere. Tali ostacoli sono di difficile individuazione nello spazio dei giunti.

La traiettoria viene scelta dall'operatore in base a criteri di varia natura, essendo i criteri di scelta svariati, e l'impiego del robot flessibile, bisogna adottare la traiettoria dell'organo terminale in relazione alla tipologia di compito che deve assolvere.

Ragioni per cui bisogna tener conto dei seguenti aspetti:

- caratteristiche del ciclo di lavoro: alcune posizioni dello spazio dovranno necessariamente appartenere al percorso, come i punti iniziali e finali, e dovranno inoltre essere realizzate determinate velocità,
- limiti della struttura meccanica del robot, legata ad esempio alla presenza del fine corsa che impediscono determinati movimenti,
- limiti meccanici ed elettrici degli attuatori: questi ultimi non possono fornire valori di forze generalizzate di velocità e di accelerazioni tali da

compromettere la resistenza meccanica della struttura, inoltre gli attuatori elettrici non possono assorbire più di una determinata potenza,

- limiti dovuti al fatto che le forze generalizzate vanno trasmesse ai link mediante organi progettati all'uopo, i quali sono dimensionati per valori di esercizio che non devono essere superati,
- limiti del controllo: sono limiti essenzialmente legati a non linearità, ad errori numerici nell'elaborazione dei dati, ed infine alla massima frequenza di campionamento del sistema di controllo.

L'utente può fornire in dettaglio le informazioni sul percorso dell'utensile; tale scelta comporta un processo oneroso e prende il nome di traiettoria calcolata, oppure fornire un numero di informazioni minori quali i punti estremi, i punti intermedi, e primitive geometriche dei punti calcolati; e tale traiettoria prende il nome di traiettoria punto-punto. In questo caso vengono spesso forniti anche le informazioni sul tempo da impiegare, vincoli relativi alle massime velocità e alle massime accelerazioni da realizzare.

Per l'applicazione sviluppata si è proceduto imponendo che lungo il percorso geometrico il modulo della velocità sia costante e pari alla velocità di avanzamento dell'utensile. Si definisca con V il modulo della velocità, allora si ha che le componenti delle velocità definite in ogni punto della curva $\Psi(\lambda)$ si ottengono come segue:

$$V_x = t_x \cdot V$$

$$V_y = t_y \cdot V$$

$$V_z = t_z \cdot V$$

In figura 6.10 è riportato un esempio di traiettoria ottenuta imponendo il vettore velocità nei punti appartenenti ad essa.

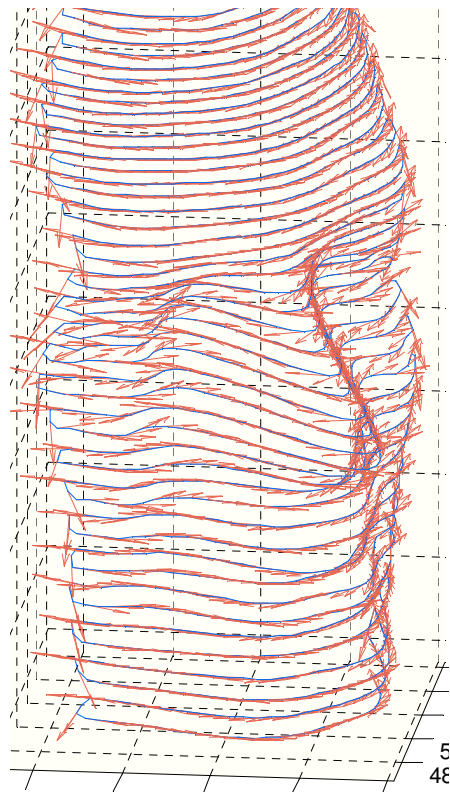
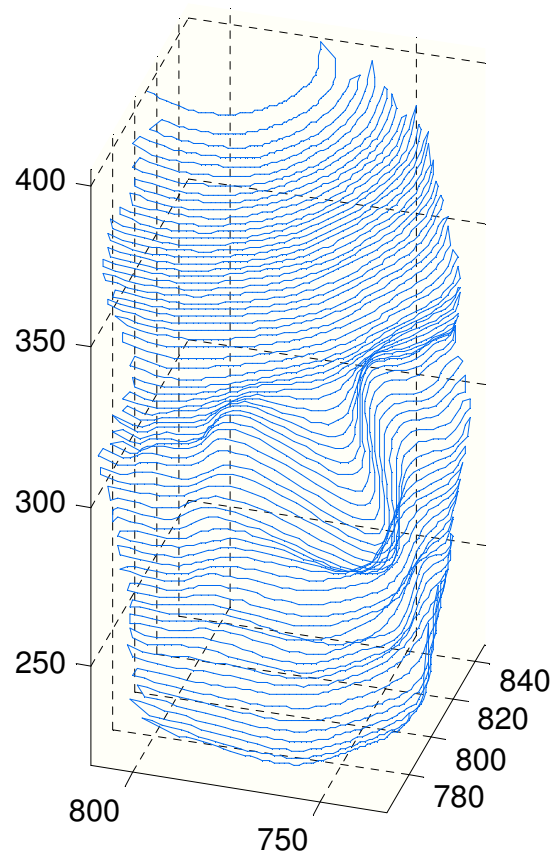


Fig. 6.10

Poiché per ogni punto della traiettoria è assegnato il vettore velocità V in modulo, direzione e verso (tangente alla traiettoria), sono facilmente calcolabili i valori delle velocità dei giunti per ciascuno dei punti della traiettoria.

Si ottiene così un insieme di valori di posizioni e velocità per ciascuno dei giunti del manipolatore che vengono forniti al sistema di controllo.

In tal modo, ciascuno dei servomotori seguirà una legge del moto, variabile istante per istante, tale da far in modo che l'organo terminale transiti per ciascuno dei punti presi sulla traiettoria con la velocità assegnata in modulo direzione e verso. In figura 6.11 è riportato uno schema nel piano raffigurante una condizione di moto dell'organo terminale adottando la tecnica di pianificazione riportata in precedenza.

