

Capitolo 3

Basi teoriche

In questo capitolo si introdurranno i concetti teorici che stanno alla base dello sviluppo di questa tesi. Si è deciso, dove necessario, di integrare l'esposizione con accenni alle problematiche più generali riguardanti la progettazione di sistemi di "visual servoing". L'esposizione di questi concetti è fondamentale, sia per poter comprendere i risultati ottenuti in questo lavoro, sia per poter delineare il panorama in cui ci si muove.

La prima parte del capitolo introdurrà il formalismo necessario alla trattazione dei problemi di robotica (paragrafo 3.1), una volta in grado di poter rappresentare la posizione e l'orientamento di un oggetto nello spazio, ci si soffermerà sul metodo di acquisizione delle immagini tramite la telecamera digitale (paragrafo 3.2) ed infine si introdurranno alcune tecniche di elaborazione delle immagini digitali (paragrafo 3.5).

3.1 Cinematica del corpo rigido

3.1.1 Generalità

La posizione di un punto P nello spazio tridimensionale, rispetto ad un sistema di riferimento rappresentato dalla terna ortonormale e levogira $O_1 - x_1 y_1 z_1$, può essere definita attraverso un vettore \vec{P} di 3 dimensioni:

$$\vec{P} = \vec{x}_1 P_{x_1} + \vec{y}_1 P_{y_1} + \vec{z}_1 P_{z_1} \quad (3.1)$$

$\vec{x}_1, \vec{y}_1, \vec{z}_1$ sono i versori del sistema di riferimento $O_1 - x_1 y_1 z_1$ e $P_{x_1}, P_{y_1}, P_{z_1}$

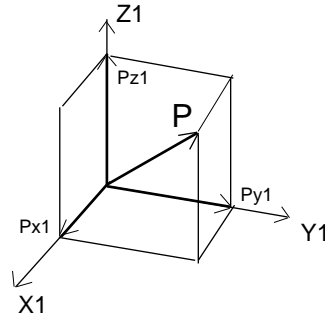


Figura 3.1: Vettore associato ad un punto P.

sono 3 scalari che indicano le proiezioni del punto \vec{P} sui rispettivi assi (vedere figura 3.1), dove con il termine proiezione si intende

$$P_{x_1} = |\vec{P}| |\vec{x}_1| \cos \alpha \quad (3.2)$$

il significato geometrico è indicato in figura 3.2 limitatamente al solo asse x.

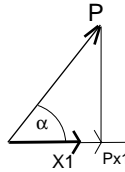


Figura 3.2: Proiezione del vettore \vec{P} sull'asse x.

Si consideri un secondo sistema di riferimento $O_0 - x_0 y_0 z_0$ (sempre costituito da una terna di vettori ortonormali e levogiri) e si rappresentino tutti e 3 i suoi versori $\vec{x}_0, \vec{y}_0, \vec{z}_0$ nella base $O_1 - x_1 y_1 z_1$ utilizzando la 3.1.

$$\vec{x}_0 = \vec{x}_1 r_{11} + \vec{y}_1 r_{12} + \vec{z}_1 r_{13} \quad (3.3)$$

$$\vec{y}_0 = \vec{x}_1 r_{21} + \vec{y}_1 r_{22} + \vec{z}_1 r_{23} \quad (3.4)$$

$$\vec{z}_0 = \vec{x}_1 r_{31} + \vec{y}_1 r_{32} + \vec{z}_1 r_{33} \quad (3.5)$$

il tutto è riscrivibile in forma matriciale

$$\begin{pmatrix} \vec{x}_0 \\ \vec{y}_0 \\ \vec{z}_0 \end{pmatrix} = \underbrace{\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}}_R \cdot \begin{pmatrix} \vec{x}_1 \\ \vec{y}_1 \\ \vec{k}_1 \end{pmatrix} \quad (3.6)$$

Notare che il termine generico r_{ij} rappresenta il prodotto scalare tra due versori, ricordando che il modulo di un versore è unitario, tramite la 3.2 si ottiene $r_{ij} = \cos\alpha_{ij}$, dove α_{ij} è detto “*coseno direttorio*”, per evidenziare ciò si può riscrivere R in un'altra forma

$$R = \begin{pmatrix} \vec{x}_1^T \cdot \vec{x}_0 & \vec{y}_1^T \cdot \vec{x}_0 & \vec{z}_1^T \cdot \vec{x}_0 \\ \vec{x}_1^T \cdot \vec{y}_0 & \vec{y}_1^T \cdot \vec{y}_0 & \vec{z}_1^T \cdot \vec{y}_0 \\ \vec{x}_1^T \cdot \vec{z}_0 & \vec{y}_1^T \cdot \vec{z}_0 & \vec{z}_1^T \cdot \vec{z}_0 \end{pmatrix} \quad (3.7)$$

La matrice R così ottenuta è una trasformazione lineare che permette di passare dalla rappresentazione di un punto P nel sistema di riferimento $O_1 - x_1 y_1 z_1$ (scritto come ${}^1\vec{P}$) alla sua rappresentazione nel secondo sistema di riferimento $O_0 - x_0 y_0 z_0$ (scritto come ${}^0\vec{P}$), cioè dal vettore ${}^1\vec{P} = \vec{x}_1 P_{x_1} + \vec{y}_1 P_{y_1} + \vec{z}_1 P_{z_1}$ al vettore ${}^0\vec{P} = \vec{x}_0 P_{x_0} + \vec{y}_0 P_{y_0} + \vec{z}_0 P_{z_0}$, per cui

$${}^0\vec{P} = {}^0R_1 \cdot {}^1\vec{P} \quad (3.8)$$

Ribadendo che 0R_1 è un operatore lineare che realizza una operazione di “*cambiamento di base*”, si può generalizzare il risultato; siano dati a questo proposito n sistemi di riferimento, si vuole allora ottenere la posizione di un punto P nella base $O_0 - x_0 y_0 z_0$, avendone la misura nella base $O_n - x_n y_n z_n$, utilizzando la 3.8 si scrivere ...

$$\begin{aligned} {}^0\vec{P} &= {}^0R_1 \cdot {}^1\vec{P} \\ {}^1\vec{P} &= {}^1R_2 \cdot {}^2\vec{P} \\ &\dots \\ {}^{n-1}\vec{P} &= {}^{n-1}R_n \cdot {}^n\vec{P} \end{aligned}$$

sostituendo opportunamente si ricava...

$${}^0\vec{P} = {}^0R_1 \cdot {}^1R_2 \dots {}^{n-1}R_n \cdot {}^n\vec{P}$$

dove $R_{tot} = {}^0R_1 \cdot {}^1R_2 \cdot \dots \cdot {}^{n-1}R_n$ è una matrice 3×3 che riassume tutti gli n cambiamenti di base. Nella letteratura la matrice R viene chiamata “*matrice di*

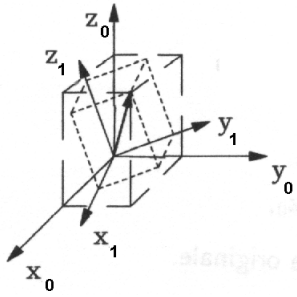


Figura 3.3: Effetto della matrice di rotazione R.

“rotazione”, questo perché se i due sistemi di riferimento $O_0 - x_0 y_0 z_0$ e $O_1 - x_1 y_1 z_1$ sono entrambi ortonormali, R rappresenta l’operazione di rotazione che applicata al sistema di riferimento $O_1 - x_1 y_1 z_1$ permette di ottenere $O_0 - x_0 y_0 z_0$, come mostrato in figura 3.3.

Esiste però un ulteriore significato cui si può attribuire alla matrice di rotazione R, riscrivendo infatti la 3.8 con una notazione diversa

$${}^0\vec{P}_1 = {}^0R_1 \cdot {}^0\vec{P}_0 \quad (3.9)$$

in questo modo ${}^0\vec{P}_1$ e ${}^0\vec{P}_0$ vengono interpretati come due vettori espressi nella medesima base $(O_0 - x_0 y_0 z_0)$ dove ${}^0\vec{P}_1$ è il risultato di una operazione di rotazione applicata a ${}^0\vec{P}_0$.

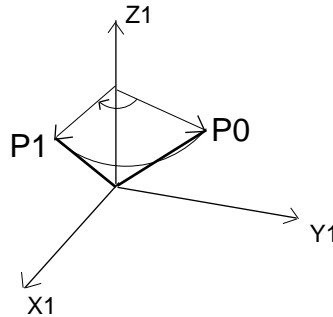


Figura 3.4: Rotazione di un vettore P.

Questa operazione è concettualmente diversa dal “cambiamento di base” di cui si è precedentemente parlato, anche se equivalente dal punto di vista matematico.

Pure in questo caso è possibile generalizzare il problema, pensando di applicare “*n rotazioni*” ad un vettore \vec{P}_0 , per cui si potrà scrivere

$$\begin{aligned} {}^0\vec{P}_1 &= R_1 \cdot {}^0\vec{P}_0 \\ {}^0\vec{P}_2 &= R_2 \cdot {}^0\vec{P}_1 \\ &\dots \\ {}^0\vec{P}_n &= R_n \cdot {}^0\vec{P}_{n-1} \end{aligned}$$

da cui discende

$${}^0\vec{P}_n = R_n \cdot \dots \cdot R_2 \cdot R_1 \cdot {}^0\vec{P}_0 \quad (3.10)$$

dove $R_{tot} = R_n \cdot \dots \cdot R_2 \cdot R_1$ è una matrice 3×3 che riassume tutte le n rotazioni applicate al vettore \vec{P}_0 per ottenere \vec{P}_n .

