Corso di Percezione Robotica

C. Modulo di Percezione Attiva

Visione robotica

Eliseo Stefano Maini
ARTS Lab, Scuola Superiore Sant’Anna
es.maini@ieee.org
s.maini@arts.sssup.it
050-883486

Rappresentazione delle immagini

Rappresentazione dell'immagine	Rappresentazione del contenuto dell'immagine	Rappresentazione del contenuto della scena rappresentata da un'immgine
↓	↓	↓
Rappresentazione singoli punti	**Rappresentazione degli oggetti estratti dall'immgine**	**Rappresentazione di oggetti nella scena**
Immagini a colori, livelli di grigio Immagini pre-elaborate	Rappresentazione di bordi, regioni, forme, proprietà…	Rappresentazione di scena 3D, oggetti visuali 2D e 3D, oggetti reali 3D, azioni …

Sommario della lezione

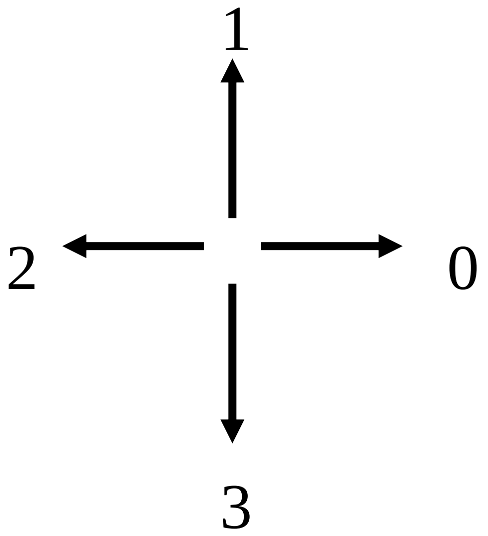
- Segmentazione dell'immagine
  - Descrizione dei contorni mediante codici a catena
  - Segmentation by fitting a model
    - Trasformata di Hough
    - Metodo ai minimi quadrati
- Visione stereoscopica
  - Fondamenti di 3D computer vision
    - Calibrazione telecamera
    - Calibrazione sistemi stereo e ricostruzione

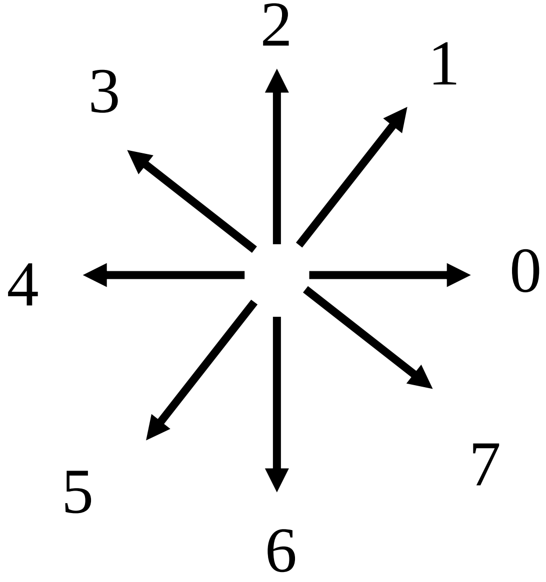
Descrittori dei contorni – codici a catena

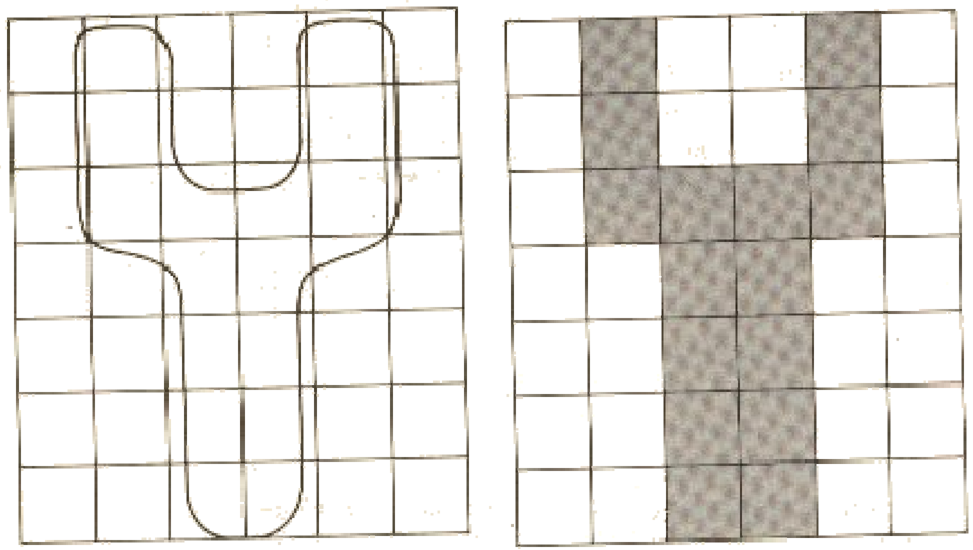
- Sono utilizzati per rappresentare un contorno come un insieme di segmenti di lunghezza e direzioni note
  - La lunghezza del segmento è data dalla dimensione della griglia
  - La direzione corrisponde al codice scelto