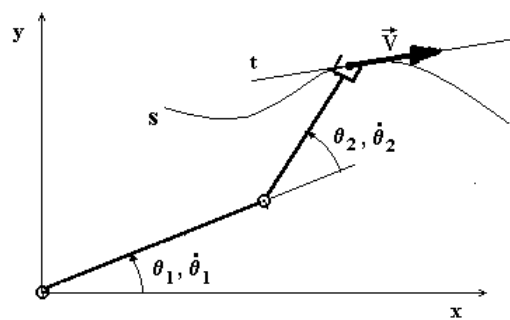


Fig. 6.11

Tale modo di procedere consente di ottenere traiettorie più precise di quelle ottenute con i metodi mediante i quali la traiettoria è costituita da un insieme di punti collegati da spostamenti dei giunti che si muovono seguendo una legge del moto fissata una volta per tutte, senza la possibilità di imporre la direzione della velocità (tangente alla traiettoria) [25].

In particolare, nella fig. 6.12 è riportata una traiettoria sinusoidale ottenuta col metodo descritto e, per confronto, la traiettoria che si ottiene con un controllo punto-punto, assegnando le coordinate di 9 punti nello spazio cartesiano di lavoro.

Nella fig. 6.13 sono riportate, insieme alla traiettoria circolare che si desidera realizzare (linea tratteggiata), la traiettoria che si ottiene col metodo su esposto (fig. 6.13,a) e quella che si ottiene assegnando 18 punti nello spazio cartesiano di lavoro (fig. 6.13,b).

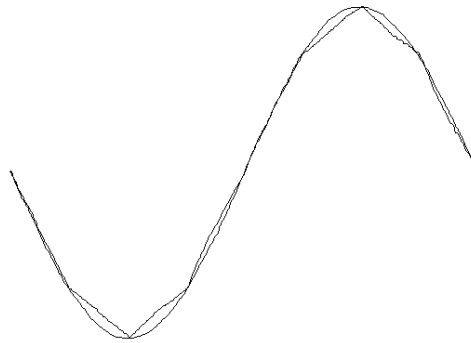


Fig. 6.12

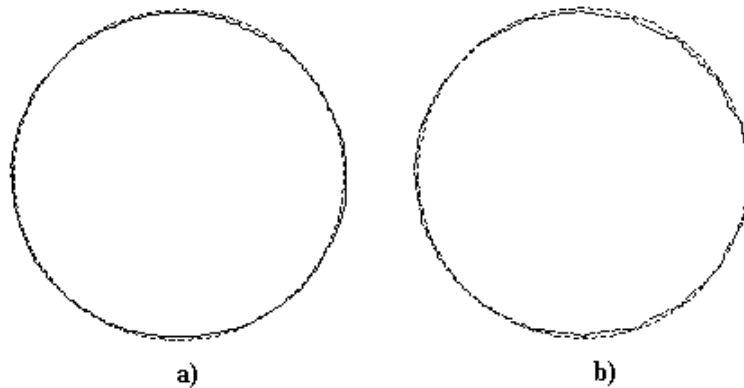


Fig. 6.13

6.3 Controllo del robot [29]

Il robot è controllato mediante la scheda controllo assi PCI-7344 della NATIONAL INSTRUMENT®, tramite la quale contemporaneamente si comandano i servomotori e si acquisiscono i dati. Fisicamente, la scheda si comporta come una periferica del PC, che interagisce con esso, per poter implementare il controllo. È possibile gestire questa operazione completamente nell'hardware della scheda, salvando su di essa precise sequenze di istruzioni. La PCI-7344 è dotata di due processori: un'unità di processo centrale (CPU) ed un processore dedicato ai segnali digitali (DSP), essa risulta essere programmabile ed è dotata di una memoria non volatile. Il processore centrale gestisce i comandi di esecuzione, la sincronizzazione col computer ed i segnali di ingresso e uscita. I comandi a disposizione della scheda la rendono flessibile ed adattabile alle più generiche funzioni di automazione industriale, consentendo di comporre le istruzioni base per eseguire un certo numero di

traiettorie. Il linguaggio di programmazione della scheda è stato gestito su piattaforma LABVIEW, sfruttando l'interfaccia grafica di semplice uso e molto intuitiva, che ha consentito di compilare il programma di gestione interamente realizzato nel dipartimento (DIME). Il programma è impostato secondo due modalità, che differiscono tra loro solo per la possibilità di eseguire il moto, salvando o meno i dati provenienti dai sensori di posizione. Questa opzione è stata introdotta, al fine di rendere più immediata l'elaborazione dei comandi, quando non vi sia la necessità di acquisire informazioni sul moto. È possibile gestire semplici spostamenti punto-punto, così come traiettorie complesse nello spazio di lavoro con passaggio per vari punti intermedi. Il compito di una scheda di controllo è quello di valutare istantaneamente la posizione e l'errore commesso, in base alla lettura precedentemente fornita dai resolver, tale informazione, elaborata dal controllore (PID), costituisce la base del comando di retroazione. Nella scheda adottata, ogni motore viene riconosciuto come servoasse, e dispone di due chip, che calcolano l'istantanea posizione per ogni aggiornamento della lettura dei sensori. Formalmente un servoasse dispone di un generatore di traiettoria, che valuta i segnali da inviare al servoamplificatore in base alla nuova posizione da raggiungere, di una uscita, costituita dal convertitore digitale-analogico (DAC), per la generazione del segnale in corrente (sinusoidale, ± 10 V), e di un anello di retroazione. La chiusura dell'anello di retroazione è realizzata dai sensori di posizione dei motori, i resolver, il cui segnale tramite il servoamplificatore, è convertito in un segnale digitale tipo encoder incrementale, da inviare in ingresso alla scheda. Il segnale encoder si compone di due segnali sfasati tra di loro di $\frac{1}{4}$ di passo, e di un segnale di riferimento con un impulso a giro, codificati a 32 bit, figura 6.14. Il circuito di conteggio della scheda conta i fronti di salita e di discesa dei segnali delle due fasi, dopo N giri dell'asse motore il conteggio si sarà incrementato di un valore pari a $4 \times N$ count, dove il count rappresenta il numero di onde quadre del segnale, generate per un giro dell'asse motore. Esso costituisce il parametro rappresentativo della risoluzione del sensore, stabilita quest'ultima, resta determinata la risoluzione angolare adottata dal sistema di controllo. Nel nostro caso, è stata adottata una risoluzione di 1024 count a giro, pari a $(360/1024)^\circ \approx 0.35^\circ$. Il segnale di riferimento serve per la valutazione

dello stato iniziale necessario al circuito di conteggio per individuare la posizione assoluta.

