

Capitolo 1

Introduzione

1.1 La robotica e la visione

Il numero dei robot operanti nel mondo è in continua crescita, ciononostante, in genere, essi vengono esclusi dalle aree di applicazione in cui non si ha sufficiente conoscenza dell'ambiente di lavoro[1]. Questa limitazione dipende in gran parte dalla inadeguatezza dei sensori negli attuali sistemi robotici commerciali. E' stato riconosciuto che l'integrazione di sensori di tipo diverso è fondamentale per incrementare la versatilità dei robot ed il loro dominio di applicazione.

I sistemi di visione tentano di riprodurre il senso della vista dell'uomo: sono tipi di sensori vantaggiosi perchè permettono di *misurare* l'ambiente senza nessun tipo di contatto. L'aumentata velocità degli elaboratori (che permette elevati *frame rate*) e il perfezionamento delle tecniche di analisi consentono elevate prestazioni, anche in videocamere dai costi ridotti.

Tipicamente la visione e l'azionamento che ne deriva sono combinati secondo un modello open-loop *looking then moving*. L'accuratezza del risultato dell'operazione dipende direttamente dalla precisione del sistema di visione e dell'end-effector del robot. Per incrementare l'accuratezza di questi sottosistemi è necessario utilizzare, invece, un *visual-feedback control loop*. E' possibile quindi utilizzare la visione per ottenere un controllo closed-loop della posizione dell'end-effector del robot. Questo genere di approccio è detto *visual servoing*. Nel *visual servoing* si fondono i risultati ottenuti in diversi campi quali l'elaborazione di immagini ad alta velocità, la cinematica, la dinamica, la teoria del controllo e l'elaborazione real-time.

1. Introduzione

1.2 Architetture di servoing

Nel 1980, Sanderson e Weiss[1] introducono una schematizzazione dei sistemi di visual servoing, che si basa essenzialmente su due questioni:

1. La struttura di controllo è gerarchica, con il sistema di visione che fornisce un insieme di punti come ingresso al controllore dei giunti del robot, oppure il sistema di visione funge anche da controllore ed elabora direttamente gli ingressi per i giunti?
2. L'errore è definito nello spazio di lavoro del robot, cioè in coordinate 3D, oppure direttamente in termini di feature¹ sull'immagine?

La schematizzazione ha, quindi, 4 strutture principali, come descritto di seguito.

Dynamic look-and-move L'architettura di controllo è gerarchica ed usa il sistema di visione per fornire un insieme di punti al controllore ai giunti, e di conseguenza usa il feedback ai giunti per stabilizzare il robot;

Direct visual servo Il controllore ai giunti è interamente rimpiazzato dal *visual servo controller* che elabora direttamente gli ingressi per i giunti. Si usa, quindi, solo la visione per stabilizzare il sistema;

Per svariate ragioni l'architettura del primo tipo è più diffusa. In primo luogo, nel passato, il basso frame rate disponibile comportava estreme difficoltà nella realizzazione di controlli *direct visual servo* dei robot, in genere dalla dinamica non lineare. Secondo, la maggior parte dei robot ha interfacce per accettare i comandi in coordinate cartesiane: questo semplifica la costruzione dei sistemi di visual servoing e li rende più portabili. Terzo, nel *look-and-move* vengono separate le singolarità cinematiche del meccanismo[2] dal sistema di visione.

Position-based control Le feature sono estratte dall'immagine e vengono usate, conoscendo i modelli geometrici del target e della videocamera, per stimare la posizione del target rispetto alla videocamera.

Image-based control I controlli sono elaborati direttamente sulla base delle feature estratte dall'immagine.

In realtà, questi due schemi possono essere considerati come punti limite di tutta una serie di differenti possibilità, attraverso le quali le informazioni

¹Vedi par.A.4

grezze ricavate dai sensori sono gradualmente astratte in rappresentazioni sempre più strutturate usando conoscenze del modello robot-ambiente.

Con l'approccio PBVS (*position-based visual servoing*), la legge di controllo può essere sintetizzata nelle usuali coordinate di lavoro del robot, e quindi, in genere, è possibile una sintesi più semplice. Per di più, l'astrazione delle informazioni sensoriali ad un livello più alto di rappresentazione permette una più agevole integrazione di differenti sensori.