3.1.2 Proprietà della matrice \mathbf{R}

Si enunceranno ora le proprietà fondamentali di cui gode questa matrice, per le dimostrazioni si rimanda a [20]. Dati due scalari a, b e due vettori \vec{P} e \vec{Q} nell’ipotesi in cui i due sistemi di riferimento da cui si ricava \mathbf{R} siano **ortonormali e levogiri**, valgono le seguenti proprietà:

1. $R \cdot R^T = R^T \cdot R = I$

\mathbf{R} è una matrice “*ortonormale*”, per cui $R^{-1} = R^T$

2. $\det(R) = 1$

Proprietà non rispettata ad esempio da un operatore di riflessione.

3. $R \cdot (a\vec{P} + b\vec{Q}) = aR \cdot \vec{P} + bR \cdot \vec{Q}$

Proprietà di linearità ed omogeneità, propria degli operatori lineari.

4. $\| R \cdot P \| = \| P \|$

Invarianza del modulo di un vettore rispetto alla rotazione.

5. $\vec{P}^T \cdot \vec{Q} = (R \cdot \vec{P})^T \cdot (R \cdot \vec{Q})$

L’angolo tra due vettori rimane costante se questi sono sottoposti alla medesima rotazione.

6. $R_0 \cdot R_1 \neq R_1 \cdot R_0$

Le rotazioni (in generale) “*non godono della proprietà commutativa*”.

3.1.3 Corpo rigido

Si definir  ora la posizione e l'orientamento nello spazio di un corpo rigido "C" rispetto ad un sistema di riferimento $O_0 - x_0y_0z_0$. Un corpo rigido   un "sistema materiale in cui le distanze tra tutti i punti che lo compongono rimangono costanti nel tempo", quindi (come visto nel paragrafo 3.1) potendo associare ad ogni punto dello spazio un vettore si pu  scrivere

$$\forall \vec{P}_0, \vec{P}_1 \in C \Rightarrow \|\vec{P}_0(t) - \vec{P}_1(t)\| = Kost. \quad (3.11)$$

Sia $O_1 - x_1y_1z_1$ un sistema di riferimento solidale a C (notare che la 3.11 vale anche in questo sistema di riferimento), preso un punto $P \in C$ si pu  rappresentare la sua posizione rispetto al corpo rigido attraverso la 3.1

$${}^1\vec{P} = \vec{x}_1P_{x_1} + \vec{y}_1P_{y_1} + \vec{z}_1P_{z_1} \quad (3.12)$$

conoscendo poi la matrice R di rotazione che permette di passare da $O_1 - x_1y_1z_1$ al sistema di riferimento principale $O_0 - x_0y_0z_0$ diventa immediato ricavare la rappresentazione di P rispetto a quest'ultimo

$${}^0\vec{P} = {}^0R_1 \cdot {}^1\vec{P} + \vec{O}_1 \quad (3.13)$$

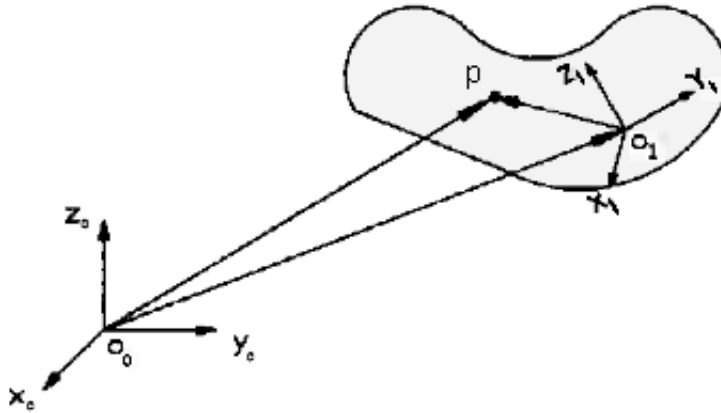


Figura 3.5: Corpo rigido.

3.1.4 Trasformazioni omogenee.

La 3.13 è una equazione che permette di ottenere la posizione di un generico punto $P \in C$ rispetto al sistema di riferimento principale, quindi può essere utilizzata per ottenere la posizione e l'orientamento del corpo rigido nello spazio tridimensionale. Il problema è che non esiste nessun operatore lineare (cioè nessuna matrice 3×3) che realizzi la 3.13 in uno spazio \mathfrak{R}^3 perché la traslazione non è una operazione lineare, infatti definendo un operatore generico che agisce su un vettore \vec{P} nella seguente maniera $T(\vec{P}) = \vec{P} + \vec{O}$. Si può allora dimostrare che la proprietà 3 del 3.1.2 non viene più rispettata, infatti

$$T(a\vec{P} + b\vec{Q}) = a\vec{P} + b\vec{Q} + \vec{O} \neq aT(\vec{P}) + bT(\vec{Q}) = a\vec{P} + b\vec{Q} + 2\vec{O}$$

per poter continuare allora ad usare la rappresentazione matriciale nel descrivere le proprietà spaziali di un corpo rigido (posizione e orientamento), è necessario introdurre “*le trasformazioni omogenee*”, le quali però richiedono la definizione di nuove grandezze.

Un punto P nello spazio rispetto ad un generico sistema di riferimento $O_m - x_m y_m z_m$ viene definito da un vettore ${}^m\vec{S}$ del tipo

$${}^m\vec{S} = \begin{pmatrix} Px \\ Py \\ Pz \\ 1 \end{pmatrix} = \begin{pmatrix} {}^m\vec{P} \\ 1 \end{pmatrix} \quad (3.14)$$

mentre un generico operatore avrà la seguente forma

$$T = \left(\begin{array}{ccc|c} r_{11} & r_{12} & r_{13} & O_x \\ r_{21} & r_{22} & r_{23} & O_y \\ r_{31} & r_{32} & r_{33} & O_z \\ \hline 0 & 0 & 0 & 1 \end{array} \right) = \left(\begin{array}{c|c} R_{3 \times 3} & O_{3 \times 1} \\ \hline 0_{1 \times 3} & 1_{1 \times 1} \end{array} \right) \quad (3.15)$$

tutto questo permette di scrivere in maniera molto compatta

$${}^0\vec{S} = T \cdot {}^1\vec{S}$$

cioè

$$\begin{pmatrix} {}^0\vec{P} \\ 1 \end{pmatrix} = \begin{pmatrix} R_{3 \times 3} & O_{3 \times 1} \\ 0_{1 \times 3} & 1_{1 \times 1} \end{pmatrix} \cdot \begin{pmatrix} {}^1\vec{P} \\ 1 \end{pmatrix}$$

che sviluppato nei suoi termini fondamentali porta a riottenere la 3.13, relazione non ottenibile, come si è visto, con matrici (3×3) .

$$\begin{cases} {}^0\vec{P} &= {}^0R_1 \cdot {}^1\vec{P} + \vec{O} \\ 1 &= 1 \end{cases}$$

Utilizzando quindi le trasformazioni omogenee, si possono rappresentare insieme sia la traslazione che la rotazione di un corpo rigido (quindi per matrici dipendenti dal tempo i moti rototraslatori) in maniera compatta, basta sostituire in T le componenti del vettore di traslazione (\vec{O}) e gli elementi della matrice di rotazione R (vedere la 3.6). Un esempio chiarirà meglio il concetto.

Esempio

Sia dato un sistema di riferimento principale $O_0 - x_0y_0z_0$ ed un corpo rigido C identificato da un secondo sistema $O_1 - x_1y_1z_1$ traslato di un vettore \vec{O}_1 e con l'asse z_1 parallelo a z_0 , come mostrato in figura 3.6

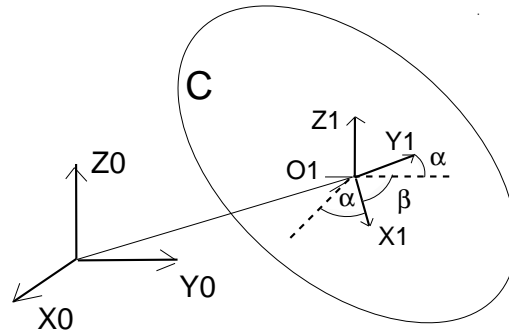


Figura 3.6: Esempio pratico di calcolo della trasformata omogenea.

Per ricavare la matrice di rotazione, basta considerare la 3.7 e calcolare i prodotti scalari tra i vari versori, risulta così

$$R = \begin{pmatrix} \cos \alpha & -\cos \beta & 0 \\ \cos \beta & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ma essendo $\alpha + \beta = \pi/2 \Rightarrow \cos \beta = \sin \alpha$ si ottiene

$$R = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.16)$$

La matrice appena ricavata rappresenta una rotazione attorno all'asse \vec{z} , per cui viene anche indicata come $R_z(\alpha)$.