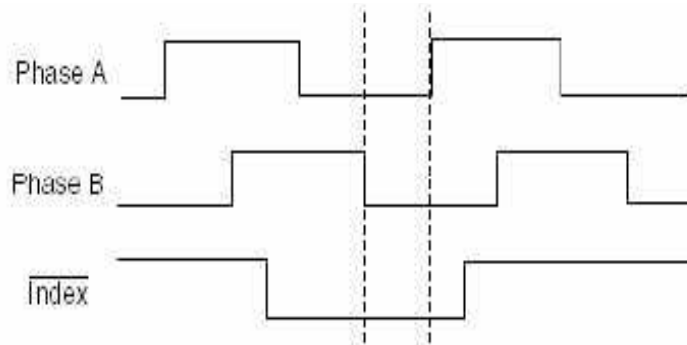


Fig. 6.14

Il controllore della scheda usa le informazioni dell'anello di retroazione per gestire il comando degli attuatori, in qualunque tipologia di traiettoria. La gestione delle leggi di moto da realizzare, è effettuata sfruttando l'insieme delle istruzioni che la scheda ha in dotazione.

6.3.1 Il comando "blend"

Per uno spostamento di tipo punto-punto, la scheda invia i segnali di moto all'amplificatore per portare i motori nella posizione finale impostata, utilizzando una legge del moto prestabilita per compiere lo spostamento. Il controllore della scheda agirà lungo l'intero spostamento per garantire che venga rispettata la legge del moto impostata. Nella maggior parte degli azionamenti industriali, la legge del moto utilizzata per uno spostamento da un punto al successivo è del tipo "ad s", con una velocità trapezoidale ed una accelerazione a gradini. La realizzazione di traiettorie complesse prevede l'assegnazione, alla scheda di controllo, di una certa quantità di informazioni relative ai punti intermedi della traiettoria stessa, tali dati saranno in quantità tanto maggiore quanto maggiore è la precisione che si vuole ottenere. La gestione del moto per punti intermedi, prevede un controllo del tipo punto-punto, iterato sull'intera traiettoria. La legge del moto utilizzata per il singolo spostamento da un punto al successivo, è caratterizzata da un tratto in accelerazione, uno a velocità costante, ed uno in decelerazione fino all'arresto nel punto finale. Una

tale legge si presta poco alla realizzazione di una traiettoria complessa per punti intermedi, infatti, si otterrebbe un movimento a scatti con arresti in ogni punto della traiettoria stessa. Per ovviare a questo inconveniente si utilizzano opportune istruzioni previste dai sistemi di controllo utilizzati. La scheda di controllo assi NI PCI-7734, prevede tra le sue funzioni di base, il comando “blend”, che permette di eseguire una qualsiasi traiettoria, passante per i punti intermedi desiderati, senza mai arrestare il moto, realizzando un unico spostamento continuo. Il “blend” è sostanzialmente un set di istruzioni presenti nel software della scheda controllo assi, che regola la sovrapposizione dei profili di velocità di due movimenti successivi riuscendo a garantire la continuità del moto.

L’istruzione opera tra due spostamenti successivi che ogni singolo motore deve realizzare, in modo tale da raccordare il profilo della velocità senza portarla a zero. La sovrapposizione dei profili di velocità interessa la fase di decelerazione del primo spostamento e quella di accelerazione del secondo, come mostrato in figura 6.15: il movimento, ancora in esecuzione, viene proseguito in modo tale che, a partire dall’istante in cui deve iniziare la decelerazione, il motore si porti alla velocità massima prevista per il successivo movimento, adottando un’opportuna accelerazione. Se i due movimenti sono caratterizzati dalla stessa velocità massima, ed avvengono senza cambiamento di direzione, si ha un perfetto raccordo dei due profili, con l’assenza di eventuali fasi di accelerazione o decelerazione.

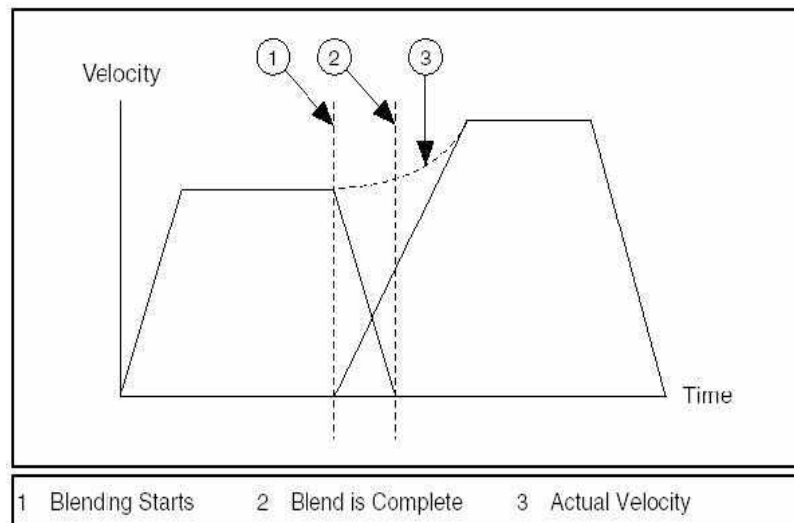


Fig. 6.15

Il comando “blend” è gestibile tramite un parametro indicato come “fattore di blend”, che può assumere valori -1, 0 o un numero positivo intero.

