

Algoritmi avanzati per applicazioni di misura 3D

Alberto Franciosi

Negli ultimi anni si è assistito ad un crescente interesse verso i sistemi di visione artificiale come parti integranti di sistemi di automazione industriale. Il continuo evolversi della tecnologia in processori sempre più potenti e la commercializzazione di sensori ottici digitali a risoluzione sempre più elevata, ha permesso notevoli sviluppi dei sistemi di visione in termini di precisione, efficacia, velocità e flessibilità, tali da poter superare il limite della visione intesa come acquisizione bidimensionale del mondo reale. Tali tecniche sono risultate di notevole interesse non solo per le applicazioni di ispezione superficiale, ma anche per tutte le applicazioni che prevedono un abbinamento dei sistemi di visione a robot antropomorfi per le più svariate applicazioni: dalla pura movimentazione di oggetti (conveyor tracking e pallettizzazione) alla loro individuazione nelle tre dimensioni per la presa da cassone o da mucchio. Nell'articolo ci si limiterà ad analizzare i più quotati sistemi di misurazione e localizzazione 3D attualmente utilizzati nell'ambito dell'automazione industriale, mostrandone i principi base ed individuandone peculiarità e limiti.

Keyword

Visione artificiale, Misura 3D, Calibrazione

L'utilizzo del 3D nel campo dell'automazione è principalmente rivolto ad esigenze di localizzazione di oggetti nelle tre dimensioni o di ispezione superficiale di oggetti non planari. Esistono diverse tecniche specifiche per diverse tipologie di applicazioni, ognuna con peculiarità e limiti ben definiti.

Localizzazione di oggetti in tre dimensioni

Gli algoritmi di localizzazione 3D aggiungono flessibilità ad un sistema di presa robot, fornendo così tre coordinate relative alla posizione dell'oggetto e tre angoli relativi alla sua posa. A valle di ciò è ovviamente necessaria da parte del sistema di visione una conoscenza totale o parziale dell'oggetto da individuare.

Ricerca prospettica di un modello

Un primo approccio è la ricerca dei contorni di un modello planare indeformabile precedentemente memorizzato, indipendentemente dalla sua deformazione prospettica.

Nella memorizzazione del modello, vengono derivate le relative informazioni e vengono memorizzate insieme al modello stesso a diverse risoluzioni in una struttura piramidale. Così facendo la ricerca del modello è più veloce, tenendo in considerazione che una deformazione prospettica è meno risentita a bassi livelli di risoluzione del modello.

Una volta individuato il modello, questo viene ricercato nei diversi layers memorizzati ottenendo una localizzazione precisa e ottenendo così la deformazione ottenuta e quindi la localizzazione e la posa dello stesso nelle tre dimensioni.



Figura 1 - Ricerca prospettica del modello basato su edge

La validità di tale sistema è legata alla natura dell'algoritmo stesso, pertanto è efficiente solo con oggetti con edge ben definiti e su superfici planari e soprattutto su oggetti non deformabili; per questo molto usata su oggetti metallici in ambito automotive.

Un secondo approccio molto simile, sfrutta la localizzazione di punti di interesse nell'immagine (mediante operatori di Harris o di LePetit) precedentemente memorizzati in un modello. Pertanto tale tecnica è utilizzata su oggetti con superfici planari ricche di texture. La perdita di accuratezza di tale tecnica rispetto



Figura 2 - Ricerca prospettica del modello basato su descrittori

alla precedente è ricompensata da una maggiore velocità e soprattutto robustezza, risultando efficace anche in presenza di variazioni di luminosità e di parziali occlusioni dell'oggetto.

Pattern matching da disegno Cad 3D dell'oggetto

Una più precisa localizzazione dell'oggetto nelle tre dimensioni è attuabile avendo a disposizione una completa descrizione della forma dell'oggetto, nello specifico con un file CAD a tre dimensioni.

Il primo passo è da ricercarsi nella generazione del modello: difatti dal disegno CAD a tre dimensioni è necessario recuperare i modelli 2D relativi a ciascuna possibile vista del modello: ipotizzando la visualizzazione dell'oggetto posto al centro di una sfera tramite telecamere virtuali poste sulla superficie della sfera stessa e rivolte verso il centro. Il problema principale è ovviamente la vastità del numero dei modelli che dovrebbero essere generati, per questo motivo la migliore soluzione è il restringimento del range di posizionamento della telecamera (come mostrato nella figura 3) abbinato ad una memorizzazione gerarchica dei modelli.

