



# Navigazione basata su inseguimento di frecce

## *Relazione di progetto*

Progetto del corso Robotica (principi e progetto)

Università degli Studi di Bergamo

A.A. 2019/2020

CALEGARI ANDREA - 1041183

PAGANESSI ANDREA - 1040464

PIFFARI MICHELE - 1040658

December 30, 2019



# Contents

<b>1</b>	<b>Stato dell'arte</b>	<b>1</b>
1.1	SAL - Stato avanzamento lavori . . . . .	1
<b>2</b>	<b>Camera</b>	<b>5</b>
2.1	Scelta della camera . . . . .	5
2.2	Come ottenere le immagini dalla camera? . . . . .	6
2.3	CPU consumption . . . . .	6
2.4	Posizionamento camera . . . . .	6
2.5	Calibrazione . . . . .	6
<b>3</b>	<b>Gestione delle maschere</b>	<b>9</b>
3.1	HSV . . . . .	9
3.2	Frecce o cerchi? . . . . .	9
3.3	Maschere . . . . .	10
3.4	Erosione e dilatazione . . . . .	10
<b>4</b>	<b>Identificazione delle forme</b>	<b>11</b>
4.1	Definizione del problema . . . . .	11
4.2	Interpolazione . . . . .	11
<b>5</b>	<b>Dalle pixel coordinates alle world coordinates</b>	<b>13</b>
5.1	Definizione del problema . . . . .	13
5.2	Da world coordinates to camera coordinates . . . . .	14
5.3	Da camera coordinates a film coordinates . . . . .	15
5.4	Da film coordinates a pixel coordinates . . . . .	15
5.5	Problema al rovescio . . . . .	16
<b>6</b>	<b>Matrici di rotazione</b>	<b>19</b>



# List of Figures

1.1	Struttura del codice . . . . .	2
1.2	Base robotica addetta alla movimentazione . . . . .	2
1.3	Base verticale sulla quale andare ad inserire la camera . . . . .	3
2.1	Camera utilizzata nello stato iniziale del progetto . . . . .	5
2.2	Camera utilizzata nello step successivo . . . . .	5
2.3	Interfaccia grafica durante la calibrazione . . . . .	7
3.1	RGB vs HSV . . . . .	9
3.2	Erosione e dilatazione . . . . .	10
5.1	Schema concettuale delle diverse coordinate . . . . .	13
5.2	Posizione del world frame . . . . .	14
5.3	Descrizione del problema . . . . .	15
5.4	Descrizione dell'ultima trasformazione . . . . .	15
5.5	Problema inverso . . . . .	16
5.6	Piano del pavimento nel camera frame . . . . .	17



# 1

## Stato dell'arte

TODO: sostituire con introduzione (tipo quella del progetto vecchio)

Obiettivo: andare a implementare sistema di *visual navigation* per la base robotica in figura 1.2.

### 1.1 SAL - Stato avanzamento lavori

- Analizzato codice Out-Of-Box del progetto dello scorso anno
  - Il codice preso non aveva main: creato
  - Compresa struttura pub/sub
  - Analizzati topic/nodes pubblicati
- Cambio camera. Perché? Prestazioni scarse al variare della luce
- Nuova camera → ueye cam
- Fatta funzionare nuova camera
  - Demo (programma già fornito con la camera)
  - Ros → utilizzato file debug-launch (inserire caratteristiche che la camera fornisce quando parte lo script).
- Scelta la posizione della camera: verticale inclinata e non orizzontale
- Progetto A.A. utilizza formule vecchie
- Problema CPU consuming (problema intrinseco della camera)
- Memory problem → risolto con *free*
- Doppie maschere: frecce di due colori
- Aggiunte queste features:
  - Gaussian blur
  - Brightness
  - Erosion - Dilatation
- Fatta erosione solo sulle frecce vicine (quelle nella metà inferiore del frame), mentre invece, le frecce nella metà superiore non vengono erose ma solo dilatate.
- Problema inizializzazione che mostrava rettangoli bianchi su alcune immagini intermedie nelle maschere
- Aggiunta distanza tra centri con tracciamento linea