L’assegnazione del valore -1 al parametro, attua la sovrapposizione dei profili di velocità, garantendo un movimento fluido e continuo, figura 6.15. La velocità nel tratto di sovrapposizione, dipende dalla velocità di crociera, dalla decelerazione e dal jerk del primo movimento, e dal jerk, dall’accelerazione e dalla velocità di crociera del secondo movimento. Il valore zero del fattore di blend, comporta l’arresto istantaneo dell’attuatore dopo l’esecuzione di ogni singolo movimento, non è realizzata nessuna sovrapposizione dei profili comportando un movimento complessivo discontinuo, figura 6.16.

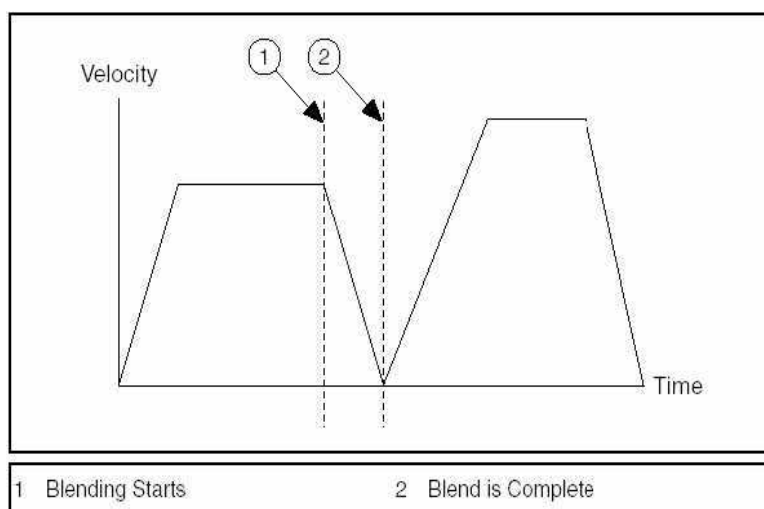


Fig. 6.16

Un valore positivo del parametro di blend rappresenta un tempo di attesa in millisecondi: il moto dopo l’esecuzione di ogni singolo spostamento da un punto al successivo, si arresta per un tempo pari a quello impostato dal parametro portando a zero la velocità, figura 6.17.

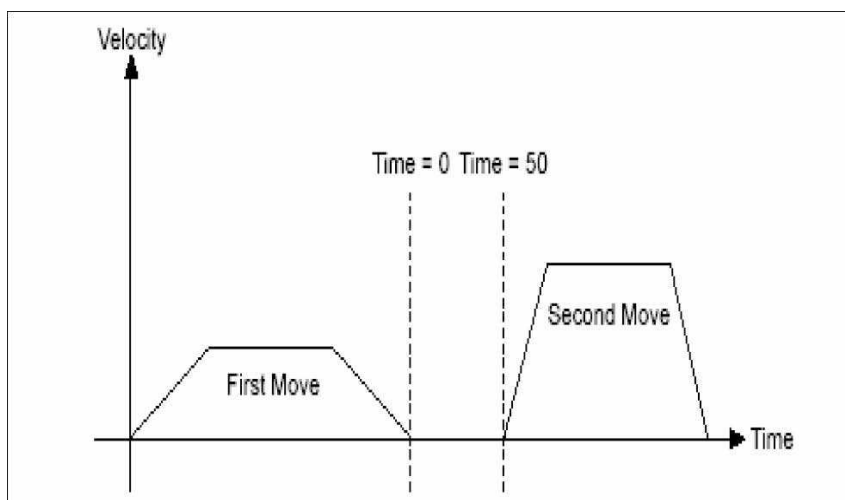


Fig. 6.18

6.3.2 Scelta dei punti intermedi in una traiettoria continua

La sovrapposizione dei profili di velocità per movimenti successivi, comporta un effetto di raccordo per i punti intermedi delle traiettorie, effetto di “smooth”. Nella figura 6.19. è mostrato un esempio dell’effetto del fattore di blend su una traiettoria costituita da due segmenti ortogonali tra di loro. Il punto intermedio non viene raggiunto con precisione, in prossimità di esso la traiettoria è arrotondata come conseguenza della sovrapposizione dei profili di velocità, ma questa situazione garantisce continuità al moto.

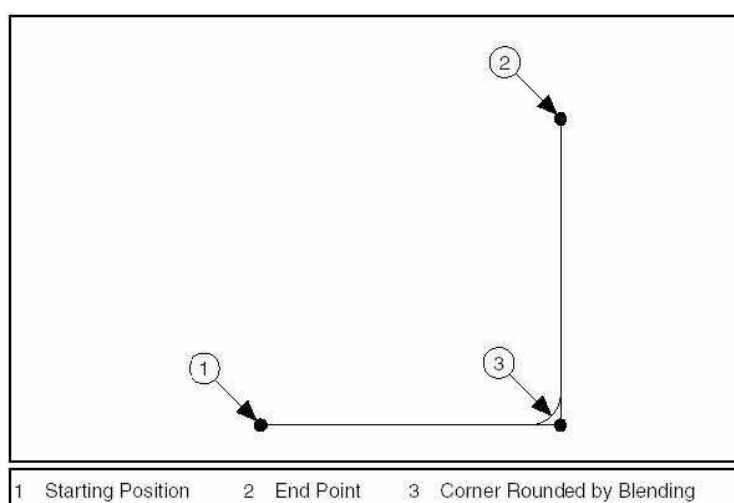


Fig. 6.19

Per garantire il passaggio per i punti intermedi assegnati è necessario ricorrere a punti di percorso fittizi. Nell'ipotesi che sia realizzata la sovrapposizione delle velocità, in ogni punto intermedio della traiettoria, bisogna garantire che sia stato percorso lo spostamento stabilito tra esso ed il precedente, e che si raggiunga la velocità desiderata nel punto in questione.