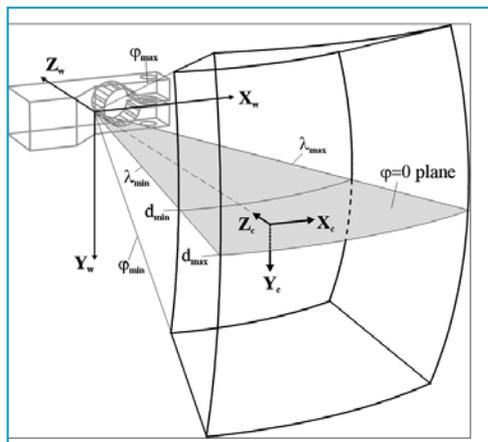


Figura 3 - Restrizione del range di posizionamento o della telecamera

Per poter fornire quindi una localizzazione più precisa mantenendo tempi di elaborazione contenuti, si procede con una ricerca ad albero esaustiva dei modelli 2D memorizzati. Le informazioni di posizionamento relative al modello individuato offrono un punto di partenza per individuare la vera posa 3D dell'oggetto, ottenibile mediante affinamenti iterativi fino alla convergenza della posizione reale. Il numero di step di iterazione, il restringimento del range di posizioni assumibili, il numero di oggetti da individuare e la mutua sovrapposizione degli oggetti presenti nell'immagine sono tutti parametri che influiscono notevolmente sui tempi di elaborazione. Un altro modo per velocizzare l'algoritmo è sfruttare le eventuali simmetrie dell'oggetto.

L'accuratezza del sistema dipende da diversi fattori, innanzitutto dalla dimensione dell'oggetto all'interno dell'immagine, in secondo luogo dagli errori risultanti dalla ricerca degli edge e del modello nell'immagine. Nella figura 4 è mostrato uno grafico di massima che mostra la stretta relazione tra accuratezza e dimen-

sione dell'oggetto, scelta dell'ottica e distanza dall'oggetto. Nella tabella sono invece presenti i risultati di due applicazioni acquisite con telecamere a risoluzione standard 640x480 con ottica da 8.5 mm e processate su processore Intel Xeon E5345 da 2,33 GHz.

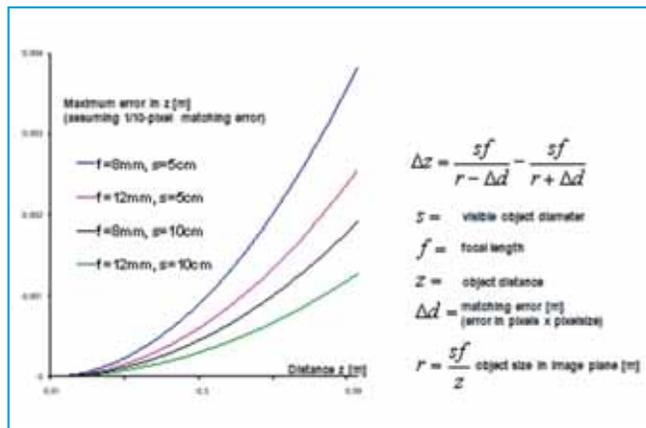


Figura 4 - Errore massimo di localizzazione lungo l'asse z

Object	Pose Range	σ_{pos} [mm]	σ_{pos} [%]	σ_{rot} [°]	Time [s]
	$\Delta\lambda = \Delta\phi = [-50,+50]^\circ$ $\Delta d = [150,200]$ mm	0.39mm	0.20	0.48	0.3
	$\Delta\lambda = \Delta\phi = [-50,+50]^\circ$ $\Delta d = [250,300]$ mm	0.74	0.21	0.60	0.9

Tabella - Benchmark



Figura 5 - Alcuni screenshot applicativi

Ispezione superficiale con tools 3D

Le informazioni nelle tre dimensioni dell'oggetto analizzato offrono un notevole interesse per tutte le applicazioni di ispezione e controllo superficiale di oggetti. Differentemente dalla localizzazione di oggetti in cui il goal è l'individuazione dell'oggetto nello spazio, in questo caso lo scopo dell'applicazione è ottenere informazioni di quota dell'oggetto inquadrato, associare cioè ad ogni pixel dell'immagine un punto del mondo reale rappresentato dalle tre coordinate reali x, y, z. La situazione è

ben differente dalla precedente, ci troviamo difatti davanti ad un oggetto di forma non nota a priori e dobbiamo individuarne le caratteristiche tridimensionali. In letteratura i sistemi esistenti più quotati si basano sulle tecniche di *depth from focus*, sistemi a luce strutturata e di visione stereoscopica.