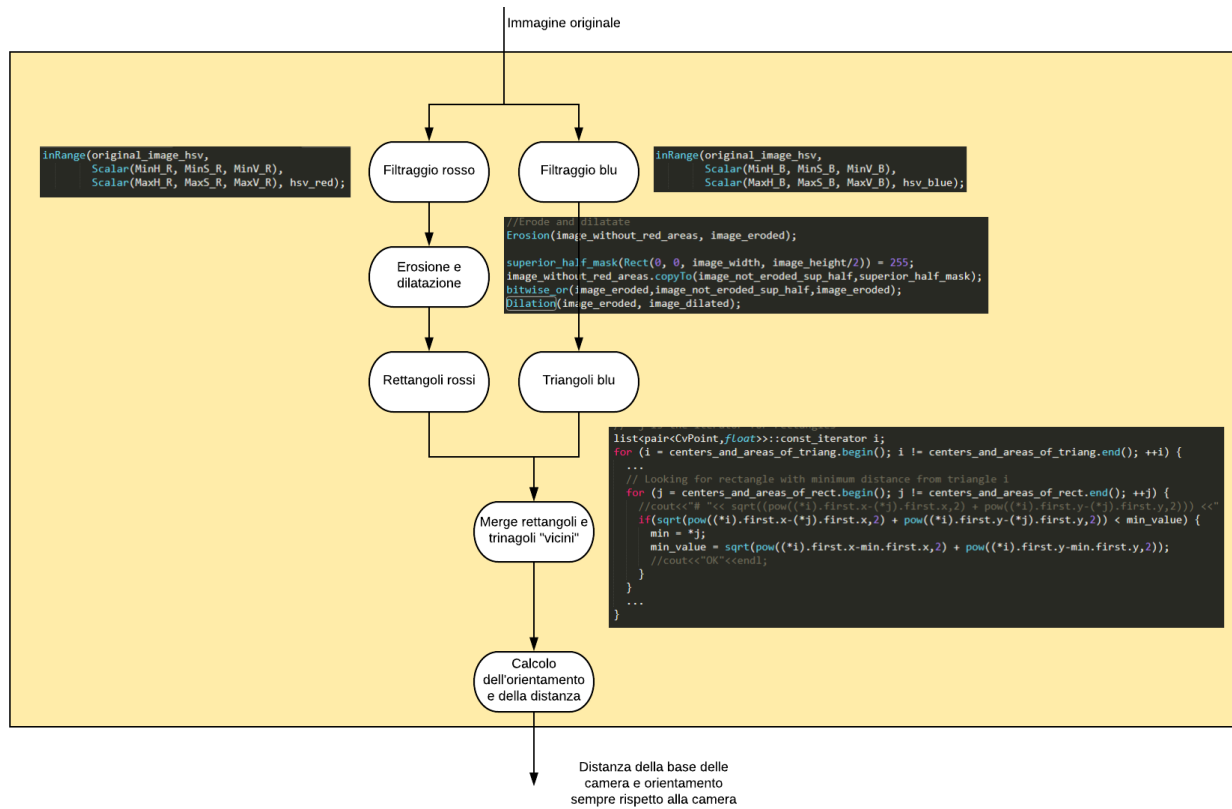


Figure 1.1: Struttura del codice



Figure 1.2: Base robotica addetta alla movimentazione





Figure 1.3: Base verticale sulla quale andare ad inserire la camera

- Prendiamo la freccia più vicina e analizziamo i dati relativi solo a questa freccia: supponendo che tutte le frecce siano uguali, è ovvio che l'area maggiore è quella della freccia più vicina (TODO: da mettere come giustificazione del codice che scriveremo)

Per creare grafi della struttura del codice ROS vedi e comando `rqt`.



## 2

# Camera

### 2.1 Scelta della camera

Il progetto, allo stato iniziale, prevedeva di andare ad utilizzare la camera *SpotLight Pro Webcam* (fig. 2.1), webcam già utilizzata in un progetto precedente, da cui abbiamo preso spunto per partire.



Figure 2.1: Camera utilizzata nello stato iniziale del progetto

È stato però notato che, utilizzando questa camera, non si era in grado di dare sufficienti garanzie di funzionamento stabile in alcune delle più comuni condizioni luminose e ambientali: infatti era alta la variabilità del comportamento della camera al variare delle condizioni luminose, il che rendeva molto instabile il riconoscimento delle frecce.

Si è deciso quindi di passare ad una camera di tipo industriale, in grado di fornire delle prestazioni più stabili e affidabili.

La scelta è ricaduta sulla camera della casa produttrice *IDS (Imaging Development System)*: si tratta del modello *UI-1221LE-C-HQ* equipaggiata con la lente *BM2420* prodotta dalla *Lensagon* (lens datasheet).

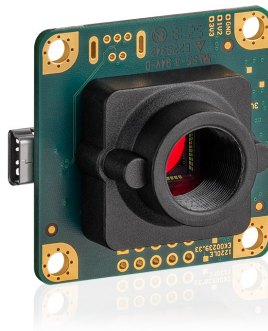


Figure 2.2: Camera utilizzata nello step successivo

La camera in Fig.2.2 ha un'otturatore globale che permette di acquisire tutta l'immagine istantaneamente e non in modo progressivo come la camera *SpotLight Pro Webcam*.

## 2.2 Come ottenere le immagini dalla camera?

Per ottenere le immagini dalla camera è stato necessario iscriversi presso il sito web della casa produttrice e scaricare i driver necessari all'installazione e al funzionamento (software).