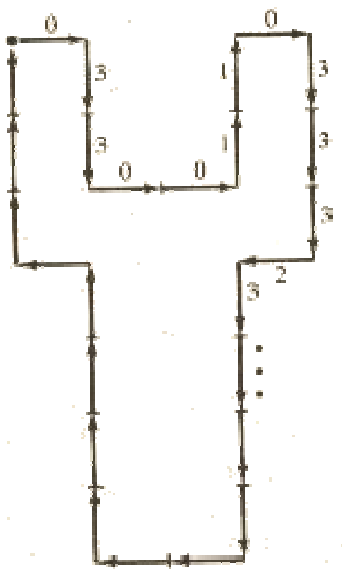
- Solitamente si usano segmenti 4-connessi o 8-connessi
- Per ridurre effetti di scala si scelgono dimensioni della griglia maggiori rispetto al singolo pixel

I codici a catena









Normalizzazione punto di partenza

- **Il codice dipende dal punto di partenza**

Risoluzione:

Tratto il codice come una sequenza circolare di codici numerici della direzione e ridefinisco il punto di partenza in modo tale che la sequenza dei numeri risultanti definisca un intero di grandezza minima

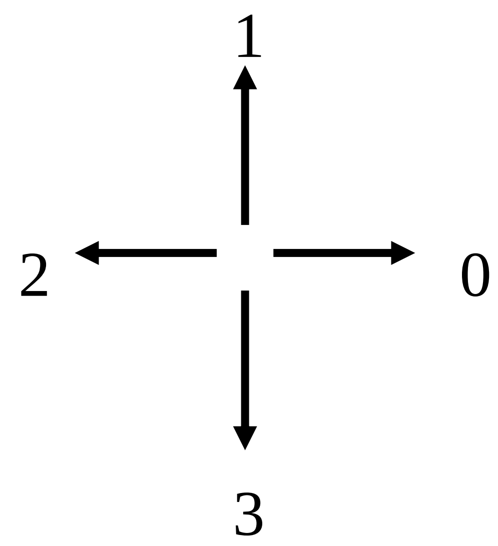
Normalizzazione rispetto alle rotazioni

- **Il codice dipende dall'orientamento del contorno**

Risoluzione

- Invece di utilizzare il codice utilizzo la differenza del codice a catena contando in senso antiorario il numero delle direzioni che separano due elementi adiacenti
- Se tratto il codice come una sequenza circolare allora il primo elemento della differenza viene calcolato utilizzando la differenza tra il primo e l'ultimo elemento della catena

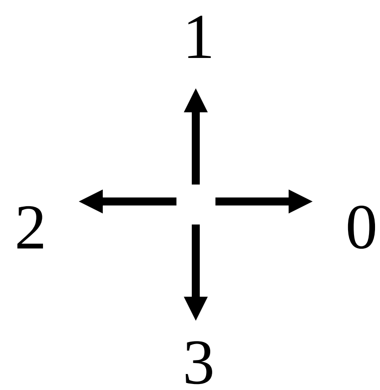
Esempio di normalizzazione delle rotazioni

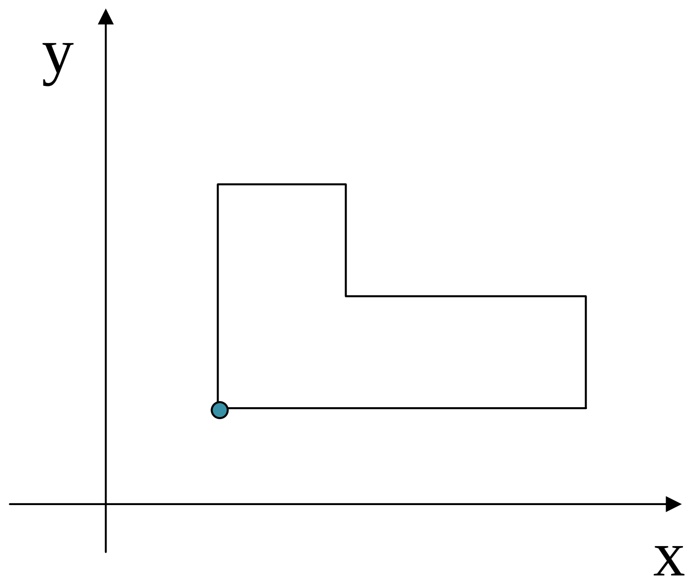


- Chain code : 1 0 1 0 3 3 2 2
- Differenza : 3 1 3 3 0 3 0
- Circolare : 3 3 1 3 3 0 3 0

Calcolo dell'area con i codici a catena

1. Area=0;
2. y=0;
3. per ogni elemento del codice a catena
4. case direzione
   - 0: area+=y;
   - 1: y++;
   - 2: area-=y;
   - 3: y--;





Segmentation by fitting a model

Segmentation by fitting a model

- Clustering voting techniques
  - Esempio della trasformata di Hough

- Least-squares methods
  - Esempio algoritmo di fitting di ellissi

Trasformata di Hough

- E' una tecnica che permette di riconoscere particolari configurazioni di punti presenti nell'immagine, come segmenti, curve o altre forme prefissate.
- E' un tipico *operatore globale*.
- Il principio fondamentale è che la forma cercata può essere espressa tramite una funzione nota che fa uso di un insieme di parametri.
- Una particolare istanza della forma cercata è quindi completamente precisata dal valore assunto dall'insieme di parametri.