Depth from focus

L'idea si basa sulla limitata profondità di campo delle ottiche, difatti in dipendenza dalla distanza dalla telecamera solo alcuni oggetti risultano a fuoco (contorni ben definiti). Il concetto si basa sulla movimentazione della telecamera e la memorizzazione delle immagini associandone la distanza di acquisizione. La determinazione della quota si baserà su un'analisi qualitativa della nitidezza per ogni pixel in ogni immagine memorizzata. Ovviamente la risoluzione relativa alla quota dell'immagine dipende strettamente dalla profondità di campo dell'ottica utilizzata, pertanto tale sistema viene utilizzato solo per misure superficiali di piccoli oggetti con piccole variazioni di quota. Per ottenere una corretta misurazione bisogna attenersi a regole ben precise, come ad esempio non sovraesporre l'immagine ed evitare qualsiasi tipo di riflesso (utilizzando quindi una luce diffusa), astenendosi inoltre dall'utilizzo di telecamere con un alto contenuto di rumore.

In ultimo, ma di fondamentale importanza, è necessario tener conto dell'errore di misura dovuto all'aberrazione (inquadrandolo un oggetto planare la distanza dalla telecamera del punto inquadrato al centro dell'immagine è inferiore a quella di un punto posto in un angolo dell'immagine stessa).

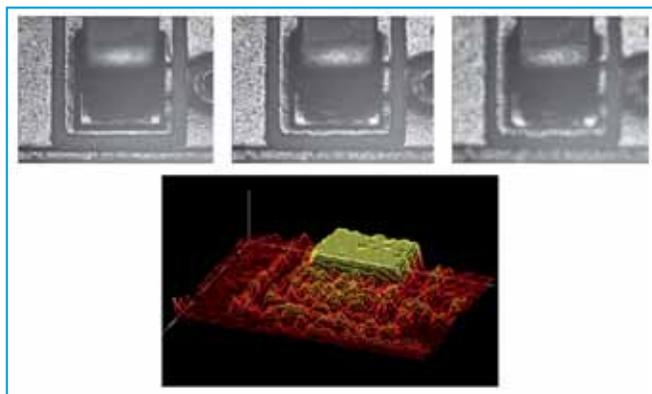


Figura 6 - Esempio di ricostruzione 3D con tecnica Depth From Focus

Sistemi a luce strutturata

I sistemi a luce strutturata sono ad oggi i sistemi di ricostruzione 3D più diffusi. Il concetto alla base di queste tecniche è l'analisi della deformazione di una lama laser proiettata su un oggetto non planare in movimento controllato. La lama laser proiettata definisce un piano nello spazio (sheet of light), questo piano rispecchia la sezione dell'oggetto in una precisa posizione. La sequenza delle immagini ottenute in seguito ad una movimentazione relativa di oggetto e telecamera viene seguito successivamente processata e in seguito ad un processo di calibrazione è possibile ottenere le coordinate reali dell'oggetto x, y e z nelle tre dimensioni, dove x è determinata dalla traslazione controllata dell'oggetto, y è rappresentata dalla coordinata longitudinale

dell'immagine e z è calcolata considerando lo scostamento laterale della linea laser.

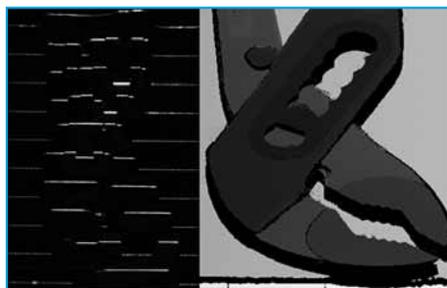


Figura 7 - Esempio di acquisizione a luce strutturata

Limitazioni derivano dalla tecnologia laser che non solo non garantisce una perfetta uniformità di spessore della lama proiettata, ma soprattutto su immagini porose introduce rumore (speckle) riducendo la nitidezza della lama. Un altro problema tipico di questi sistemi, dovuto al posizionamento relativo telecamera lama laser, è il manifestarsi di zone d'ombra, come mostrato nella figura 8. Infine possono sussistere problemi dovuti alla geometria dell'oggetto stesso: riflessi e inspessimento della linea proiettata in corrispondenza di curvature sul materiale, come mostrato nella figura 9. Una corrispondenza precisa in coordinate reali è possibile solamente in seguito ad un'attenta calibrazione del sistema ottenuta utilizzando un target di calibrazione noto.