Per ogni spostamento tra un punto ed il successivo, si assegna una velocità fittizia, tale che con la sovrapposizione dei profili di velocità, ad ogni punto della traiettoria finale corrisponda la velocità corretta. Facendo riferimento alla figura 6.20, sono state indicate con l'apice tutte le grandezze "calcolate" per applicare la sovrapposizione dei profili di velocità, mentre non ne sono dotate quelle grandezze che risultano assegnate. Lo spostamento $S_{(i+1),i}$ tra i punti i ed $i+1$, è assegnato alla scheda di controllo con una velocità di crociera calcolata pari a v_i' , tale che con la sovrapposizione dei profili di velocità, esso venga realizzato arrivando nel punto i con la reale velocità v_i richiesta in questa posizione.

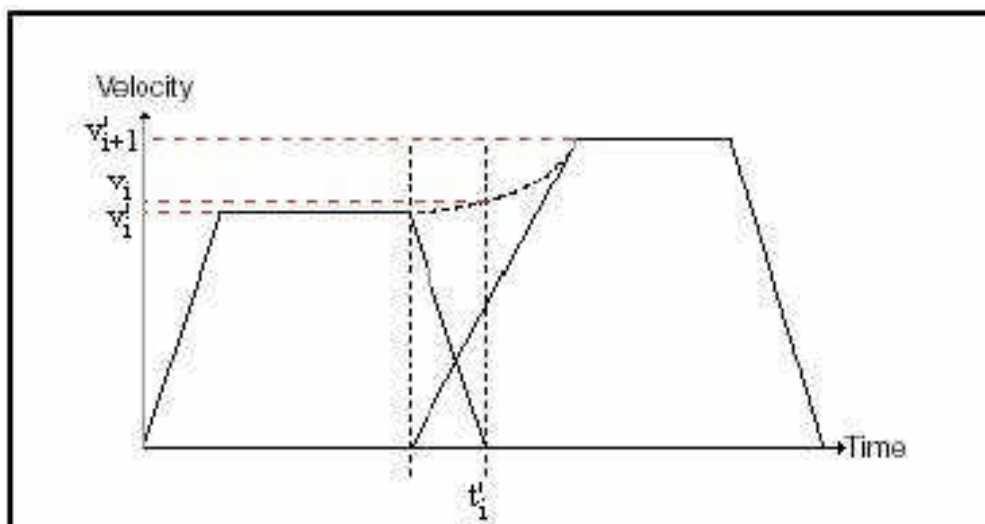


Fig. 6.20

In prima approssimazione è possibile aumentare le accelerazioni dei singoli spostamenti da un punto al successivo, al fine di ottenere dei raccordi piccoli ed approssimabili a tratti rettilinei; questa operazione consente di ricavare più semplicemente i valori delle velocità "calcolate" v_i' . L'istruzione blend è stata introdotta nei programmi realizzati per la gestione del prototipo, può essere attivata

per un solo asse come per tutti in contemporanea. La legge del moto prevista per ogni singolo spostamento punto-punto, ha un profilo di velocità di tipo trapezoidale, nel quale lo jerk (derivata della accelerazione) assume valore infinito in corrispondenza dei cambi di pendenza. Esiste la possibilità di prevedere un valore non infinito dello jerk con un arrotondamento del profilo di velocità in prossimità dei cambi di pendenza.

6.3.3 Software di gestione della scheda di controllo assi

La movimentazione e l'acquisizione dati, come già accennato, viene realizzata tramite un programma interamente sviluppato nel dipartimento. Questo programma permette di eseguire un qualsiasi movimento dell'estremità del robot all'interno dello spazio di lavoro. L'insieme delle istruzioni implementate dall'hardware della scheda di controllo sono state gestite attraverso l'ambiente di programmazione LabView, realizzando un codice con elevate capacità di interfacciamento con l'utente. La schermata principale del programma è mostrata in figura 6.21. È subito evidente nella parte bassa l'opzione per la scelta tra le due modalità di esecuzione, e di lato l'intervallo di campionamento adottato per la lettura ed il salvataggio dei dati provenienti dai sensori, impostato a 50 millisecondi. Il programma prevede sia la possibilità di eseguire il movimento semplicemente, sia di acquisire contemporaneamente i dati dei sensori relativi alla posizione ed alla velocità. Con tale opzione è stato possibile valutare ed escludere l'effetto di ogni possibile causa di disturbo, come l'acquisizione ed il salvataggio dei dati, che potrebbe influenzare il sistema di controllo durante l'esecuzione di una traiettoria complessa o descritta da molti punti.