Per completezza le matrici di rotazione attorno ai due rimanenti assi sono

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \quad R_y(\alpha) = \begin{pmatrix} \cos \alpha & 0 & -\sin \alpha \\ 0 & 1 & 0 \\ \sin \alpha & 0 & \cos \alpha \end{pmatrix} \quad (3.17)$$

Concludendo l'esempio, la trasformazione omogenea che rappresenta la figura 3.6 è

$$T = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 & O_{1x} \\ \sin \alpha & \cos \alpha & 0 & O_{1y} \\ 0 & 0 & 1 & O_{1z} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.18)$$

3.1.5 Teoria della robotica

Come fatto per la matrice R , è possibile comporre una serie di n trasformazioni omogenee rappresentanti, o n *cambiamenti di sistemi di riferimento rototraslati l'uno rispetto l'altro*, o *l'applicazione di n rototraslazioni ad un vettore \vec{P} in un unico sistema di riferimento*.

Nel primo caso si ottiene

$${}^0\vec{P} = {}^0A_1 \cdot {}^1A_2 \dots {}^{n-1}A_n \cdot {}^n\vec{P} = {}^0T_n \cdot {}^n\vec{P} \quad (3.19)$$

mentre nel secondo caso

$${}^0\vec{P}_n = A_n \cdot \dots \cdot A_2 \cdot A_1 \cdot {}^0\vec{P}_0 = T_{tot} \cdot {}^0\vec{P}_0 \quad (3.20)$$

Una applicazione immediata della 3.19 è rappresentata dal manipolatore robotico

Questo costituisce una catena cinematica aperta, formata da $n+1$ corpi (detti "*link*") connessi da n giunti, dove il primo corpo è la base del manipolatore e l'ultimo costituisce l'organo terminale (end-effector). In figura 3.7 troviamo

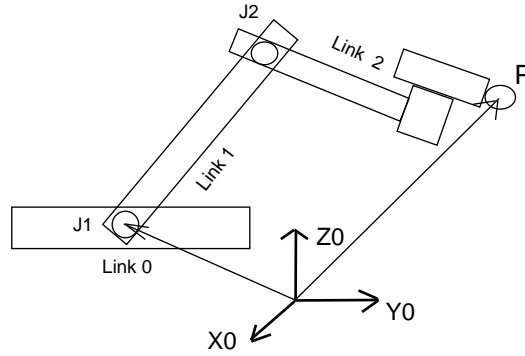


Figura 3.7: Manipolatore a 2 gradi di libertà.

l'esempio di un manipolatore costituito da 3 corpi e 2 giunti. è possibile quindi, come visto nel paragrafo 3.1.4, associare ad ogni corpo della catena cinematica una matrice di forma omogenea A , rappresentante la posizione e l'orientamento del corpo stesso rispetto al link precedente. Nell'esempio riportato, definendo per ogni corpo un sistema di riferimento, si ottiene

- A_0 = Posizione ed orientamento del corpo 0 (base del manipolatore) rispetto al sistema di riferimento $O_0 - x_0y_0z_0$.
- A_1 = Posizione ed orientamento del corpo 1 rispetto alla base del manipolatore.
- A_2 = Posizione ed orientamento dell'organo terminale ($link_2$) rispetto al corpo 1 del manipolatore.

Per cui le coordinate del punto \vec{P} solidale con il sistema di riferimento dell'organo terminale saranno rappresentabili nel sistema di riferimento $O_0 - x_0y_0z_0$ tramite

$${}^0\vec{P} = T \cdot {}^2\vec{P} \quad \text{dove} \quad T = A_0A_1A_2 \quad (3.21)$$

che corrisponde appunto alla 3.19.

Convenzione di Denavit-Hartenberg

La convenzione di Denavit-Hartenberg è un procedimento sistematico per il calcolo della matrice omogenea A fra coppie di bracci (link). Si assegna prima di

tutto ad ogni *link*_{*i*} della catena cinematica un sistema di riferimento, si applica poi un algoritmo (vedi [15] e [20]) che porta a determinare i parametri che caratterizzano la singola matrice ${}^{i-1}A_i$, procedendo in modo iterativo si arriva a ricavare la T della 3.19

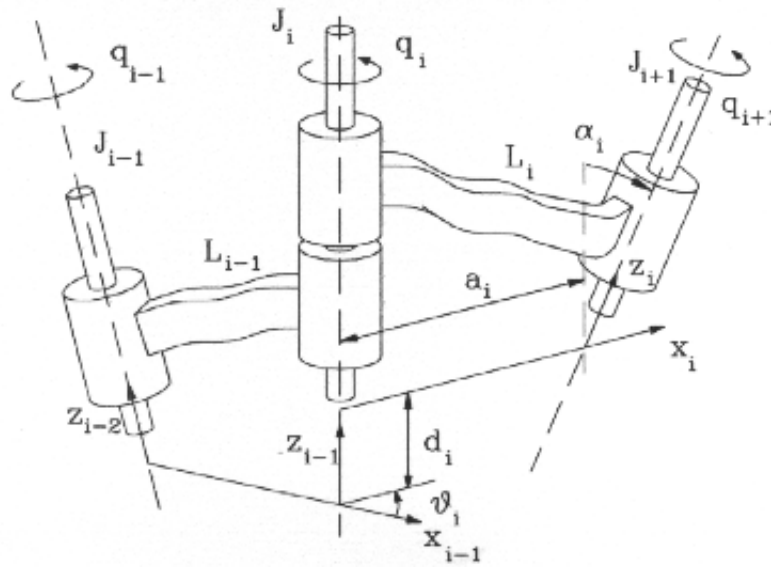


Figura 3.8: Parametri utilizzati nella notazione Danavit-Hartenberg

Osservando la figura 3.8 si possono notare i 4 parametri che caratterizzano gli elementi della matrice ${}^{i-1}A_i$

1. a_i = lunghezza della normale comune tra i due assi consecutivi dei giunti z_{i-1} e z_i .
2. α_i = identifica l'angolo (detto *torsione*) in senso antiorario tra gli assi z_{i-1} (riferito al giunto i) e z_i (riferito al giunto $i + 1$).
3. d_i = distanza tra l'origine o_{i-1} e l'intersezione con l'asse z_{i-1} della normale comune riferita all'asse del giunto precedente (cioè z_{i-2}), d_i è anche chiamata *scorrimento*.
4. θ_i = angolo in senso antiorario tra l'asse x_{i-1} e l'asse x_i , detto anche angolo di *torsione*. Notare che il generico asse x_j giace sulla normale comune a_j .

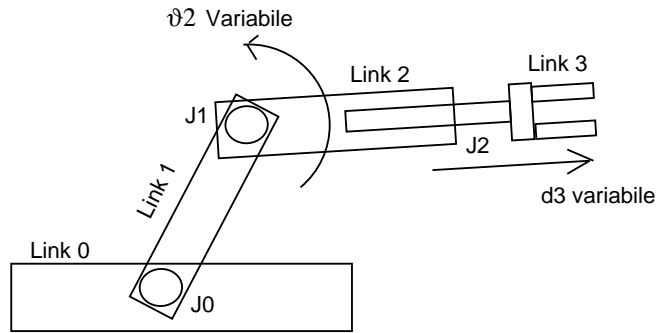


Figura 3.9: Esempio di giunti prismatici e rotoidali.

I parametri a_i e α_i sono costanti perché dipendono dalla geometria del singolo link. Se il giunto è *prismatico* poiché il suo asse coincide con l'asse z_{i-1} varierà lo scorrimento d_i , mentre θ_i rimarrà costante. Se invece il giunto è *rotoidale* poiché la rotazione intorno all'asse z_{i-1} è descritta dall'angolo fra x_i ed x_{i-1} sarà allora θ_i ad essere variabile mentre d_i costante.

3.1.6 Cinematica diretta e inversa

Per completare la panoramica sulla teoria della robotica introduciamo due problemi:

Problema cinematico diretto

Si vuole passare dalla conoscenza delle posizioni, velocità, accelerazioni dei vari giunti che costituiscono la catena cinematica di un manipolatore alle rispettive grandezze di un generico punto \vec{P} dell'organo terminale del manipolatore. Per fare ciò è sufficiente applicare la metodologia delle trasformazioni omogenee appena descritta, estendendo i discorsi alle velocità e alle accelerazioni [20].

Problema cinematico inverso

Si vuole passare dalle grandezze cinematiche di un punto \vec{P} (P , \dot{P} , \ddot{P}) alle rispettive grandezze dei vari giunti che costituiscono la catena cinematica aperta del manipolatore. Per questo secondo problema *non esiste nessuna tecnica di carattere generale che applicata dia una soluzione*, si utilizzano allora due tipi di

approcci, il primo *numerico*, il secondo *analitico* che porta in genere ai risultati migliori [20].