La videocamera ha diverse opzioni di acquisizione tra cui anche una grandezza dell'immagine diversa dal classico 640x480 pixel ma, per ragioni di semplicità, si è scelto di lasciare invariate le impostazioni di default.

In ogni caso è molto semplice verificare lo stato di funzionamento della camera stessa: è sufficiente, una volta scaricati e installati i software proprietari della casa produttrice, andare ad aprire il software *uEyeDemo* e modificare le impostazioni in base alle proprie necessità.

E' necessario sottolineare come, per essere in grado di ottenere le immagini dalla camera, sia importante assicurarsi che *l'ueye daemon* sia in funzione: esso parte automaticamente dal momento in cui si avvia il PC con la camera già connessa. Nel caso in cui essa venga connessa a caldo è necessario andare ad avviare il *daemon* utilizzando questo comando da terminale, che mette in evidenza come si tratti di un comando per camera usb (nel caso si andasse a lavorare con camera ethernet, servirebbe andare a selezionare la giusta opzione, proposta a terminale):

```
sudo /etc/init.d/ueyeusbdrc start
```

## 2.3 CPU consumption

Un fattore importante nella scelta della camera è l'elevato tempo di utilizzo della CPU da parte della camera *UI-1221LE-C-HQ*. In contrasto, la camera inizialmente scelta vantava un consumo di CPU nettamente inferiore e quindi che più si potrebbe adattare all'installazione su dispositivi mobili e con bassa potenza di calcolo.

## 2.4 Posizionamento camera

Durante i test e gli esperimenti svolti in laboratorio la videocamera è stata legata ad un palo metallico e fissata con un angolo di inclinazione rispetto ad esso di circa 35 gradi; l'altezza dal pavimento è, inoltre, di circa 83.5 cm

## 2.5 Calibrazione

La fase di calibrazione della camera permette di ricavare (alcuni o tutti) i parametri che permettono al modello *pin-hole* di poter essere utilizzato per proiettare punti da coordinate mondo a coordinate camera.

In inglese la calibrazione della camera, ovvero il ricavare i parametri intrinseci e/o estrinseci, si chiama *Camera resectioning* in quando il concetto di Camera Calibration si può riferire anche al problema della calibrazione fotometrica del sistema.

Una telecamera è infatti generalmente modellata mediante il modello proiettivo centrale pinhole camera (o foro stenopeico). Tale modello è definito da due famiglie di parametri:

- Parametri intrinseci (da tre a cinque): descrivono la telecamera indipendentemente dalla sua posizione nello spazio.
- Parametri estrinseci (sei): descrivono la posizione della telecamera nello spazio indipendentemente dalle sue caratteristiche interne.

Nel nostro caso specifico, per ottenere questi parametri di interesse, abbiamo sfruttato la forte interconnessione che la camera *IDS* offre con l'ambiente ROS: infatti, installando uno specifico pacchetto software tramite il comando

```
rosdep install camera_calibration
```

che permette di ottenere un insieme di pacchetti software scritti in *Python*, i quali permettono di ottenere tutti i parametri di interesse.

Nello specifico, il comando da eseguire tramite il terminale è il seguente:

```
roslaunch camera_calibration cameracalibrator.py --size 8x6 --square 0.108 image:=/
camera/image_raw camera:=/camera
```