Trasformata di Hough – Caso della retta

- Per esempio, assumendo come rappresentazione della retta la forma:

$$y=ax+b$$

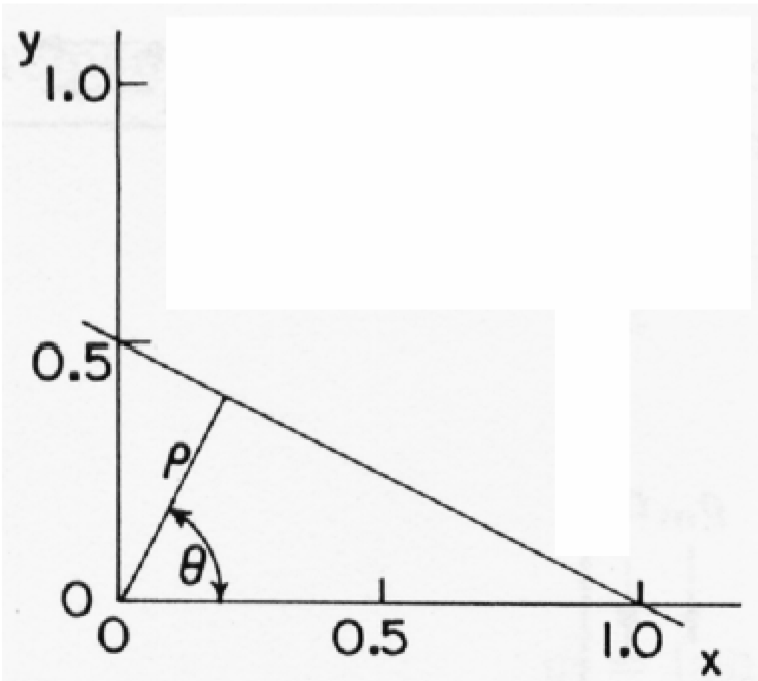
- qualunque retta è completamente specificata dal valore dei parametri (a,b).

- Se si assume un tipo di rappresentazione diversa, quale la forma normale di Hesse

$$\rho = x \cos\vartheta + y \sin \vartheta$$

- la retta è completamente specificata dalla coppia (ρ,ϑ).

Esempio



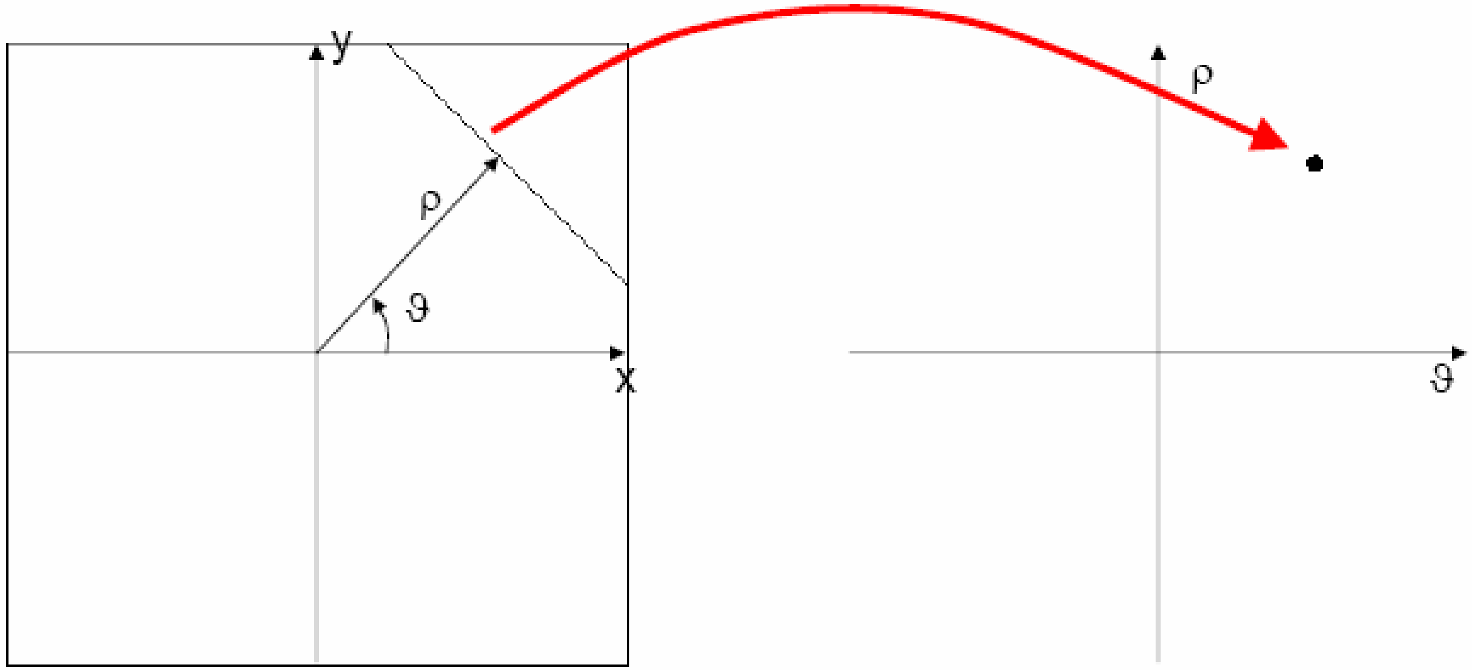
- La retta in figura è identificata dalla coppia:

  (a,b)=(-0.5,0.5)

- o dalla coppia:

(ρ ,ϑ)=(0.447,1.107)

Il piano trasformato



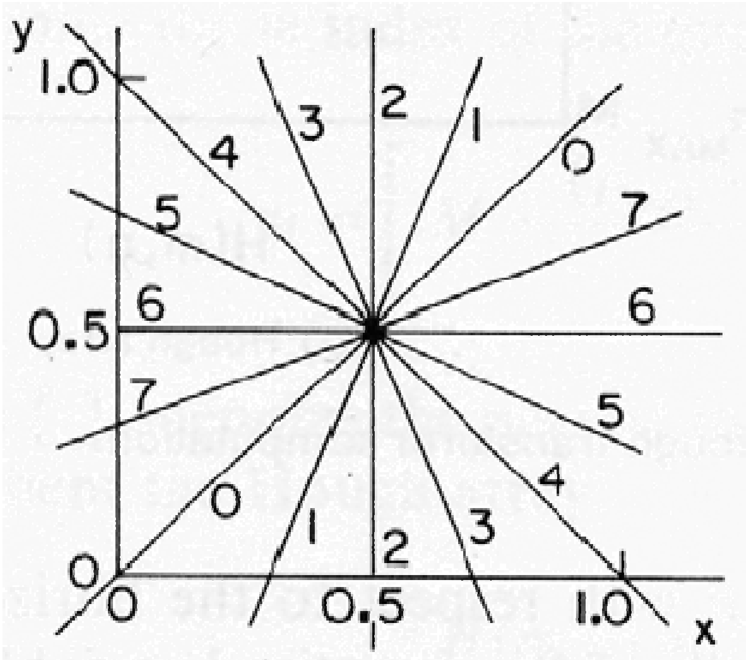
Domande

- Nell'immagine in analisi, l'unica informazione disponibile è costituita dall'insieme di punti che appartiene al foreground.

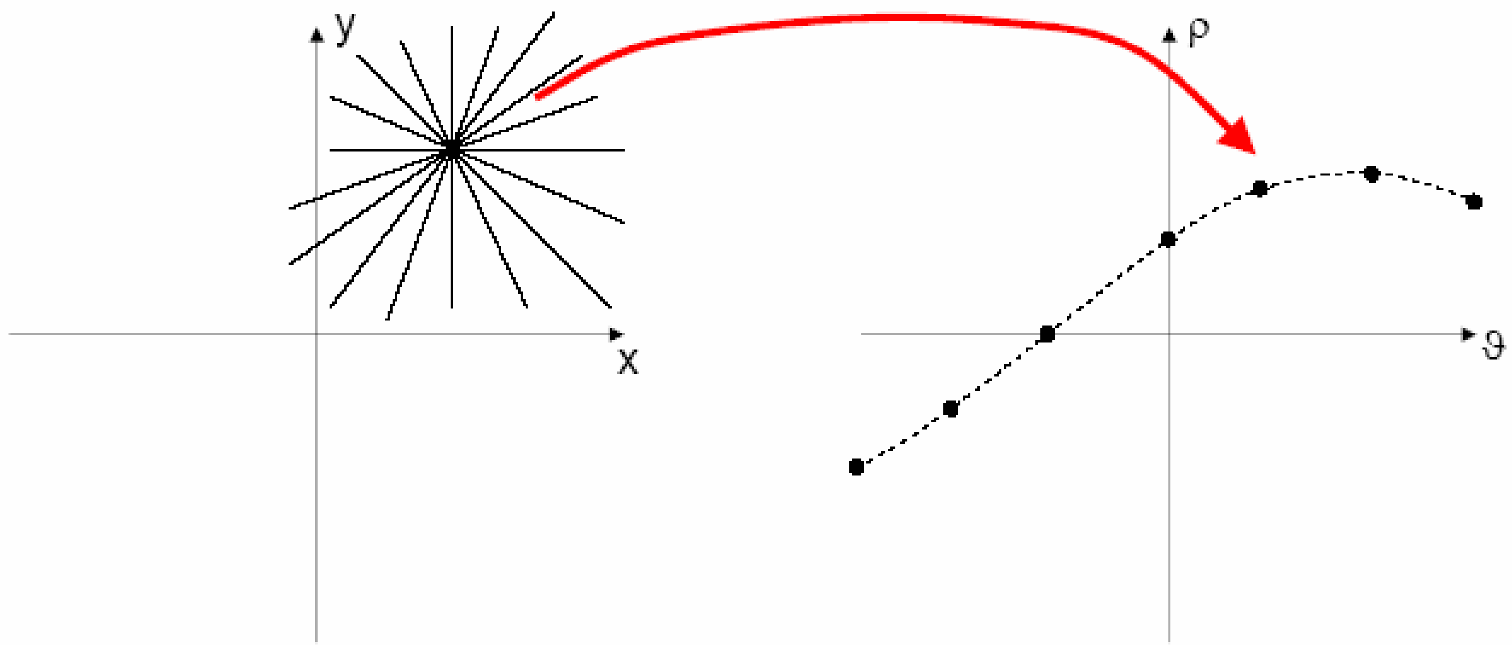
1. Come poter sfruttare questa trasformata ai fini della individuazione di segmenti in un'immagine ?
2. Qual è la trasformata di un punto nell'immagine ?