3.2 Modello della telecamera

¹ La telecamera è un dispositivo passivo in grado di riprodurre un'immagine bidimensionale della porzione di spazio compresa nel suo campo visivo. Il meccanismo di acquisizione è del tutto simile a quello presente in natura nell'occhio umano ed è illustrato in figura 3.10, dove il punto F indica il centro ottico della lente (chiamata “*crystallino*”). La trasformazione che porta alla generazione dell'immagine viene chiamata “*trasformazione prospettica*”, è una relazione di tipo non lineare che lega un punto \vec{P} di coordinate (X, Y, Z) nello spazio ad un punto \vec{p} di coordinate (x, y) sul piano visivo della telecamera (nel caso dell'occhio il piano visivo coincide con la retina).

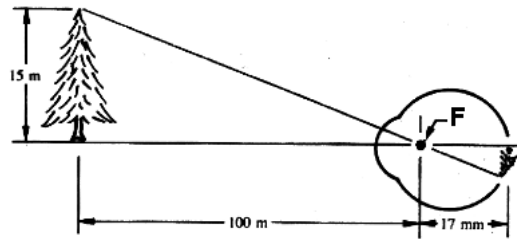


Figura 3.10: Schema di acquisizione di un'immagine tramite l'occhio.

Nella figura 3.11 è rappresentato il *piano visivo 1* che schematizza il meccanismo geometrico presente nell'occhio umano² (figura 3.10), in aggiunta è stato evidenziato un secondo piano visivo (*piano visivo 2*) che sarà invece quello utilizzato in questa esposizione, le differenze tra i due consistono solo nella rappresentazione matematica ma i risultati che ne derivano sono analoghi.

Si fissi un sistema di riferimento $O_c - x_c y_c z_c$ con l'origine $O_c \equiv F$, sia chiamata f (*distanza focale*) la distanza tra l'origine F ed il piano visivo. Attraverso l'utilizzo di semplici proporzioni è possibile definire le coordinate del punto $p(x, y)$ in funzione delle coordinate spaziali del punto $P(X, Y, Z)$

¹vedere [9] [10]

²Chiamato in letteratura “*Pin-Hole Camera*”

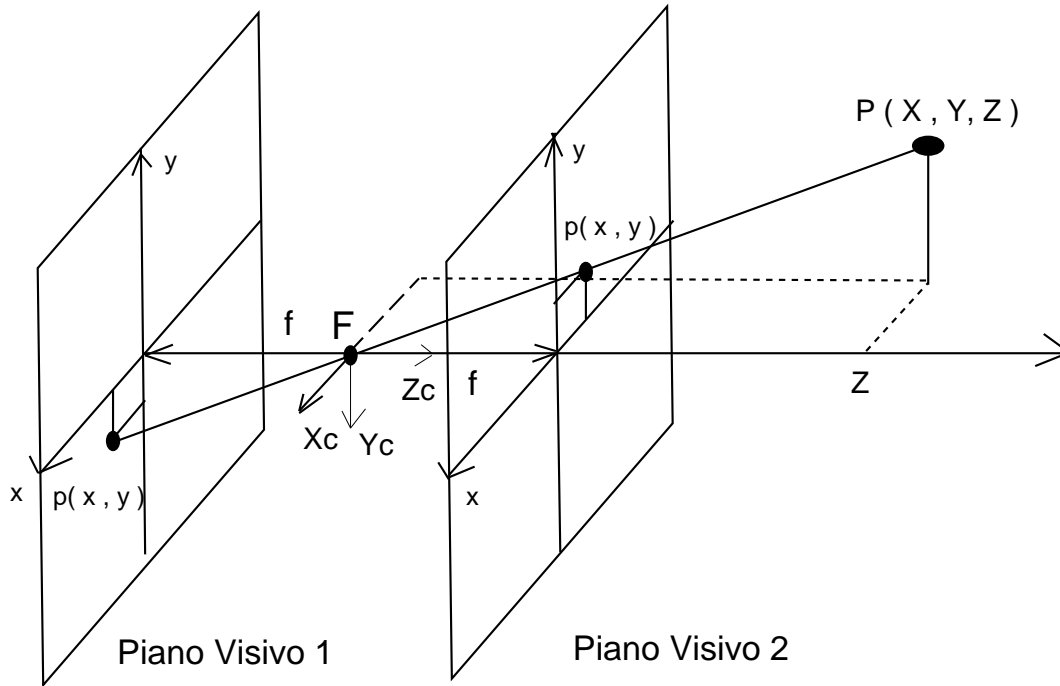


Figura 3.11: Relazione geometrica tra un punto reale ed un punto dell'immagine.

$$\frac{x}{f} = \frac{X}{Z} \quad \frac{y}{f} = \frac{Y}{Z} \quad (3.22)$$

Questa conversione è di tipo non lineare perché si introduce una divisione per Z , sarebbe allora conveniente poter trovare un metodo per rappresentare la 3.22 in forma lineare.

3.2.1 Coordinate omogenee

Esprimiamo un punto \mathbf{p} ($\mathbf{p} = (x, y)$) appartenente ad un piano come funzione di 3 coordinate cioè $\tilde{\mathbf{P}} = (x_1, x_2, x_3)$ in modo tale che la relazione che lega le due rappresentazioni sia data da

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x_1}{x_3} \\ \frac{x_2}{x_3} \end{pmatrix} \quad (3.23)$$

se ora si moltiplicasse $\tilde{\mathbf{P}}$ per un generico numero non nullo λ ottenendo quindi $\lambda\tilde{\mathbf{P}} = (\lambda x_1, \lambda x_2, \lambda x_3)$ questo non influenzerebbe il valore di \mathbf{p} sensibile solo al rapporto degli elementi di $\tilde{\mathbf{P}}$. Nel caso in cui x_3 fosse nullo la 3.23 non potrebbe

essere definita avendo tutti i suoi termini infiniti, mentre riscrivendo il punto \mathbf{p} tramite $\tilde{\mathbf{P}}$ si avrebbe una rappresentazione del tipo $(x_1, x_2, 0)$.

La rappresentazione a 3 variabili del punto \mathbf{p} è detta *rappresentazione in coordinate omogenee*, mentre la 3.23 realizza il passaggio tra le due rappresentazioni.

L'aggettivo "omogeneo" deriva dal fatto che data un'equazione, funzione delle coordinate (x, y) , si può dedurre tramite la 3.23 un'equazione omogenea funzione di (x_1, x_2, x_3) . Si prenda come esempio una relazione che individua una retta nel piano x, y

$$a_1x + a_2y + a_3 = 0$$

questa può venire rappresentata in maniera omogenea, infatti utilizzando la 3.23

$$a_1 \frac{x_1}{x_3} + a_2 \frac{x_2}{x_3} + a_3 = 0$$

e moltiplicando per x_3 si ottiene la forma omogenea.

$$a_1x_1 + a_2x_2 + a_3x_3 = 0$$

Per poter apprezzare fino in fondo la duttilità di questa rappresentazione bisogna ritornare al caso tridimensionale descritto nella figura 3.11, possiamo allora rappresentare i punti $p(x, y)$ e $P(X, Y, Z)$ in coordinate omogenee. Definiti due scalari λ, s non nulli si ottiene

$$\begin{aligned} (X, Y, Z) &\rightarrow (\lambda X, \lambda Y, \lambda Z, \lambda) \\ (x, y) &\rightarrow (sx, sy, s) \end{aligned}$$

fatto ciò si può scrivere la 3.22 utilizzando le coordinate omogenee

$$\begin{pmatrix} sx \\ sy \\ s \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \lambda X \\ \lambda Y \\ \lambda Z \\ \lambda \end{pmatrix} \quad (3.24)$$

riscrivibile in forma compatta

$$\tilde{\mathbf{p}} = P_p \cdot \tilde{\mathbf{P}} \quad (3.25)$$

dove P_p è detta *matrice prospettica*

$$P_p = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (3.26)$$

risolvendo il prodotto e riconvertendo il punto $p(x, y)$ da $(sx, sy, s) \rightarrow (x, y)$ tramite la 3.23, si ottiene come anticipato proprio la 3.22, per cui si è riusciti a rappresentare un'operazione non lineare come la *trasformazione prospettica* con una applicazione lineare³.

3.2.2 Generalizzazione

Un altro vantaggio della rappresentazione in coordinate omogenee è portato dal fatto che è possibile trattare assieme (quindi in maniera uniforme) trasformazioni omogenee applicate a corpi rigidi come visto nel paragrafo 3.1.4 e trasformazioni prospettiche. Si definisca a questo proposito un sistema di riferimento principale $O_0 - x_0y_0z_0$ ed un secondo sistema di riferimento $O_c - x_cy_cz_c$ solidale con la telecamera (vedi figura 3.12). Se si conosce la matrice di rototraslazione ${}^cT_0(4 \times 4)$ che permette di passare da $O_0 - x_0y_0z_0$ a $O_c - x_cy_cz_c$ ed è nota la posizione del punto \vec{P} nel sistema di riferimento principale, allora è possibile attraverso la 3.19 passare alla rappresentazione del punto \vec{P} nella base della telecamera (cP), per cui si potrà scrivere

$${}^c\vec{P} = {}^cT_0 \cdot {}^0\vec{P} \quad (3.27)$$

ma in questo modo ci si è ricondotti alla situazione schematizzata in figura 3.11, per cui applicando la 3.25 ricaviamo la proiezione (\vec{P}_i) del punto \vec{P} sul piano visivo della telecamera

$$\vec{P}_i = P_p \cdot {}^c\vec{P} \quad (3.28)$$

unendo la 3.28 con la 3.27 si ricava

$$\vec{P}_i = P_p {}^cT_0 \cdot {}^0\vec{P} \quad (3.29)$$

dove chiamando

³notare che in genere l'utilizzo della 3.24 è indipendente dal valore assegnato a λ per cui si pone $\lambda = 1$.