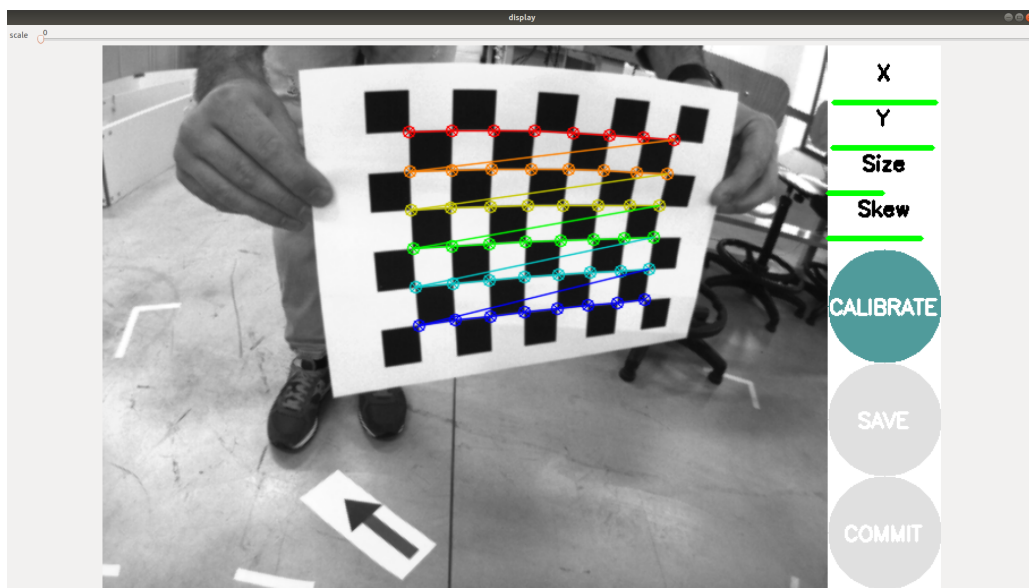


Figure 2.3: Interfaccia grafica durante la calibrazione

Questi invece sono gli esiti forniti in uscita dal processo di calibrazione:

```
width
640

height
480

[narrow_stereo]

camera matrix
575.407063 0.000000 392.952424
0.000000 572.733663 256.805741
0.000000 0.000000 1.000000

distortion
-0.310832 0.140048 -0.005256 -0.007508 0.000000

rectification
1.000000 0.000000 0.000000
0.000000 1.000000 0.000000
0.000000 0.000000 1.000000

projection
517.612549 0.000000 400.977651 0.000000
```

26	0.000000	540.829285	256.819184	0.000000
27	0.000000	0.000000	1.000000	0.000000

## 3

# Gestione delle maschere

### 3.1 HSV

Nella strutturazione del progetto ci è venuto molto naturale andare a lavorare con una scala di colori HSV.

Ma perchè non applicare un filtraggio basato su RGB?

Come noto nella letteratura, nell'ambito dell'*image recognition* è usuale il problema di andare a mascherare un colore piuttosto che un altro, come nel nostro caso.

Potremmo voler trovare, sempre per esempio, oggetti di colore rosso, scannerizzando nell'immagine solamente colori (255,0,0) nella scala RGB: con questo approccio andremmo ad applicare una condizione troppo stringente ai colori. Si potrebbe pensare, come soluzione a questa condizione parecchio stringente, di trovare colori in un range di rossi, come per esempio (130,0,0);(255,0,0): il problema comunque persisterebbe proprio per il fatto che il rosso è ottenuto come combinazione di più colori primari, e non come un solo singolo colore.

Potremmo pensare dunque, di andare a fondo del problema, cambiando gli intervalli dei valori RGB per tutti i colori primari ma sarebbe uno sforzo non secondario e probabilmente con risultati poco significativi. È necessario un metodo che abbia meno parametri per semplificare l'identificazione dei colori: questo metodo è rappresentato proprio dall'utilizzo del HSV.

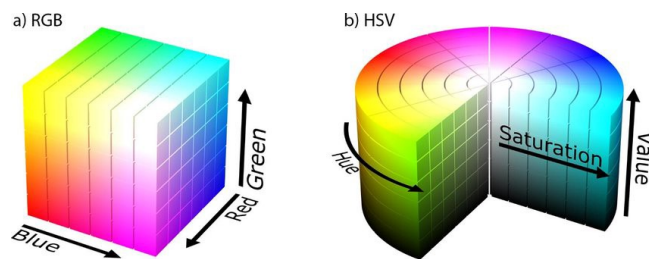


Figure 3.1: RGB vs HSV

Si può notare come, in Fig 3.1, il colore rosso sia definito in un range ben chiaro di un solo valore (*Hue*) mentre *Saturation* and *Value* servono solo a limitare ulteriormente la gamma di tonalità di rosso accettabili.

### 3.2 Frecce o cerchi?

Nel progetto utilizzato come punto di partenza, si erano individuati cerchi di diversa grandezza per testare le funzionalità della libreria OpenCV nell' object detection.

Ovviamente l'utilizzo di cerchi limita di molto l'espressività del simbolo in quanto un cerchio non può, per sua natura identificare una direzione, men che meno un verso. Le frecce sono quindi la soluzione più naturale al problema di identificare con un simbolo una direzione ed un verso che poi, un eventuale robot mobile, potrà seguire.

TODO mettere foto cerchi e frecce foto maschera

### 3.3 Maschere

Le maschere sono fondamentali nel processo di object detection tramite la libreria OpenCV.

Queste ultime sono immagini binarie (bianco e nero) e sono utilizzate da tutte le funzioni di identificazione dei contorni e interpolazione di punti presenti in OpenCV. La seguente linea di codice

```
28 inRange(original_image_hsv, Scalar(MinH_R, MinS_R, MinV_R), Scalar(MaxH_R, MaxS_R, MaxV_R), hsv_red);
```

permette di identificare tutti gli oggetti di colore rosso e di salvarli nella Matrice *hsv\_red*.

### 3.4 Erosione e dilatazione

L'erosione e la dilatazione sono tecniche note allo stato dell'arte attuale utili per rimuovere rumore da una immagine.