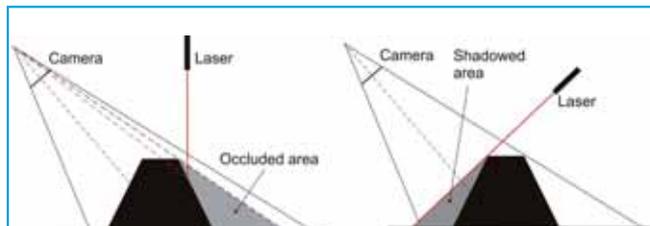


Figura 8 - Zone d'ombra dovute alla geometria del sistema

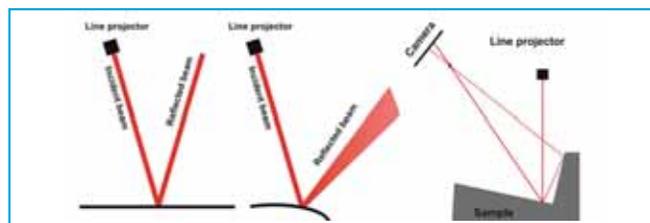


Figura 9 - Limitazioni causate dalla geometria dell'oggetto

Visione stereoscopica

L'utilità di un sistema di visione stereoscopica è senza dubbio la possibilità di recuperare informazioni tridimensionali relative alla superficie di un oggetto di dimensione e forma non note a priori. La necessità è di avere a disposizione due immagini inquadranti lo stesso oggetto acquisite al medesimo istante ma da posizioni diverse, il sistema è composto quindi da due telecamere inquadranti lo stesso oggetto poste a distanza fissa e nota. Le immagini acquisite devono essere calibrate e rettificare in modo tale da poter permettere una misura accurata.

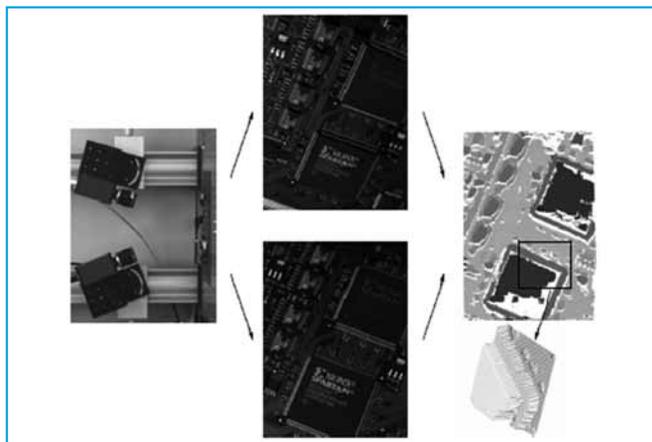


Figura 10 - Schema di principio di un sistema di visione stereoscopico

La distanza delle telecamere deve essere studiata in modo tale da massimizzare l'area coperta da entrambe le telecamere evitando possibilmente occlusioni, tenendo conto però che tale distanza è direttamente proporzionale all'accuratezza del sistema. Inoltre è consigliato posizionare le telecamere in modo tale che la congiungente risulti parallela al piano inquadrato.

Il funzionamento di principio è molto semplice e si basa sulla ricerca di punti corrispondenti nelle due immagini e su una successiva triangolazione essendo note le coordinate di posizionamento delle due telecamere. Se le telecamere sono state posizionate correttamente e se è stata implementata la rettificazione per entrambe le immagini, i punti corrispondenti giacciono su una stessa linea detta epipolare, come mostrato nella figura 11. Tale accorgimento riduce sensibilmente i tempi di elaborazione in fase di ricerca dei punti corrispondenti.