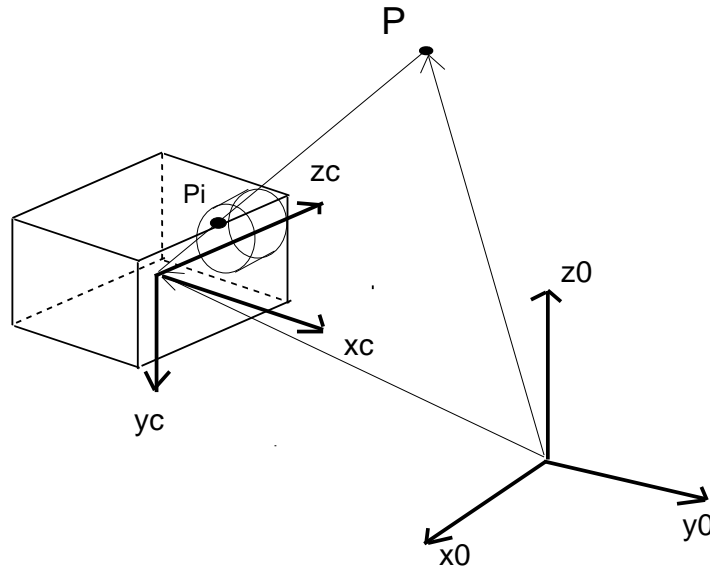


Figura 3.12: Esempio di posizionamento telecamera.

$$P_{tot} = P_p {}^c T_0 \quad (3.30)$$

si può scrivere la 3.29 come

$$\vec{P}_i = P_{tot} \cdot {}^0 \vec{P} \quad (3.31)$$

cioè abbiamo ottenuto l'immagine del punto ${}^0 \vec{P}$ ripresa dalla telecamera, attraverso un semplice prodotto per la matrice P_{tot} 3×4 .

3.2.3 Matrice CCD

Per apportare una trattazione completa, bisogna tenere conto del fatto che P_i nella realtà concreta non è un punto su un piano cartesiano, ma è costituito da un singolo elemento di una griglia di pixel, per poter considerare anche questo aspetto si introduce una matrice aggiuntiva nella 3.30 ottenendo

$$P_{tot} = P_{ccd} P_p {}^c T_0 \quad (3.32)$$

Preso uno scalare non nullo r si può indicare un punto $P_i(u, v)$ nella griglia dei pixel in forma omogenea

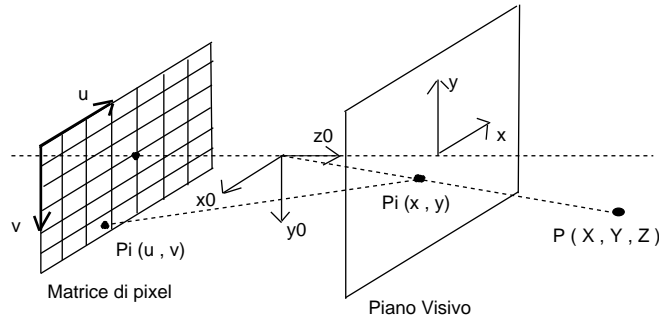


Figura 3.13: Relazione tra piano visivo e griglia dei pixel.

$$\begin{pmatrix} ru \\ rv \\ r \end{pmatrix} = \begin{pmatrix} K_u & 0 & u_0 \\ 0 & K_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} sx \\ sy \\ s \end{pmatrix} \quad (3.33)$$

Questa è un'equazione in coordinate omogenee che determina una mappatura tra il piano visivo ed un piano costituito dalla matrice di pixel in cui l'immagine viene memorizzata, dove

$$P_{ccd} = \begin{pmatrix} K_u & 0 & u_0 \\ 0 & K_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.34)$$

Tramite P_{ccd} si considera la forma non ideale del pixel (termini K_u e K_v) e la non coincidenza tra l'origine del piano visivo xy e l'origine della griglia di pixel uv (termini u_0 e v_0) (vedere a questo proposito [10] [13]). Una volta ricavato \vec{P}_i basta applicare la 3.23 per abbandonare la rappresentazione omogenea e passare a quella cartesiana ottenendo le coordinate (x,y) della proiezione del punto sul piano visivo⁴.

3.3 Calibrazione

Sostituendo quindi la 3.32 nella 3.31 si ha una immediata relazione che permette dato un punto \vec{P} di ottenerne l'immagine acquisita dalla telecamera, è naturale

⁴Piú precisamente se si è utilizzato la 3.34 si ottengono le coordinate del punto nella matrice dei pixel (u, v) . In seguito si utilizzerá il termine *piano visivo* anche riferendosi alla griglia dei pixel (u,v) .

che affinché ciò possa realizzarsi bisogna conoscere tutti i parametri che caratterizzano P_{tot} che come si è visto sono (f, k_u, k_v, u_0, v_0) , detti *parametri intrinseci della telecamera*, mentre quelli *estrinseci* sono rappresentati dalla posizione e dall'orientamento della telecamera rispetto al sistema di riferimento principale (matrice cT_0)⁵.

In molte applicazioni alcuni di questi parametri sono sconosciuti, molto spesso sono la posizione e l'orientamento della telecamera, in questo caso è richiesta la risoluzione di un problema detto di “*calibrazione della telecamera*”⁶. Per ottenerne la soluzione si possono seguire diversi approcci:⁷

1. Considerare una serie sufficiente ampia di punti.
2. Considerare più immagini successive di stessi punti, con la telecamera in movimento.
3. Utilizzare riprese contemporanee di diversi punti (vista stereoscopica).

Prendendo in considerazione l'approccio 1, si procederà utilizzando delle immagini di punti (u_i, v_i) di cui si conosce l'esatta posizione nello spazio tridimensionale (coordinate X_i, Y_i, Z_i). Partendo allora dalla 3.31 e riscrivendola nella forma

$$\begin{pmatrix} ru \\ rv \\ r \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \cdot \begin{pmatrix} \lambda X \\ \lambda Y \\ \lambda Z \\ \lambda \end{pmatrix} \quad (3.35)$$

Si vuole ricavare quante coppie di vettori (ru, rv, r) e $(\lambda X, \lambda Y, \lambda Z, \lambda)$ servono per poter determinare tutti gli elementi P_{ij} della matrice P_{tot} .

P_{tot} contiene 12 elementi incogniti, risolvendo il prodotto 3.35 e ricavando u e v tramite la 3.23, si ottengono

$$\begin{cases} u = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}} \\ v = \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}} \end{cases} \quad (3.36)$$

⁵D'ora in avanti si intenderà con il termine “*locazione*”, la posizione e l'orientamento di oggetto nello spazio

⁶In letteratura “*External Camera Calibration*”, [13] Cap.1 .

⁷[19] pag. 17, [13] Cap. 2.

Le equazioni della 3.36 rappresentano due vincoli, affinché quindi il sistema abbia una sola soluzione bastano 6 punti (12 equazioni in 12 incognite, sistema che può essere risolto attraverso il *metodo dei minimi quadrati*⁸). Sono stati sviluppati metodi che permettono di ottenere la calibrazione della telecamera con un numero di punti $n < 6$, il problema è che salvo casi particolari bisogna risolvere delle equazioni polinomiali, ed in generale la soluzione non è unica ([13] e [14]). Quest'ultima osservazione appare evidente nel caso in cui si provi ad ottenere la matrice P_{tot} solo attraverso la conoscenza di 3 punti (questo problema è chiamato P3P⁹), osservando la figura 3.14, si può notare che esistono infinite combinazioni di 3 punti che generano la stessa immagine acquisita dalla telecamera, questa relazione *molti ad uno* è indicativa della non linearità della trasformazione prospettica. Utilizzando questi metodi si riescono a selezionare le possibili soluzioni compatibili con l'effettiva distribuzione spaziale dei tre punti, ma in generale le soluzioni trovate sono sempre più di una, solo per $n \geq 6$ (PnP) si riesce a garantirne l'unicità [13] come anticipato in questo paragrafo.

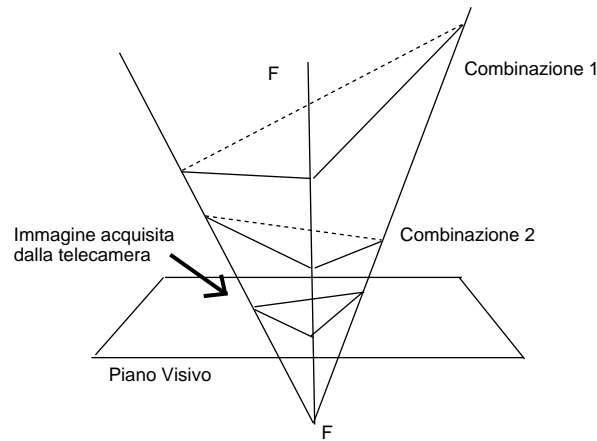


Figura 3.14: Esempio di due combinazioni in un P3P.