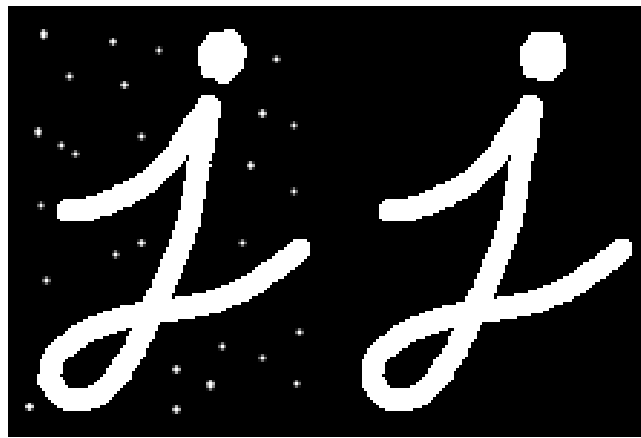


Figure 3.2: Erosione e dilatazione

Nel caso preso in considerazione da questo report è stato necessario applicare la tecnica di erosione solamente alla parte inferiore dell'immagine in quanto, se applicata nella parte superiore, avrebbe eliminato oltre al rumore anche un'eventuale freccia che sarebbe apparsa molto piccola.



## 4

# Identificazione delle forme

## 4.1 Definizione del problema

Il capitolo precedente ha mostrato come sia possibile ottenere dei contorni, delle figure data un'immagine a colori. È ora necessario identificare la forma di ciascuno di questi contorni riconoscendo così i vari quadrati e rettangoli presenti nell'immagine che avessero, una volta acquisiti dalla camera, il colore specificato. Come ultimo passaggio va svolto il controllo per verificare che esista o meno una freccia; quest'ultima altro non è che un triangolo e un rettangolo sufficientemente vicini fra di loro.

## 4.2 Interpolazione

Per approssimare il contorno ottenuto e filtrato attraverso le mask, come spiegato nel Capitolo 3, è necessario usare una funzione che approssima il contorno individuato con un altro poligono avente meno vertici così che la distanza tra di essi sia inferiore ad una certa soglia. Tale funzione è così definita nella libreria OpenCV:

```
void approxPolyDP(InputArray curve, OutputArray approxCurve, double epsilon, bool
closed)
```

Si è reso necessario effettuare un tuning del parametro *double epsilon* in quanto, per frecce diverse, a distanza variabile e con orientazione non fissa sono stati individuati differenti valori ottimali. Il valore che più si adattava a tutti i casi presi in considerazione è stato ottenuto sperimentalmente e corrisponde a *double epsilon=0.045*.

La funzione di cui sopra restituisce quindi una lista di poligoni ognuno dei quali è descritto da una lista di vertici. Il passo successivo è stato cercare nella lista dei poligoni un elemento che avesse 4 lati nel caso di un rettangolo e 3 in quello di un triangolo:

```
29 if(result->total >= 3 && result->total <= 3 && fabs(cvContourArea(result,
CV_WHOLE_SEQ))>lower_area_triang)
```

```
30 if(result->total >= 4 && result->total <= 6 && fabs(cvContourArea(result,
CV_WHOLE_SEQ))>lower_area_rect)
```

Sempre per via sperimentale è stato possibile scoprire che vincolando il poligono che approssima il quadrilatero cercato ad avere tra i 4 e i 6 lati, la probabilità di riconoscere correttamente un rettangolo aumentava. Per il triangolo questo non si è reso necessario vista la buona probabilità di successo nella ricerca vincolata a 3 lati.

Ottenuti ora tutti i triangoli e i rettangoli sufficientemente grandi nella figura va affrontato il problema del riconoscimento di ogni freccia presente nel seguente modo:

- per ciascun rettangolo identificato, si calcola la distanza che intercorre tra esso e ogni triangolo riconosciuto. Per calcolare la distanza tra due figure è necessario:

- definire il centro del rettangolo, tramite funzioni della libreria OpenCV
- ottenere il baricentro del triangolo
- calcolare la distanza cartesiana tra i due punti appena individuati
- si tiene in considerazione solamente la distanza minore calcolata.
- si confronta suddetto valore con una soglia sperimentale; se questo valore è minore allora si può assumere che il triangolo e il quadrato presi in considerazione siano una freccia, altrimenti si scarta la coppia.
- la freccia appena rilevata viene aggiunta alla lista delle frecce rilevate nell'immagine.

Per ogni freccia, che ora altro non è che una coppia di punti,

$$\begin{aligned} C_{triangolo} &= (x_t, y_t) \\ C_{rettangolo} &= (x_r, y_r) \end{aligned} \tag{4.1}$$

vanno identificati nell'ordine:

- il centro della freccia, ottenuto come il punto medio del segmento che collega i due centri che definivano la freccia precedentemente.

$$C_{freccia} = \left( \frac{x_t + x_r}{2}, \frac{y_t + y_r}{2} \right)$$

- L'inclinazione della freccia nel piano, calcolata come:

$$\begin{aligned} \phi &= \text{atan}\left(\frac{\Delta y}{\Delta x}\right), \text{dove} \\ \Delta y &= y_t - y_r \\ \Delta x &= x_t - x_r \end{aligned} \tag{4.2}$$

- L'area dell'oggetto freccia, ricavata come somma dell'area del triangolo e del quadrato.