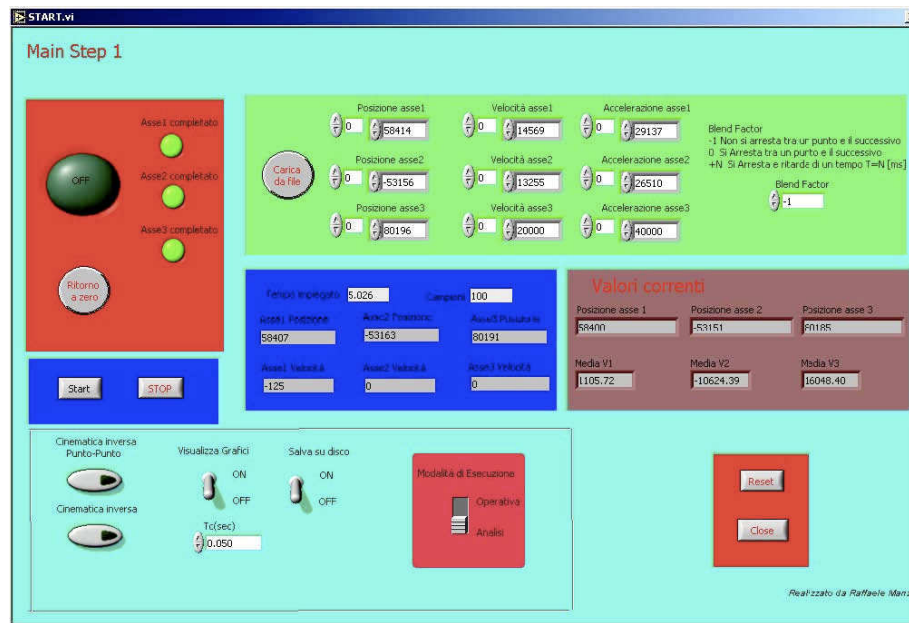


Fig. 6.21

Al centro della schermata sono presenti due riquadri che riportano le posizioni correnti e le velocità dei tre giunti, mentre in alto a sinistra vi è il pulsante di ritorno allo zero (posizione di riferimento), corredato di tre indicatori luminosi, uno per ogni motore, che si accendono a moto compiuto. È presente, in alto a destra, un riquadro in cui è possibile assegnare direttamente posizioni, velocità ed accelerazioni per uno o più punti, oppure caricare le stesse informazioni da un file esterno. Altre due task, attivate dai pulsanti “Cinematica inversa punto-punto” e “Cinematica inversa”, consentono di comandare lo spostamento punto – punto, scalato nel tempo o nell’ampiezza, e non scalato, rispettivamente per due punti la prima, e per più punti la seconda, in quest’ultima è anche possibile assegnare dall’esterno il file delle posizioni, figure 6.22, 6.23.

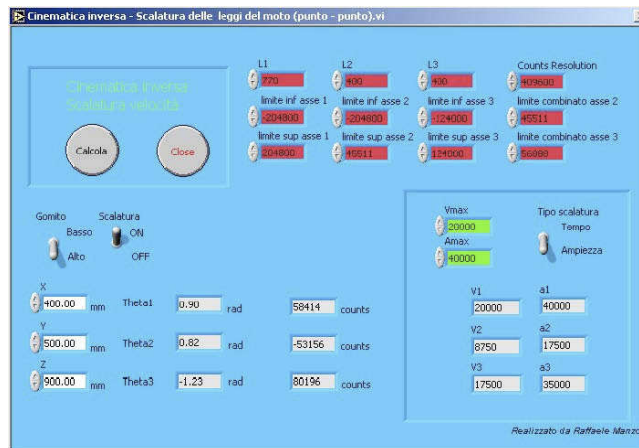


Fig. 6.22

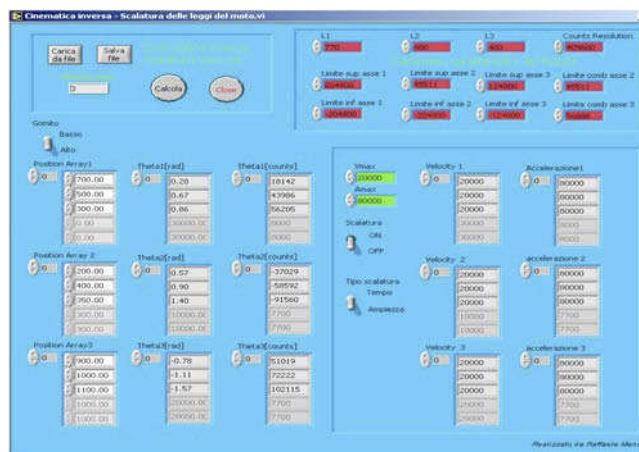


Fig. 6.23

Nelle due task di comando delle operazioni di cinematica inversa è possibile scegliere la modalità, gomito alto o gomito basso, con cui far realizzare il moto dal robot, per tenere conto del fatto che la soluzione della cinematica inversa non è univocamente definita. Inoltre è possibile impostare le dimensioni dei link, la risoluzione di un giro completo a valle del motore, valutato nell'unità di misura di riferimento impostata (count), i finecorsa software per ogni link e quelli combinati che considerano la somma degli spostamenti compiuti dai singoli link.

Se è attivata la modalità di esecuzione che comprende anche l'acquisizione dei dati, tutte le traiettorie eseguite dal prototipo, tramite l'uso di questo programma, possono essere analizzate. È possibile visualizzare i grafici dello spostamento e delle velocità dei motori a valle dei riduttori, figura 6.24, e di salvarne i dati su file in formato testo, attraverso i pulsanti "Visualizza grafici" e "Salva su disco" della schermata principale.

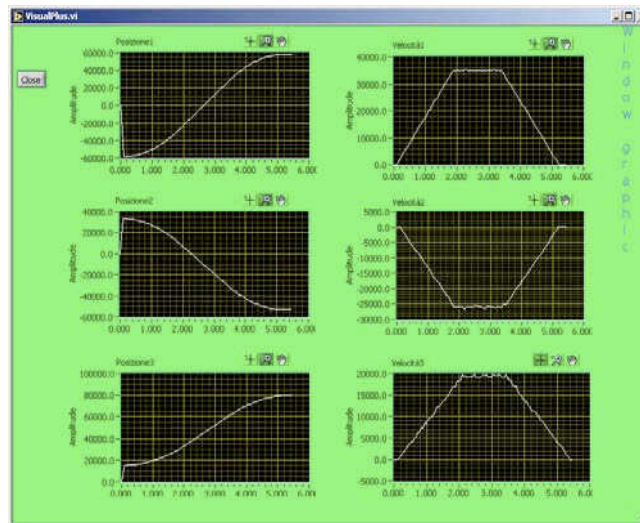


Fig. 6.24

6.4 Risultati

Una volta ottenute le nuvole di punti degli oggetti scannerizzati ci si è proposto di realizzare mediante una lavorazione meccanica le superfici utilizzando come materiale polistirolo e come dispositivo per la lavorazione un Dremel equipaggiato con una fresa sferica di diametro 4 mm; grazie a questa geometria si garantisce il taglio qualsiasi sia la direzione di avvicinamento dell'utensile verso la superficie da lavorare, evitando quindi problemi dovuti a vincoli di taglio che si possono avere per esempio nel caso di fresa cilindrica, figura 6.25.