L'approccio IBVS (*image-based visual servoing*) può ridurre il ritardo computazionale, la necessità di interpretare l'immagine ed eliminare gli errori dovuti alla modellizzazione (vedi App.A) ed alla calibrazione della videocamera [3]. D'altro canto il progetto del controllore è più complesso perchè il sistema è non lineare ed altamente accoppiato.

Un'ulteriore distinzione è la seguente:

Sistemi EOL (*endpoint open-loop*) in cui il sistema di visione osserva solo il target;

Sistemi ECL (*endpoint closed-loop*) in cui il sistema di visione osserva sia il target che l'end-effector del robot.

La differenza primaria è che i sistemi EOL richiedono una esplicita calibrazione *hand-eye* quando la specifica del task viene implementata nell'algoritmo di visual servoing. Quindi, la precisione di posizionamento dei sistemi EOL dipende direttamente dall'accuratezza della calibrazione. Di converso, i sistemi ECL possono operare con un'accuratezza indipendente dalla loro calibrazione. Comunque, poichè i sistemi ECL richiedono il tracking dell'end-effector e del target, l'implementazione di controllori ECL richiede tecnologie avanzate nella computer vision.

Una questione fondamentale è la stima dei parametri da utilizzare nel feedback. Nel *position-based visual servoing* ciò è strettamente legato al problema di ricostruzione della geometria della scena ripresa da una o più videocamere. Consideriamo esclusivamente il problema della stima nella situazione di singola videocamera.

Un punto in un'immagine ripresa da una singola videocamera corrisponde ad una linea nello spazio (vedi App.A.1). E' possibile eseguire una ricostruzione geometrica della scena usando una sola videocamera in movimento (è la situazione che analizzeremo in questo lavoro di tesi), ma le equazioni che governano questo processo sono spesso mal condizionate e danno problemi di stabilità. Risultati migliori si ottengono avendo una qualche conoscenza geometrica del target. In genere è possibile stimare accuratamente il posizionamento di un oggetto conosciuto se il sistema di visione osserva delle feature appartenenti a quell'oggetto ed usa queste per effettuare la stima.

1. Introduzione

Questo approccio è stato utilizzato da Wilson[4] per il posizionamento di un end-effector a 6 DOF.

1.3 Localizzazione del robot

Formalmente, il problema della localizzazione consiste nel determinare la posizione di un robot sulla base delle informazioni sensoriali ricevute. Analiticamente può essere espresso come la valutazione della distribuzione di probabilità $p(\mathbf{l}|S)$ dove $\mathbf{l} = [x, y, z, \theta, \phi, \psi]^T$ rappresenta posizione ed orientamento di un robot nello spazio ed S le informazioni sensoriali. $p(\mathbf{l}|S)$ è la probabilità che il robot si trovi nella posizione \mathbf{l} note le informazioni sensoriali S . Accade spesso di operare in ambienti planari dove una configurazione può essere espressa come $\mathbf{l} = [x, y, \theta]$ [5] (è questo il nostro caso). Distinguiamo in metodi di posizionamento assoluti e relativi.

I metodi di posizionamento assoluti si basano sulla conoscenza della geometria del workspace o della posizione di determinati landmark[6].

I metodi di posizionamento relativi sono quelli che forniscono una stima della attuale configurazione del robot rispetto alla sua configurazione precedente. Se si prevede l'utilizzo di soli sensori di spostamento (odometria), la posizione \mathbf{l}_i al tempo t_i è determinata sulla base della posizione \mathbf{l}_{i-1} al tempo t_{i-1} e dello spostamento u_i effettuato nell'intervallo $[t_{i-1}, t_i]$. Si valuta cioè $p(\mathbf{l}|\mathbf{l}_{i-1}, u_i)$.

L'impiego esclusivo di dati di spostamento in un processo di localizzazione comporta due vincoli:

- calcolo della posizione all'istante t_i mediante un processo di integrazione degli spostamenti relativi agli istanti t_0, \dots, t_{i-1} ;
- mancanza della possibilità di percepire riferimenti assoluti.

Il fatto di utilizzare tecniche di integrazione per calcolare la posizione porta all'accumulo di eventuali errori sistematici, poichè il procedimento integra anche questi. Tale errore cresce direttamente con la lunghezza del percorso compiuto, portando rapidamente la stima della posizione ad essere scorretta. A causa della mancanza della possibilità di percepire riferimenti assoluti, inoltre risulta impossibile correggere l'errore accumulato.

Nei prossimi capitoli studieremo metodi di localizzazione relativi che si basano su informazioni visive e che non sono affetti dagli inconvenienti appena descritti.