Trasformata di un punto

- Nel piano dell’immagine, un punto è identificato dall’intersezione di più rette.
- Quindi, ad ogni punto P corrisponde, nel piano dei parametri, la curva formata dai punti immagine delle rette passanti per P.



Trasformata del punto

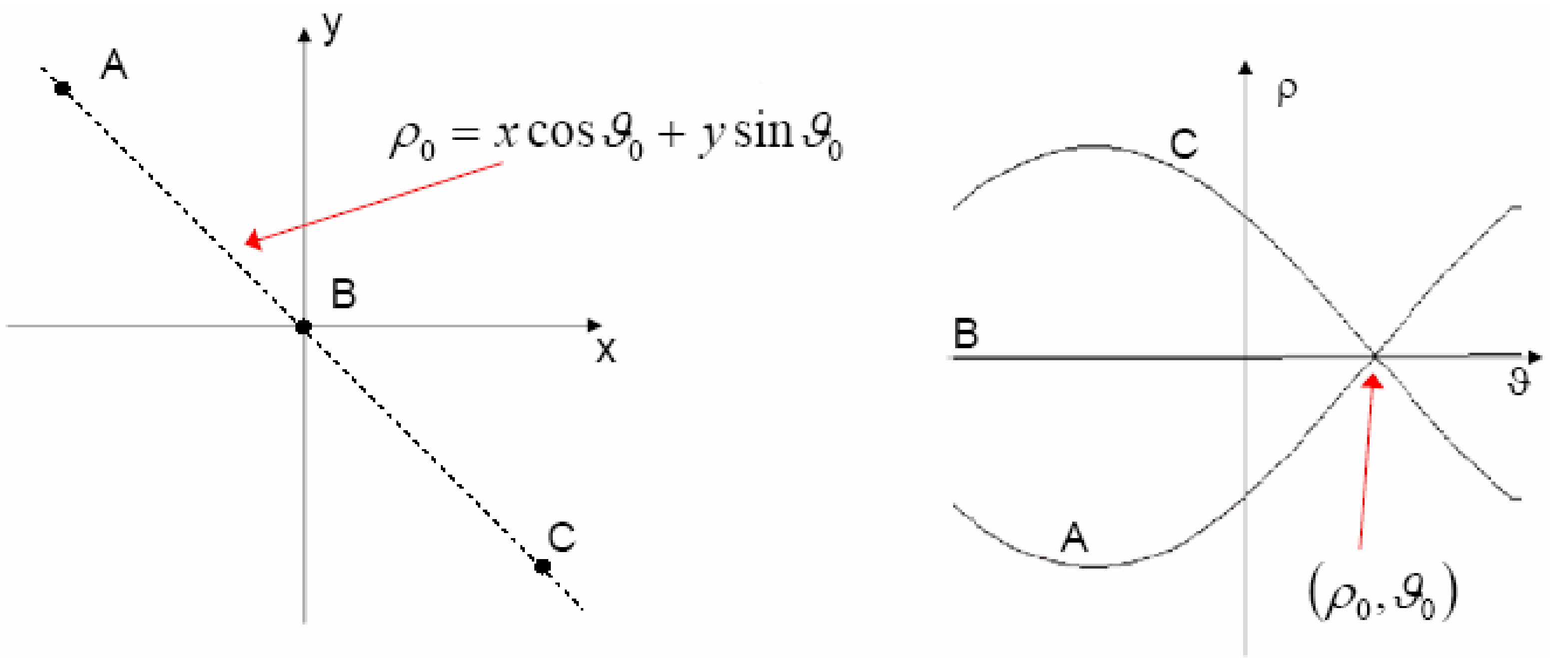


Domanda

- Che cosa succede se nell'immagine ci sono dei punti allineati su una stessa retta ?

- Sul piano dei parametri, le curve che corrispondono alle trasformazioni dei vari punti si intersecano in un punto del piano trasformato che è l'immagine della retta su cui giacciono i punti.

Individuazione della retta sul piano trasformato



Implementazione trasformata di Hough – caso della retta

- Si consideri una discretizzazione del piano dei parametri (ρ,ϑ).
- Ciò permette di rappresentare tale piano su una matrice H(m,n) i cui indici di riga e di colonna corrispondono ai valori quantizzati di ρ e ϑ.