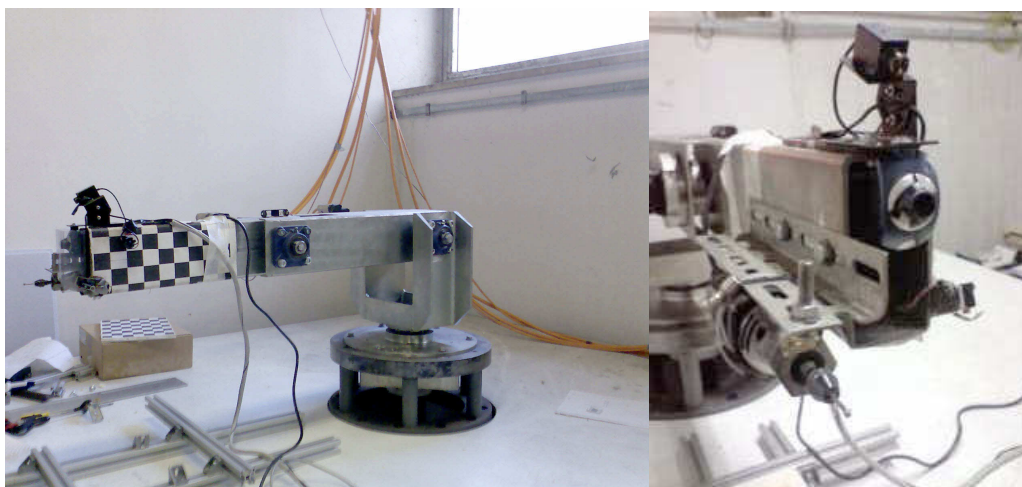


Fig. 6.25

Per validare la tecnica di ricostruzione sono stati scelti due oggetti, uno riconducibile alla classe di dispositivi del settore industriale, cioè una paletta di un compressore assiale, l'altro scelto come esempio per applicazioni di ricostruzione nel campo dei beni culturali, cioè la testa di un manichino. Nelle figure 6.26 e 6.27 sono riportati i due oggetti test e parti della loro superficie realizzate in differenti scale.

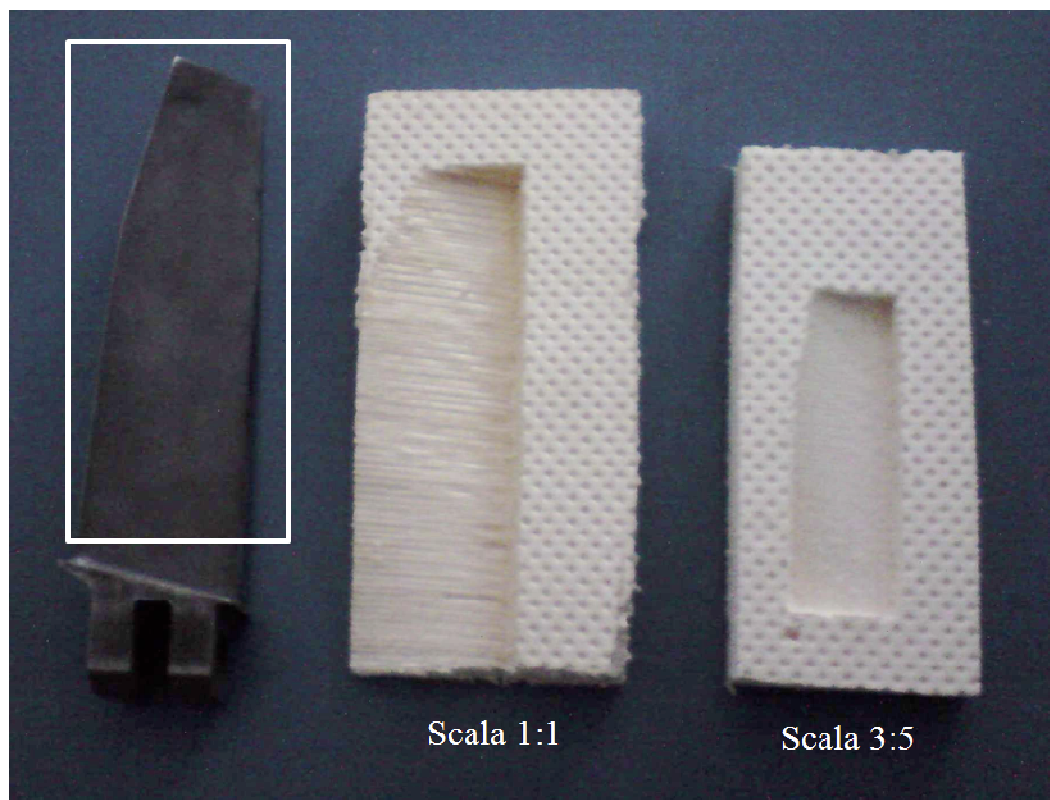


Fig. 6.26



Fig. 6.27

E' da notare che è stato possibile assegnare due tipologie di traiettorie, una di sgrossatura e l'altra di finitura ottenuta dalla prima interpolando i punti dati e assegnando una velocità di avanzamento dell'utensile più bassa.

Il materiale utilizzato è stato scelto prendendo in considerazione i vincoli dovuti agli utensili a disposizione, ciò non toglie la possibilità di lavorare altri materiali scegliendo in modo adeguato i dispositivi per la lavorazione.