# Dalle pixel coordinates alle world coordinates

## 5.1 Definizione del problema

Nel capitolo precedente è stato illustrato un metodo atto all'identificazione delle frecce che rientrano nel campo visivo della camera. Viene ora trattato come sia possibile ottenere la posizione della freccia, precedentemente ottenuta, nelle coordinate 3D (U,V,W) rispetto alla base del robot.

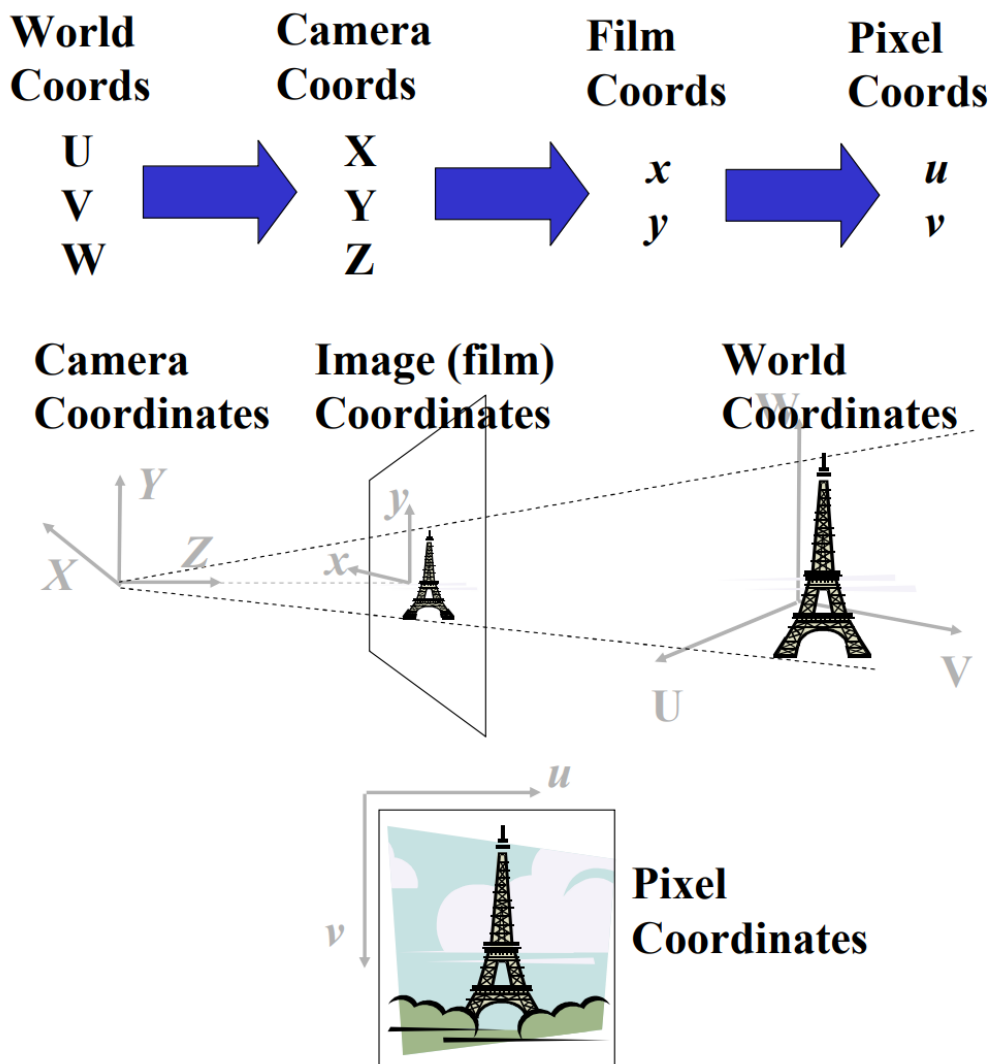


Figure 5.1: Schema concettuale delle diverse coordinate

Come si vede in Fig.5.1 si devono effettuare tre trasformazioni per ottenere a partire dalle world coordinates le pixel coordinates. Il problema, nel caso specifico preso in considerazione dal report, risulta essere l'opposto: dalle coordinate nella camera è necessario ottenere la posizione globale dell'oggetto effettuando una trasformazione inversa. Si analizzeranno ora le singole trasformazioni che permetteranno alla fine di ottenere il risultato voluto.