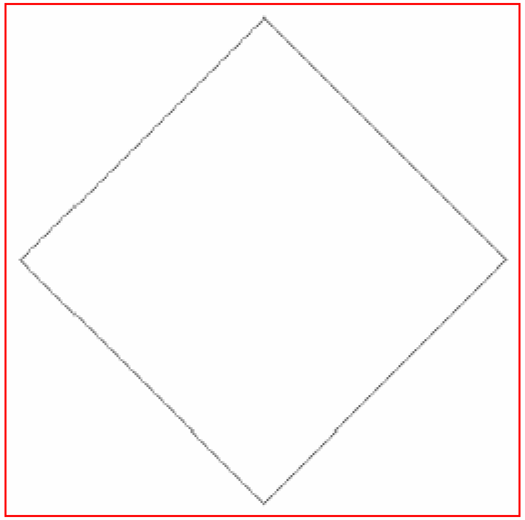
- Gli intervalli di variazione di ρ e ϑ sono fissati sulla base delle caratteristiche dell'immagine originale.
  - Tipicamente $-\rho_{max} \leq \rho \leq \rho_{max}$, $-\pi/2 \leq \vartheta \leq \pi/2$,
  - dove $\boldsymbol{\rho_{max}=0.5*(NR^2+NC^2)^{1/2}}$ e (NR,NC) sono le dimensioni dell'immagine originale.
  - Il numero dei livelli di quantizzazione va poi scelto in base all'accuratezza desiderata.
  - Una scelta quasi sempre soddisfacente è max(NR,NC).

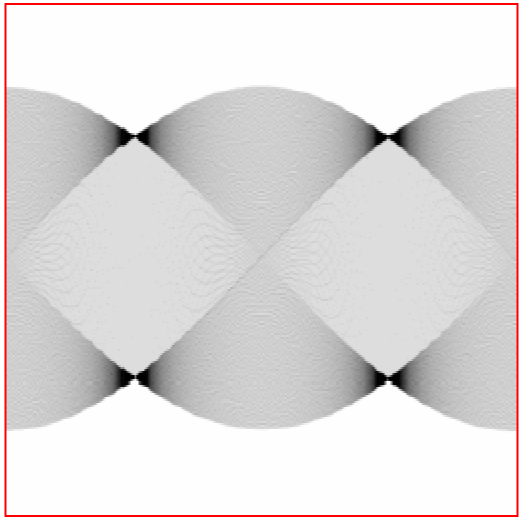
Algoritmo

1. Si azzeri la matrice H(.,.);
2. Per ogni punto $P \in F$, $P=(x,y)$
   1. per ϑ_n che varia tra $-\pi/2$ e $\pi/2$ con passo $d\vartheta$
      1. si valuta $\rho(n)=x*\cos(\vartheta_n)+y*\sin(\vartheta_n)$
      2. si ricava l'indice m corrispondente a $\rho(n)$
      3. si incrementi H(m,n)
   2. end
3. End

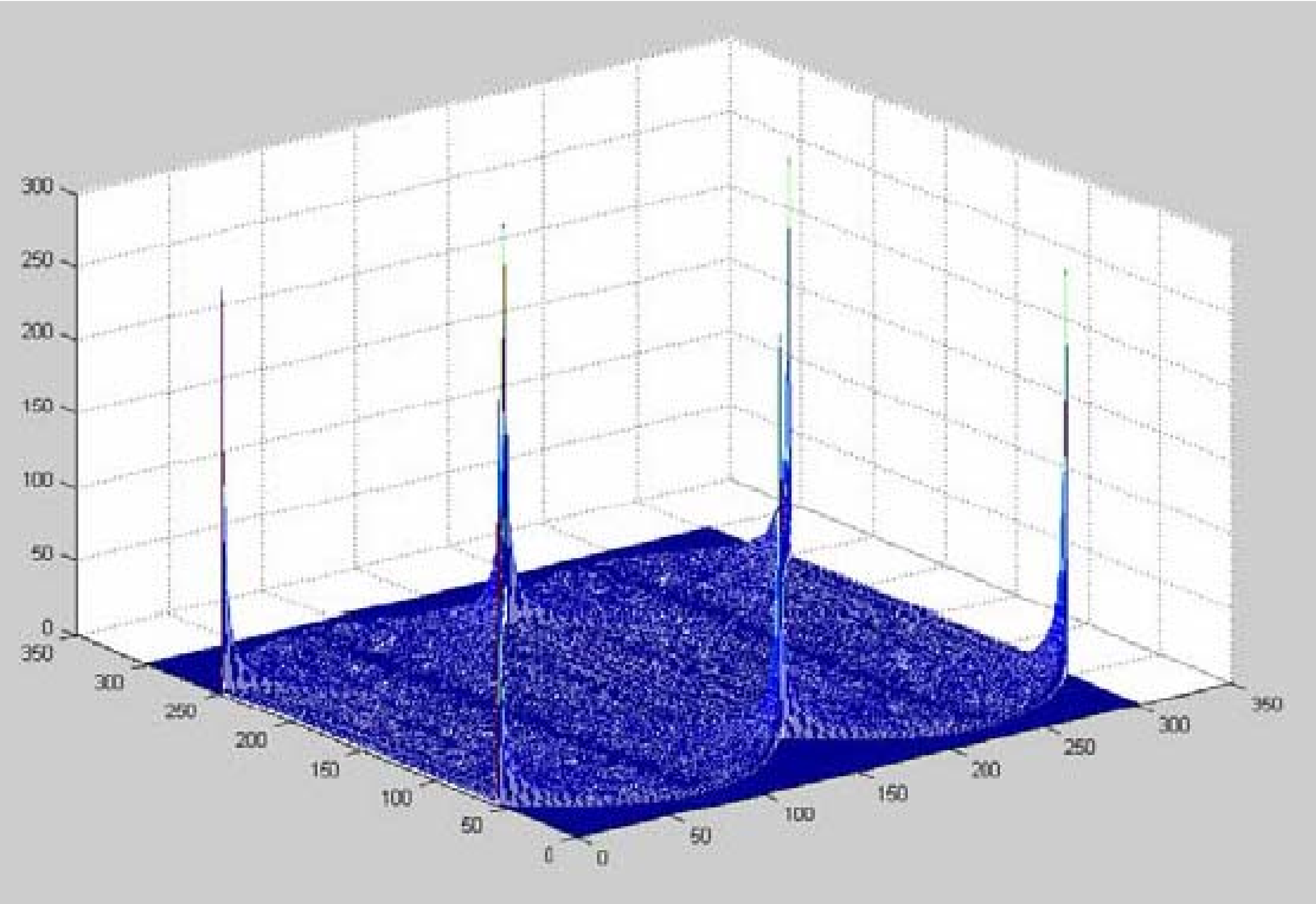
4. Si individuino i massimi locali su H(.,.) corrispondenti ai parametri dei segmenti individuati