Per finire si risottolinea che una volta ricavato P_{tot} si ottengono i parametri che caratterizzano la telecamera, in particolar modo avendo ottenuto quelli *estrinseci* (cioè la matrice cT_0 della 3.27), si è riusciti a dedurre dalla immagi-

⁸Vedere [9, 10, 13].

⁹“*Perspective 3 point*”

ne acquisita dalla telecamera informazioni sulla posizione ed orientamento nello spazio della telecamera stessa.

3.3.1 Stima della posizione di un oggetto.

Si ipotizzi di conoscere tutti i parametri (sia estrinseci che intrinseci) di una telecamera, per cui anche la sua *locazione* rispetto al sistema di riferimento principale. Sotto queste ipotesi si vuole ricostruire la posizione di un oggetto inquadrato dalla telecamera, ipotizzando che questo sia identificato da un insieme di N punti¹⁰. Questo è un “*problema di stima della posizione*”¹¹ ed è strettamente legato al problema della calibrazione di cui si è parlato nel paragrafo precedente. Per la soluzione si seguono principalmente due tipi di approccio

1. Riprendere gli N punti con una singola telecamera in movimento.
2. Riprendere gli N punti utilizzando 2 telecamere (vista stereoscopica).

Per entrambi gli approcci troviamo principalmente 2 tipi di soluzioni, quelle *analitiche* e quelle *ai minimi quadrati*. Soluzioni analitiche sono state ricavate nella stima della posizione di un oggetto identificato da 3, 4 e 5 punti, con 6 o più punti, come nel caso della calibrazione, si trova sempre un’unica soluzione.

Metodo dei minimi quadrati

Le soluzioni derivate dal *metodo dei minimi quadrati* si basano sulla definizione di una funzione errore ϵ che deve essere minimizzata¹² agendo sui parametri che caratterizzano la *locazione* dell’insieme dei punti rispetto alla telecamera.

Siano P_1, \dots, P_N gli N punti nello spazio tridimensionale osservati dalla telecamera, mentre P_{i1}, \dots, P_{iN} le loro proiezioni sul piano visivo e siano poi definiti N *pesi* ω_i di modo che $\sum_1^N \omega_i = 1$. Chiamando per finire la trasformazione prospettica π , si può allora scrivere

$$\epsilon^2 = \sum_{n=1}^N \omega_n \|\vec{p}_{i_n} - \pi(R\vec{P}_n + \vec{t})\|^2 \quad (3.37)$$

¹⁰Questi punti potrebbero ad esempio rappresentare particolari proprietà dell’oggetto stesso, si veda a questo proposito il paragrafo 3.5.

¹¹In letteratura chiamato “*Pose Estimation*” [11, 19].

¹²Vedere [19]

la matrice di rotazione R ed il vettore di traslazione \vec{t} che minimizzano la 3.37 indicano la *locazione* dell'oggetto (cioé del sistema formato dagli N punti) rispetto alla telecamera, conoscendo per ipotesi i parametri estrinseci della telecamera, ecco allora ricavata la posizione dell'oggetto rispetto al sistema di riferimento principale e quindi risolto il problema della stima della posizione¹³.

3.4 Piano di lavoro

Si analizzerà ora una situazione particolare in cui la calibrazione della telecamera e la successiva stima della posizione spaziale di alcuni punti, risulta assai semplificata. Affinché ciò possa realizzarsi bisogna che tutti i punti giacciono su un piano che chiameremo *piano di lavoro* e questo sia parallelo al piano visivo, in modo che ogni punto abbia la medesima componente Z rispetto la sistema di riferimento della telecamera $O_c - x_c y_c z_c$, in questa circostanza allora i termini della 3.36

$$\begin{cases} p_{13}Z + p_{14} \\ p_{33}Z + p_{34} \\ p_{23}Z + p_{24} \\ p_{33}Z + p_{34} \end{cases} \quad (3.38)$$

sono tutti costanti e quindi la matrice P_{tot} diviene ridondante, infatti per le ipotesi fatte dovendo essere la somma dei termini $p_{i3}Z + p_{i4}$ costante, si ha che uno dei gradi di libertà (p_{i3}, p_{i4}) può essere eliminato. Si può allora riscrivere la 3.35 nella forma¹⁴

$$\begin{pmatrix} ru \\ rv \\ r \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (3.39)$$

La 3.39 è una trasformazione molto semplice e veloce da calcolare¹⁵, utilizzata in tutte quelle applicazioni in cui la profondità degli oggetti inquadrati è molto piccola se rapportata alla distanza della telecamera dagli oggetti, un esempio di utilizzo di una tale trasformazione prospettica è in cartografia¹⁶

¹³Vedere per una descrizione più dettagliata [11]

¹⁴Ricordando che in generale $\lambda = 1$

¹⁵Chiamata *Affine projection*, vedi [10] pag.38, [13], [19].

¹⁶[19] pag. 8, [13] pag. 21.

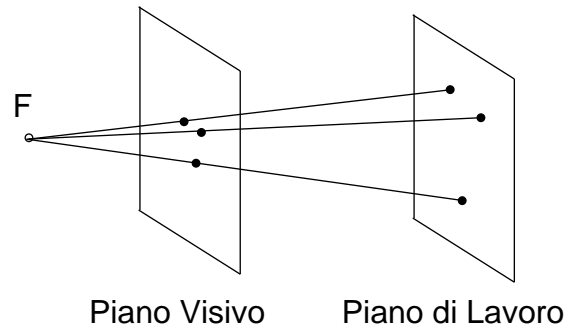


Figura 3.15: Piano visivo parallelo al piano di lavoro.

3.4.1 Matrice di calibrazione

Sempre nelle ipotesi in cui il piano visivo sia parallelo al piano di lavoro, ci si porrà ora il problema inverso, cioè si vuole determinare la matrice K che permette di ottenere un punto nello spazio tridimensionale conoscendone solamente l'immagine $P(u, v)$ ¹⁷, cioè

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \cdot \begin{pmatrix} ru \\ rv \\ r \end{pmatrix} \quad (3.40)$$

Avendo a disposizione le coordinate di soli tre punti immagine (rispetto al sistema di riferimento principale $O_0 - x_0y_0z_0$) è possibile calibrare immediatamente la matrice K (vedere figura 3.15), infatti

$$\begin{pmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \cdot \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ 1 & 1 & 1 \end{pmatrix} \quad (3.41)$$

si ottiene subito che

$$\begin{pmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{pmatrix} \cdot \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ 1 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \quad (3.42)$$

¹⁷confrontando la 3.40 con la 3.39 si ricava immediatamente che K è l'inversa sinistra di P_{tot} .

l'unico vincolo è che la matrice formata dai punti immagine sia invertibile. Una condizione necessaria e sufficiente affinché ciò sia vero è che i vettori rappresentanti i punti immagine $(u_1, v_1, 1)$, $(u_2, v_2, 1)$, $(u_3, v_3, 1)$ siano linearmente indipendenti¹⁸.

Stima della posizione

Una volta ricavato \mathbf{K} (cioè a calibrazione avvenuta), utilizzando la 3.40 è possibile ricavare per ogni punto presente nell'immagine la corrispondente posizione sul piano di lavoro cioè si effettua una “*stima di posizione*”.

Generalizzando, dati m punti ripresi dalla telecamera, è possibile utilizzare la seguente relazione

$$\underbrace{\mathbf{E}}_{3 \times m} = \underbrace{\mathbf{K}}_{3 \times 3} \cdot \underbrace{\mathbf{C}}_{3 \times m} \quad (3.43)$$

per ricavare con un semplice prodotto matriciale le posizioni degli m punti nello spazio tridimensionale, dove

$$\mathbf{E} = \begin{pmatrix} X_1 & X_2 & \dots & X_{m-1} & X_m \\ Y_1 & Y_2 & \dots & Y_{m-1} & Y_m \\ Z_1 & Z_2 & \dots & Z_{m-1} & Z_m \end{pmatrix} \quad (3.44)$$

e

$$\mathbf{C} = \begin{pmatrix} u_1 & u_2 & \dots & u_{m-1} & u_m \\ v_1 & v_2 & \dots & v_{m-1} & v_m \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} \quad (3.45)$$

rappresentano rispettivamente la matrice degli m punti sul piano di lavoro e la matrice degli stessi m punti proiettati sul piano visivo.

3.4.2 Matrice di calibrazione (secondo caso)

Come già anticipato a pagina 35 punto 2, è possibile arrivare alla calibrazione della telecamera anche attraverso più immagini di uno stesso oggetto. In questo caso, conoscendo la posizione assoluta di questo, è possibile ottenere la *locazione*

¹⁸Per un esempio concreto dell'utilizzo di tale metodo nella calibrazione della telecamera vedere il codice dell'applicazione *ad anello aperto* sviluppata in questa tesi.

assoluta della telecamera, mentre non avendo alcun dato sulla sua posizione si riesce comunque ad ottenere la locazione relativa della telecamera rispetto all'oggetto.