## 5.2 Da world coordinates to camera coordinates



Figure 5.2: Posizione del world frame

Partendo da un sistema di riferimento solidale al robot e all'altezza del pavimento, che identifichiamo come sistema globale, è possibile definire una matrice di rototraslazione per ottenere il sistema di riferimento solidale al centro della camera.

$$R_{worldcam} = R_{traslazione} * R_{rotazione} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & h_{cam} \\ 0 & 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} \cos(\frac{\pi}{2} + \alpha) & -\sin(\frac{\pi}{2} + \alpha) & 0 & 0 \\ \sin(\frac{\pi}{2} + \alpha) & \cos(\frac{\pi}{2} + \alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

È stata effettuata una traslazione lungo l'asse W in quanto la camera è posta esattamente sopra all'origine del sistema  $O_{UVW}$  ed una rotazione rispetto all'asse U di  $\frac{\pi}{2}$  in quanto, per convenzione, si associa all'asse delle Z la profondità nel frame solidale alla camera. A questo punto è necessario effettuare un'altra rotazione di  $\alpha$  gradi rispetto all'asse X a seconda dell'inclinazione alla quale si sceglie di far lavorare la camera **INSERISCI LA FOTO DELLA CAMERA SUL PALO PER FAR CAPIRE IL CONCETTO!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!**

### 5.3 Da camera coordinates a film coordinates

#### Basic Perspective Projection

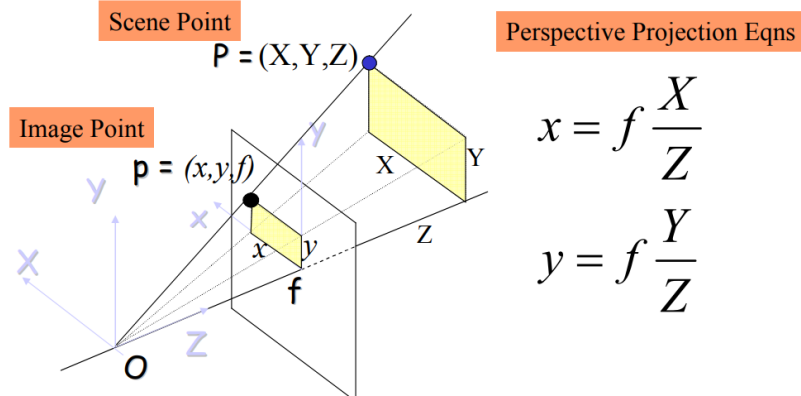


Figure 5.3: Descrizione del problema

In Fig.5.3  $f$  rappresenta il fuoco della camera, parametro ottenibile attraverso una procedura di calibrazione.

### 5.4 Da film coordinates a pixel coordinates

#### Intrinsic parameters (offsets)

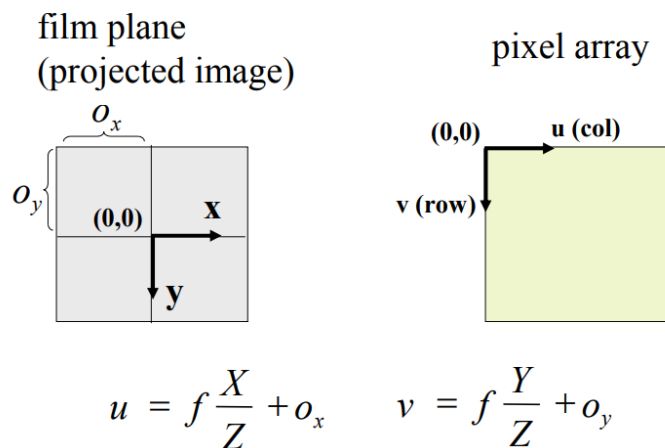


Figure 5.4: Descrizione dell'ultima trasformazione

In Fig.5.4  $O_x$  e  $O_y$  sono anch'esse ricavabili tramite la procedura di calibrazione della camera

## 5.5 Problema al rovescio

### Backward Projection

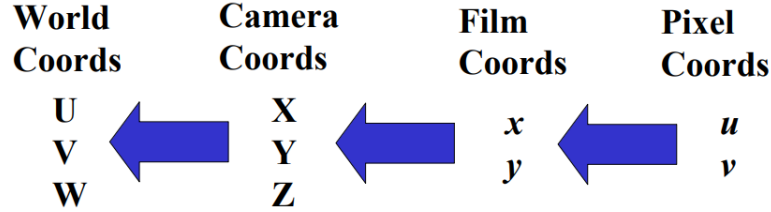


Figure 5.5: Problema inverso

Come si vede in Fig.5.5 si deve ora affrontare il processo inverso siccome, nel caso trattato, si hanno a disposizione  $u$  e  $v$  e si vogliono ottenere  $U, V, W$ . Come già detto prima  $O_x$  e  $O_y$  e  $f$  sono ottenibili tramite calibrazione; dunque:

$$\begin{aligned} x &= u - O_x \\ y &= v - O_y \end{aligned} \tag{5.1}$$

Sono state ottenute le equazione del film coordinates. È ora necessario ottenere le camera coordinates. Per fare ciò è necessario conoscere il valore di  $Z$ , cosa ottenibile in due modi:

- utilizzando una camera con sensore di profondità
- assumendo che gli oggetti siano sempre posti su un piano di cui si conosce l'equazione.