CONCLUSIONI

E' stato studiato e sperimentato il processo di ricostruzione 3D con metodi ottici mostrando che con l'ausilio di un manipolatore noto, pur non adottando un sensore di elevata precisione, si ottengono dei risultati soddisfacenti. Si è scelto di adottare come tecnica di acquisizione di forma la triangolazione laser, così facendo si consente al robot di riconoscere l'ambiente circostante in real time con precisione, robustezza e ripetibilità. E' stata messa a punto una tecnica di calibrazione che consente di calibrare l'intero sistema robot – laser – scanner riducendo gli errori legati al posizionamento relativo tra i dispositivi che condiziona in modo sostanziale la ricostruzione 3D. La possibilità di poter conoscere dettagliatamente il sistema di controllo ha permesso di implementare un software per la riproduzione delle superfici scannerizzate fissando sul robot un Dremel con fresa sferica. Il sistema realizzato consente quindi di scannerizzare oggetti e riprodurli eventualmente in scala con estrema flessibilità. Inoltre è possibile riacquisire la forma lavorata per l'analisi della qualità della riproduzione. Questa macchina può essere utilizzata in tantissime applicazioni dove si richiede di acquisire una forma 3D senza contatto e riprodurre la stessa onde farne una copia. Con la procedura descritta per esempio è possibile non solo realizzare la stessa superficie acquisita ma con semplici operazioni definire anche lo stampo per eventuali copie. Uno sviluppo già in atto della macchina è quello di installare sul robot un polso, aggiungendo ulteriori gradi di libertà, avendo quindi la possibilità di analizzare e replicare superfici più complesse.

BIBLIOGRAFIA

- [1] V. Ciccarelli, *Un sistema di visione per robot industriali: realizzazione e sviluppo*. Tesi di Laurea in Ingegneria Meccanica AA 2004/05, Napoli.
- [2] A. Fusiello, *Visione Computazionale: appunti delle lezioni*. Marzo 2008, Verona.
- [3] D. Scaramuzza, *Progetto e realizzazione di un sistema di visione stereoscopica per la robotica, con applicazione all'inseguimento di corpi in moto e all'autolocalizzazione*. Tesi di Laurea in Ingegneria Elettronica AA 2003/04, Perugia.
- [4] J. Heikkilä, O. Silven, *A Four-step Camera Calibration Procedure with Implicit Image Correction*. Infotech Oulu and Department of Electrical Engineering University of Oulu.
- [5] V. Niola, E. Pollasto, C. Rossi, S. Savino *An Algorithm for Kinematic Calibration of Robot Arms*. RAAD 2006.
- [6] G. D'Orsi, *Calibrazione Cinematica di un Braccio Meccanico mediante un Sistema di Visione*. Tesi di Laurea in Ingegneria Meccanica AA 2004/05, Napoli.
- [7] V. Morreale, *Visione Artificiale per Robot Mobile: tecniche per il riconoscimento e l'interpretazione di immagini a colori*. Tesi di Laurea in Ingegneria Elettronica AA 1998/99, Palermo.
- [8] S. Savino, *Progettazione e realizzazione di un sistema di acquisizione delle traiettorie di un Robot mediante telecamera digitale*. Tesi di Laurea in Ingegneria Meccanica AA 1999/00, Napoli.
- [9] B. A. Rosenfeld, *A History of Non-Euclidean Geometry, Evolution of the Concept of a Geometric Space*, Springer Verlag. 1988, New York.
- [10] D. Marr, L. Vaina, *Representation and recognition of the movements of shapes*. 1982, Londra.
- [11] R. Bajcsy, *Active perception*. Agosto 1988.
- [12] A.V. Aho, R. Sethi, J.D. Ullman, *Compilers: Principles, Techniques, and Tools*. 1985, Massachusetts.
- [13] J. S. Albus, A. J. Barbera, *A cognitive architecture for intelligent multi-agent systems*. 1988.

- [14] Azevedo, T. C. S., *3D object reconstruction from uncalibrated images using a single off-the-shelf camera*, Proceedings of VIP IMAGE , Thematic conference on computational vision and medical image processing, October 17-19, 2007.
- [15] Chalidabhongse, T.H.; Yimyam, P.& Sirisomboon, P. *2D/3D Vision-Based Mango's Feature Extraction and Sorting*, ICARCV 2006 (1-6).
- [16] Fremont, V.; Chellali, *Turntable-Based 3D Object Reconstruction*, Proceedings of IEEE Conference on Cybernetics and Intelligent Systems, pp. 1276-1281, Singapore, 2004.
- [17] R. Y. Tsai, *A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses*. IEEE Trans. on Robotics and Automation, Vol. RA-3, No. 4, pp. 323–344, 1987.
- [18] Jean-Yves Bouget – *Camera Calibration Toolbox for Matlab* - http://www.vision.caltech.edu/bougetj/calib_doc/
- [19] D. Acosta, O. García and J. Aponte, *Laser Triangulation for shape acquisition in a 3D Scanner Plus Scanner*, Proceedings of the Electronics, Robotics and Automotive Mechanics Conference, 2006.
- [20] J. Forest, *New methods for triangulation-based shape acquisition using laser scanners*, PhD thesis, University of Girona, 2004.
- [21] C. Colombo, D. Comanducci e A. Del Bimbo, *Low-Cost 3D Scanning by Exploiting Virtual Image Symmetries*, Journal of Multimedia, Vol. 1, No. 7, 2006.
- [22] N. Koller, *Fully Automated Repair of Surface Flaws using an Artificial Vision Guided Robotic Grinder*, PhD thesis, University of Leoben, 2007.
- [23] C. Matabosch, *Hand-held 3D-scanner for large surface registration*, PhD thesis, University of Girona, 2007.
- [24] M. Ritter, M. Hemmleb, O. Sinram, J. Albertz and H. Hohenberg, *A Versatile 3D Calibration Object for Various Micro-range Measurement Methods*, Proceedings of ISPRS, pp. 696-701, Istanbul, 2004.
- [25] C. Rossi, *Appunti dalle lezioni di Meccanica dei Robot*. A.A. 2008-2009, Napoli 2008
- [26] B. Siciliano, *Robotica industriale*, Mc-Hill Libri Italia, Napoli 2000

- [27] S. Larsson and J.A.P. Kjellander, *Motion control and data capturing for laser scanning with an industrial robot*, Robotics and Autonomous Systems, Vol.54, No.6, pp. 453-460, 30 June 2006.
- [28] R.A. Jarvis and Y.L. Chiu, *Robotic Replication of 3D Solids*, Proceedings of IROS, 1996.
- [29] S. Savino, *Progettazione, Realizzazione e Messa a punto di un robot revoluto*, Tesi di dottorato A.A. 2003-2004, Napoli.