Di seguito si vedrà il metodo utilizzato in questa tesi nell'applicazione ad *anello chiuso* per ricavare la matrice K della 3.40

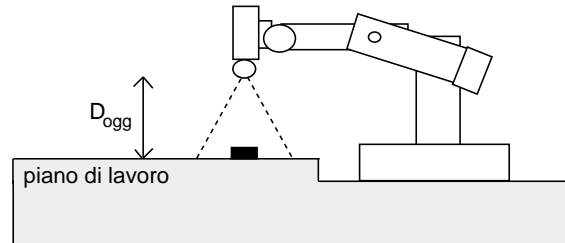


Figura 3.16: Situazione tipica al momento della calibrazione.

Nella ipotesi in cui il piano di lavoro sia parallelo al piano visivo, si consideri di operare con un manipolatore robotico a 6 gradi di libertà¹⁹ e di trovarsi nella situazione raffigurata in figura 3.16 ed in più di non conoscere la distanza D_{ogg} che intercorre tra la telecamera montata sulla flangia terminale del robot e l'oggetto inquadrato.

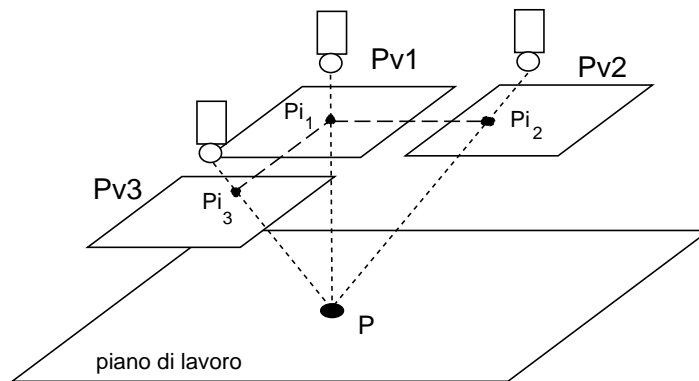


Figura 3.17: Calibrazione tramite successive immagini di un oggetto fermo.

Per calibrare la telecamera basta muoverla in tre posizioni successive (Pv_1 , Pv_2 , Pv_3) ottenute tramite spostamenti di valore noto ed espresse secondo il sistema di

¹⁹Condizione sufficiente per essere sicuri che i movimenti di calibrazione successivamente descritti siano fisicamente realizzabili dal manipolatore.

riferimento dell'end-effector²⁰. Considerando poi le tre immagini successive dell'oggetto $(P_{i_1}, P_{i_2}, P_{i_3})$ e scegliendo una particolare posizione (e la rispettiva immagine) come riferimento, ad esempio la coppia (Pv_1, P_{i_1}) si può allora scrivere

$$\mathbf{E} = \mathbf{K} \cdot \mathbf{C} \quad (3.46)$$

dove

$$\mathbf{E} = \begin{pmatrix} Pv_{3x} - Pv_{1x} & Pv_{2x} - Pv_{1x} & Pv_{1x} - Pv_{1x} \\ Pv_{3y} - Pv_{1y} & Pv_{2y} - Pv_{1y} & Pv_{1y} - Pv_{1y} \\ 1 & 1 & 1 \end{pmatrix} \quad (3.47)$$

$$\mathbf{C} = \begin{pmatrix} Pi_{3x} - Pi_{1x} & Pi_{2x} - Pi_{1x} & Pi_{1x} - Pi_{1x} \\ Pi_{3y} - Pi_{1y} & Pi_{2y} - Pi_{1y} & Pi_{1y} - Pi_{1y} \\ 1 & 1 & 1 \end{pmatrix} \quad (3.48)$$

Ecco allora che se \mathbf{C} è invertibile²¹ si ricava immediatamente la matrice \mathbf{K}

$$\mathbf{K} = \mathbf{E} \cdot \mathbf{C}^{-1}$$

Tale matrice rappresenta il legame che intercorre tra lo spazio del robot e lo spazio dell'immagine e da questa è possibile dedurre la posizione ed orientamento della telecamera rispetto alla flangia terminale del robot. Conoscendo (se necessario) la locazione dell'end-effector rispetto al sistema di riferimento principale $O_0 - x_0y_0z_0$ è dunque possibile ricavare anche la locazione assoluta della telecamera.

Notare che la matrice \mathbf{K} può essere espressa come

$$K = \left(\begin{array}{c|c} \mathbf{B}_{2 \times 2} & \mathbf{0}_{2 \times 1} \\ \hline \mathbf{1}_{1 \times 2} & \mathbf{1}_{1 \times 1} \end{array} \right) \quad (3.49)$$

dove

$$\mathbf{B} = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \quad (3.50)$$

²⁰Vedere le coordinate Tool del Puma 260 paragrafo 2.6 ed il paragrafo 3.1.5.

²¹Per avere l'invertibilità basta imporre spostamenti dell'end-effector perpendicolari tra loro.

È la matrice \mathbf{B} che racchiude la vera e propria trasformazione dal piano visivo al piano di lavoro ed è proprio a questa che da ora in poi si farà riferimento quando si tratterà la trasformazione del piano visivo al piano di lavoro nell'applicazione ad *anello chiuso*.

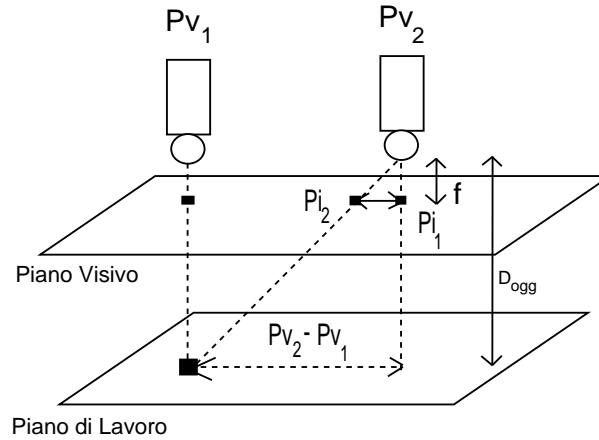


Figura 3.18: Calcolo della profondità dall'acquisizione di 2 sole immagini.

Per concludere, è possibile ricavare il valore della distanza D_{ogg} attraverso la semplice relazione

$$D_{ogg} = f \cdot \frac{Pv_2 - Pv_1}{P_{i_2} - P_{i_1}} \quad (3.51)$$

Tale relazione è stata ricavata dallo schema rappresentato in figura 3.18 dove il fuoco f è noto²².

Una volta avvenuta la calibrazione è quindi possibile, in maniera del tutto analoga al procedimento indicato a pagina 40, stimare la posizione di un oggetto ripreso dalla telecamera.

²²Ricavato quindi una volta per tutte da una calibrazione fuori linea.

3.5 Elaborazione dell'immagine

Nel paragrafo precedente è stata ricavata la matrice K che permette di ottenere la posizione spaziale dei punti utilizzando semplicemente i valori dalle loro proiezioni sul piano visivo. I risultati raggiunti sono notevoli, ma in realtà, difficilmente applicabili al mondo reale. Nei casi concreti la telecamera si trova ad inquadrare oggetti veri e propri ognuno dei quali caratterizzato da innumerevoli proprietà tra le quali il colore, la dimensione e la forma.

A questo punto è necessario disporre di una serie di strumenti che permettano, analizzando l'immagine acquisita, di poter estrarre da questa alcune proprietà utili agli obiettivi che il sistema deve raggiungere. La disciplina che si prefigge la risoluzione di questo tipo di problematica, cioè l'estrazione di proprietà da una immagine, è chiamata "*Image Processing*".

Di seguito si esporranno i concetti teorici che permettono di implementare delle tecniche di elaborazione dell'immagine per poter ricavare la posizione e l'orientamento di un oggetto inquadrato dalla telecamera.

Posizione di un oggetto

Se si suppone che la telecamera acquisisca immagini di tipo binario²³, cioè in bianco e nero, si può affermare che l'area di una immagine C' rappresentante la proiezione di un oggetto C sul piano visivo (vedere figura 3.19) è data da

$$M_{00} = \iint b(x, y) dx dy \quad (3.52)$$

Dove la funzione $b(x, y)$ è chiamata *funzione caratteristica* e vale 1 nei punti del piano visivo dove è presente l'immagine C' e 0 altrove. Utilizzando la 3.52, è possibile anche definire la posizione dell'immagine C' tramite il calcolo del suo centro di massa²⁴, infatti ponendo

$$M_{10} = \iint x b(x, y) dx dy \quad M_{01} = \iint y b(x, y) dx dy \quad (3.53)$$

si ottiene

$$\bar{x} = \frac{M_{10}}{M_{00}} \quad \bar{y} = \frac{M_{01}}{M_{00}} \quad (3.54)$$

²³Ipotesi semplificativa, ma che non lede la generalità dei risultati ottenuti.

²⁴Pur essendo una convenzione, l'utilizzo del centro di massa è una schematizzazione molto utile per rappresentare la posizione di un oggetto, per un esempio di veda l'appendice A.