La seconda assunzione è, nella realtà dei fatti, una ipotesi corretta in quanto gli oggetti e le frecce giaceranno sempre sul pavimento. È quindi richiesto di calcolare l'equazione del pavimento nel camera frame:

•

$$z = 0 \tag{5.2}$$

rappresenta l'equazione del piano se fosse nel world frame

•

$$z * R_{worldcam} = 0 \tag{5.3}$$

il piano così calcolato è la descrizione dal punto di vista matematico del pavimento dal punto di vista della camera

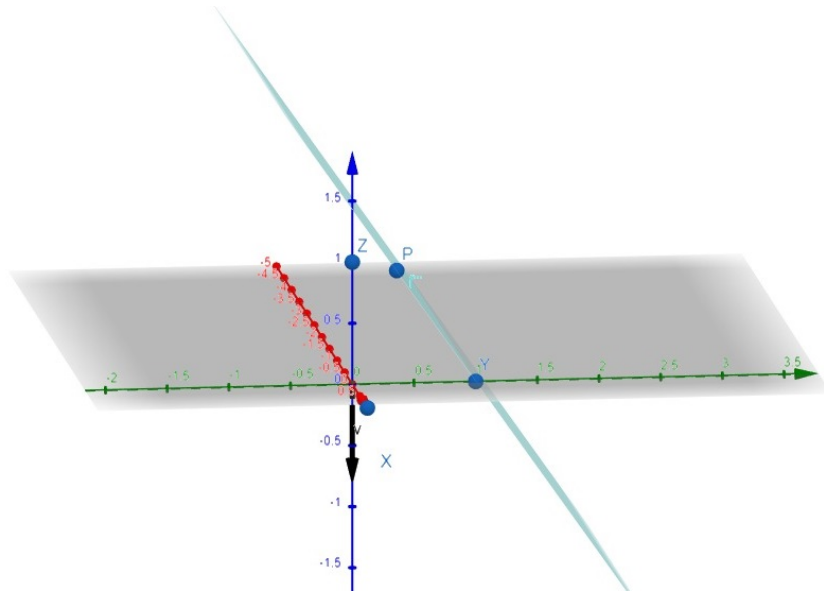


Figure 5.6: Piano del pavimento nel camera frame

- si può dunque calcolare il fattore di scala  $s$  dato un generico piano  $ax + by + cz + d = 0$

$$\frac{-d}{aX' + bY' + c} \quad (5.4)$$

dove

$$\begin{aligned} X' &= \frac{x}{f} = \frac{X}{Z} \\ Y' &= \frac{y}{f} = \frac{Y}{Z} \end{aligned} \quad (5.5)$$

che sono le coordinate normalizzate rispetto a  $Z$ .

- Dunque, come ultimo passaggio si moltiplica tutto per il fattore di scala:

$$\begin{aligned} X &= X' * s \\ Y &= Y' * s \\ Z &= S \end{aligned} \quad (5.6)$$

Per ottenere le equazione nel world frame si deve dunque utilizzare la matrice inversa:

$$\begin{pmatrix} U \\ V \\ W \\ 1 \end{pmatrix} = R_{worldcam}^{-1} * \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$





**6**

## **Matrici di rotazione**



# Bibliography

- [1] *Descrizione della camera* <https://en.ids-imaging.com/store/ui-12211e-rev-2.html>
- [2] *Manuale della camera* [https://en.ids-imaging.com/IDS/datasheet\\_pdf.php?sku=AB02422](https://en.ids-imaging.com/IDS/datasheet_pdf.php?sku=AB02422)
- [3] *Manuale della lente* <https://www.lensation.de/product/BM2420/>
- [4] *Ueye cam e ROS* <http://wiki.ros.org/ueye>
- [5] *HSV vs RGB* <https://medium.com/neurosapiens/segmentation-and-classification-with-hsv-8f2406>
- [6] *HSV vs RGB* <https://handmap.github.io/hsv-vs-rgb/>
- [7] *Camera Projection I* <http://www.cse.psu.edu/~rtc12/CSE486/lecture12.pdf>
- [8] *Camera Projection II* <http://www.cse.psu.edu/~rtc12/CSE486/lecture13.pdf>
- [9] *Coordinate omogenee* [http://robotics.unibg.it/teaching/robotics/pdf/14\\_Geometria3D.pdf](http://robotics.unibg.it/teaching/robotics/pdf/14_Geometria3D.pdf)
- [10] *Calibrazione camera I* <http://www.ce.unipr.it/people/medici/geometry/node145.html>
- [11] *Calibrazione camera II* [http://wiki.ros.org/camera\\_calibration](http://wiki.ros.org/camera_calibration)