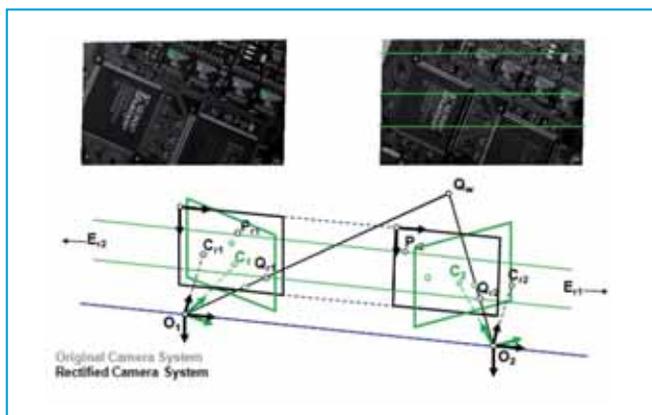


Figura 11 - Linee epipolari su immagini rettificate

Il cuore dell'applicazione è quindi la ricerca dei punti corrispondenti. L'approccio standard è il matching tramite correlazione, molto veloce e robusto in corrispondenza di variazioni di luminosità, ma con il limite di funzionare solo in prossimità di aree ricche di texture. Un approccio alternativo denominato multigrigrid stereo matching si basa su un approccio legato al calcolo delle variazioni con algoritmi *multigrigrid*, in grado così di ricostruire mediante interpolazione le informazioni relative a quelle aree in

cui non è presente nessuna texture. Tale approccio consente risultati molto precisi soprattutto in corrispondenza di edge, d'altro canto non è molto robusta in presenza di variazioni di luminosità ed è computazionalmente molto più pesante. Esempi di risultati di queste tecniche sono visibili nella figura 12.

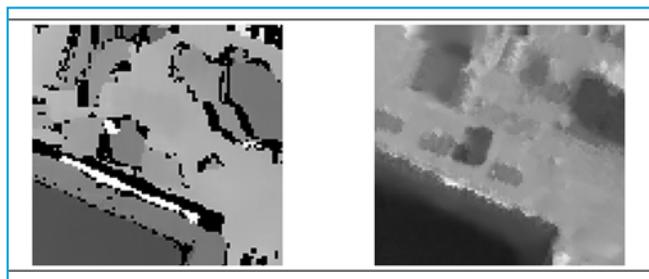


Figura 12 - Differenza tra ricostruzione per correlazione (sx) e multigrigrid (dx)

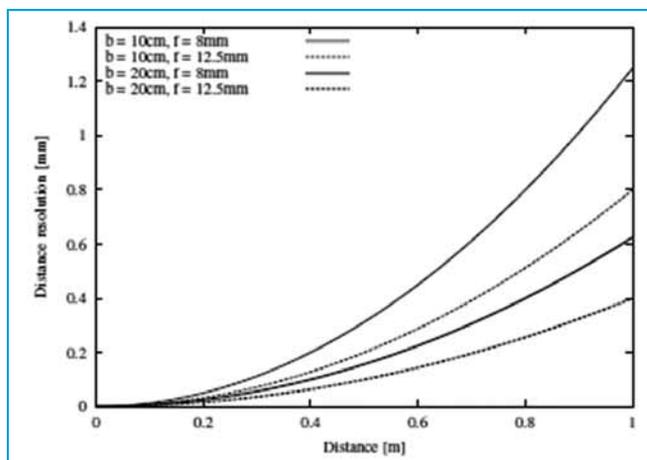


Figura 13 - Accuratezza di sistemi di visione stereoscopici

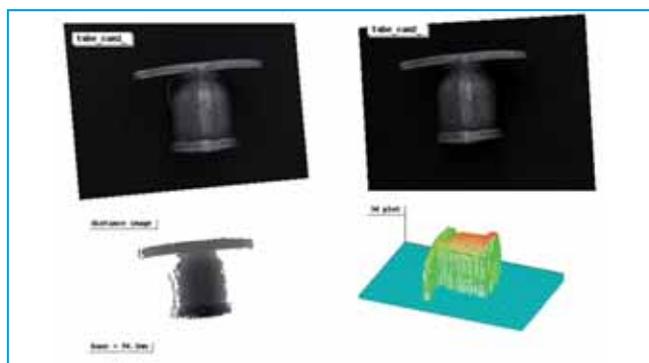


Figura 14 - Esempio di ricostruzione 3D da due immagini stereografiche

In termini di accuratezza tutto dipende dal setup del sistema, in particolar modo è inversamente proporzionale alla distanza tra le due telecamere ed alla lunghezza focale dell'ottica scelta, e direttamente proporzionale al quadrato della distanza dal piano inquadrato, un'idea è visibile nel grafico mostrato nella figura 13.