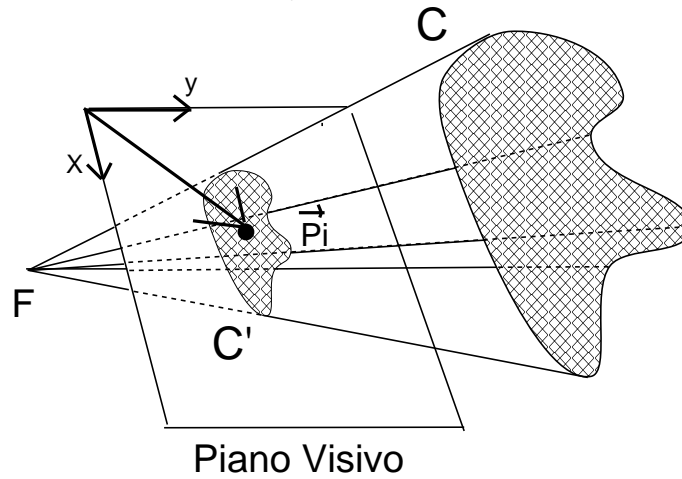


Figura 3.19: Esempio di proiezione di un oggetto.

dove $\vec{P}_i = (\bar{x}, \bar{y})$ rappresenta il vettore indicante la posizione (centro di massa) di C' rispetto al sistema di riferimento x, y del piano visivo²⁵.

Orientamento di un oggetto

Per determinare l'orientamento θ di C' si farà riferimento al suo asse di inerzia, questo coincide con la retta sul piano x, y che minimizza la quantità

$$I = \iint r^2 b(x, y) dx dy \quad (3.55)$$

dove r indica la distanza del generico punto di integrazione (x, y) dalla retta in questione. Una volta determinata tale retta, basta calcolarne il coefficiente angolare θ rispetto al sistema di riferimento x, y del piano visivo, come indicato in figura 3.20²⁶.

Le grandezze trovate si sono indicate con la nomenclatura M_{ij} , perché coincidono con i momenti di ordine 'i' rispetto all'asse x e di ordine 'j' rispetto all'asse y , dove per momento di intende

$$M_{ij} = \iint x^i y^j b(x, y) dx dy \quad (3.56)$$

²⁵Vedere [12, 19].

²⁶Vedere Appendice A per il procedimento dettagliato.

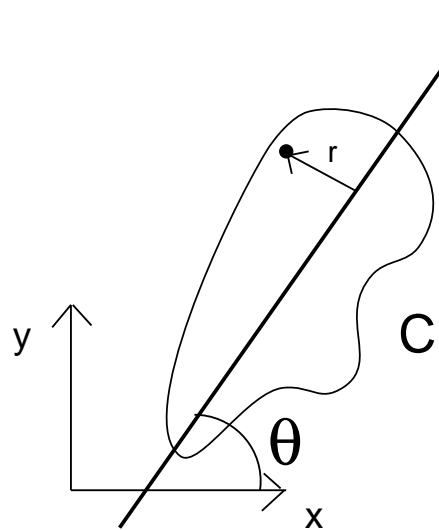


Figura 3.20: Orientamento di un oggetto sul piano visivo.

3.5.1 Immagini binarie discrete

Come anticipato nel paragrafo 3.2.3, concretamente la proiezione di un punto sul piano visivo viene in realtà acquisita dalla telecamera (digitale) come un elemento di una griglia di pixel, ed è appunto questa struttura dati che rappresenta l'immagine che poi verrà successivamente processata, quello che però non è ancora stato evidenziato è che questa trasduzione in pixel determina una quantizzazione dell'immagine ricevuta, con una conseguente perdita di informazione, questa discretizzazione però permette di ottenere una trattazione matematica semplificata. Si ipotizzi a questo proposito di avere una griglia di pixel di dimensione $(N \times M)$ ed un'immagine C' di un oggetto proiettato su di essa, allora l'area di C' sarà data dalla 3.52 che discretizzata diviene

$$M_{00} = \sum_{i=1}^N \sum_{j=1}^M b(x_i, y_j) \quad (3.57)$$

nello stesso modo è possibile riscrivere gli altri integrali, facendo ciò si ottiene una formulazione direttamente implementabile in un linguaggio di programmazione²⁷.

²⁷Vedere il codice della libreria *Imel*.

3.5.2 Etichettamento

In alcune applicazioni può risultare utile poter contare il numero di oggetti inquadrati dalla telecamera, il procedimento che si segue è quello di etichettare ogni oggetto presente nella griglia di pixel attraverso un algoritmo detto di *etichettamento*²⁸, cioè si vuole passare dalla situazione indicata in figura 3.21

```

000000000000000000000000000000
00000000110000000110000
00000011111100000111110
0000011111110000011100
00000011111110000000000
00000001111111000000000
00000001111111000000000
000000011110000001111000
000000010000000001111000
000000000000000001111100
000000000000000000000000

```

Figura 3.21: Esempio di valori di una griglia di pixel per una immagine binaria.

alla situazione finale in cui, ad ogni oggetto presente nella schermo, è stato associato un identificatore (nel caso indicato in figura 3.22 l'identificatore è un numero naturale appartenente all'intervallo $[1, N]$ dove N è il numero di oggetti presenti sullo schermo).

```

000000000000000000000000000000
00000000110000000220000
00000011111100000222220
0000011111110000022200
00000011111110000000000
00000001111111000000000
00000001111000000333000
00000001000000000333000
00000000000000000333100
000000000000000000000000

```

Figura 3.22: Esempio di etichettamento immagine.

²⁸In letteratura chiamato "*Labeling*".

3.5.3 Connettività

Per arrivare al risultato indicato in figura 3.22, bisogna introdurre il concetto di “connessione tra pixel” o “connettività”²⁹, si consideri il seguente insieme di pixel

$$\begin{array}{ccc} X_1 & X_2 & X_3 \\ X_4 & X & X_5 \\ X_6 & X_7 & X_8 \end{array}$$

1. Si dirà che il pixel X è 4-connesso se $X = 1$ ed almeno uno dei pixel X_2, X_4, X_5, X_7 è uguale ad 1.
2. Si dirà che il pixel X è 8-connesso se $X = 1$ ed almeno uno dei pixel $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8$ è uguale ad 1.

000	000	000
010	011	010
000	000	001
Pixel Isolato	Pixel 4-connesso ed 8 connesso	Pixel 8-connesso ma non 4- connesso

Figura 3.23: Esempio di catalogazione pixel.

Queste due definizioni quantificano il grado di connessione³⁰ tra pixel adiacenti, sono schematizzazioni necessarie per poter definire l'appartenenza o meno di un pixel ad un insieme. Quale dei due tipi di connessione utilizzare dipende unicamente dal tipo di risultato che si vuole ottenere, ad esempio la 4-connessione intende pixel confinanti diagonalmente come appartenenti ad oggetti separati, mentre per la 8-connessione tali pixel appartengono allo stesso oggetto (vedere figura 3.23).

²⁹Vedere [12, 9].

³⁰In letteratura chiamato “*Neighborhood*”.

Algoritmo di etichettamento

L'algoritmo che verrà implementato³¹ è di tipo iterativo e si scompone fondamentalmente in 3 passi logici

1. Inizializzazione.

Fase in cui si associa a ciascun pixel "acceso" dell'immagine una label diversa³² rappresentata da un numero naturale, via via crescente. Al termine di questa fase si sarà creata una nuova immagine in cui a ciascun pixel che aveva valore "1" all'interno dell'immagine originaria corrisponde un numero (naturale) diverso.

2. Passo top-down.

Si passa ora all'elaborazione dell'immagine appena ottenuta, scandendo pixel per pixel, a partire dall'angolo superiore sinistro fino all'angolo inferiore destro. Si associa al pixel p che si sta studiando (che chiameremo *pivot*) l'etichetta³³ di valore minimo trovata tra i 4 pixel confinanti che si trovano vicino al pivot e che sono già stati attraversati dal passo corrente, cioè i seguenti pixel:

$$\begin{array}{ccc} 1 & 2 & 3 \\ 4 & p & 0 \\ 0 & 0 & 0 \end{array}$$

Applicando questo passo ad ogni pixel, si viene a generare una nuova immagine caratterizzata da un numero di etichette minore od uguale al numero di etichette che erano presenti nell'immagine prima dell'elaborazione.

3. Passo bottom-up.

Si scandisce l'immagine, pixel per pixel, a partire dall'angolo inferiore destro dell'immagine fino all'angolo superiore sinistro, poi si esegue lo stesso procedimento seguito nel passo 2 sui seguenti pixel:

$$\begin{array}{ccc} 0 & 0 & 0 \\ 0 & p & 4 \\ 3 & 2 & 1 \end{array}$$

³¹Vedi la libreria *Imel*.

³²Questo algoritmo si applica ad immagini di tipo binario, cioè del tipo acceso/spento.

³³Rappresentata come detto da un numero naturale.

I passi 2 e 3 vengono ripetuti iterativamente e in sequenza fino a quando la loro esecuzione non apporta più alcuna modifica all'immagine³⁴, cioè non viene alterata neanche una etichetta, a questo punto il numero di etichette rimaste nell'immagine coincide con il numero di oggetti presenti nell'immagine binaria iniziale e l'etichettamento è concluso.

³⁴Per un esempio vedere [12] pag.40.