Esempio: figura piana

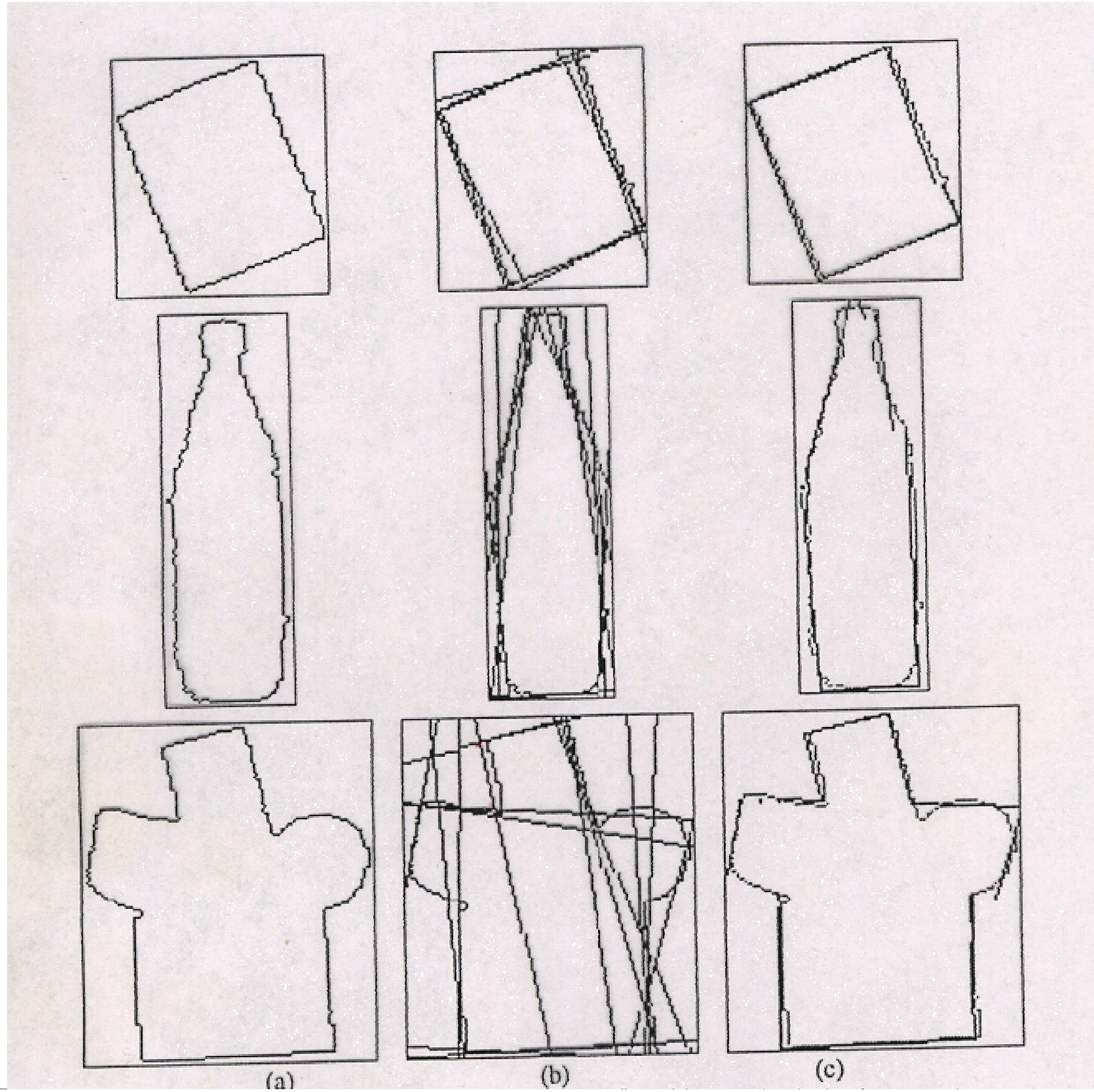




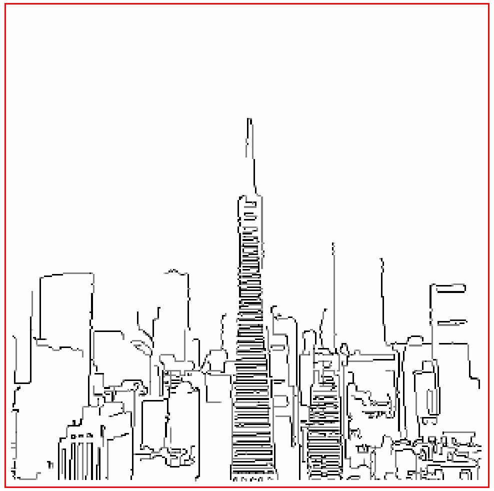
Esempio: matrice accumulazione

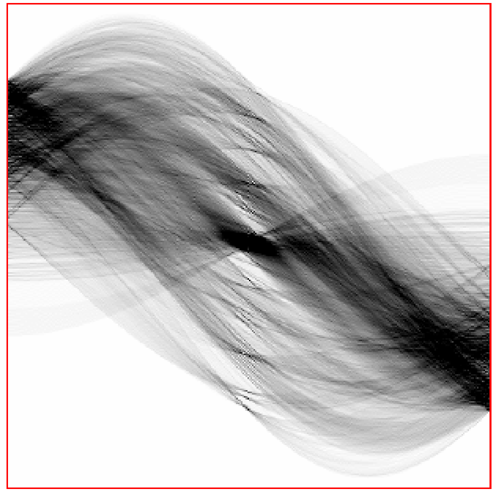


Il problema della soglia

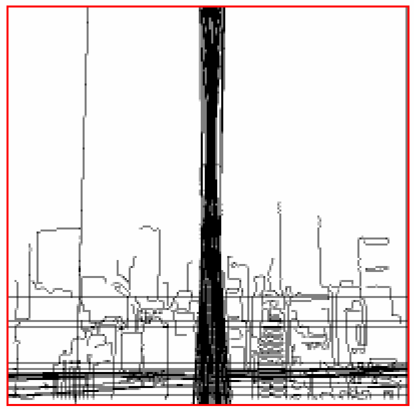


Sogliatura immagini reali

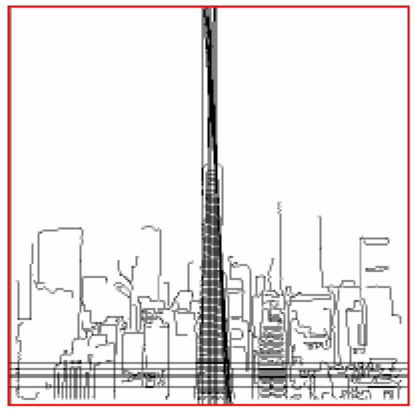




Soglia: 101



Soglia: 140



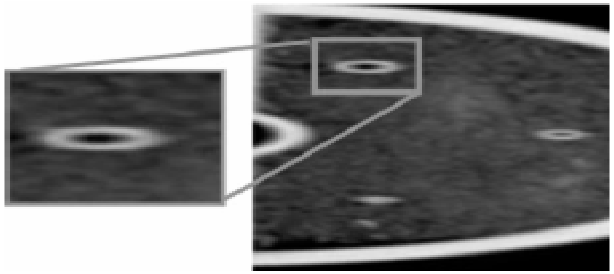
Soglia: 160

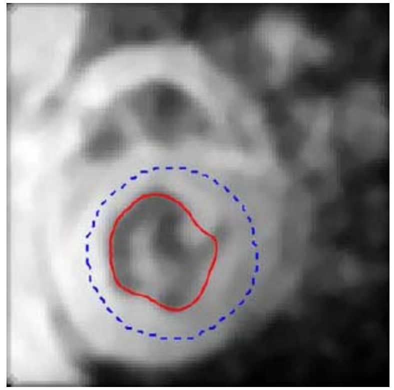
Conclusioni

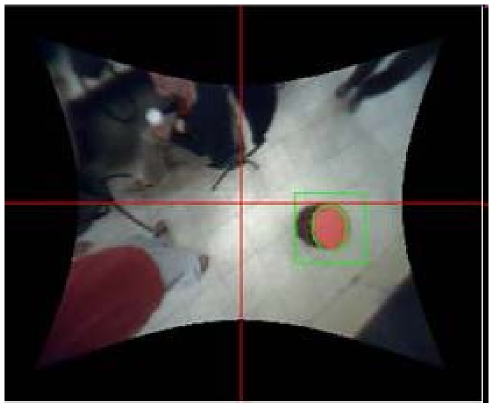
- E' possibile – <u>**ma molto oneroso**</u> - utilizzare la trasformata di Hough per individuare cerchi, tenendo conto dell'equazione $(x-a)^2+(y-b)^2=c^2$
  - In questo caso è possibile lavorare su:
    - un piano dei parametri (a,b), fissando il raggio c dei cerchi da individuare
    - uno spazio dei parametri (a,b,c), facendo variare c in un intervallo finito.
- E' stata inoltre proposta (Ballard) una generalizzazione della trasformata di Hough che permette di individuare oggetti di forma qualunque.
- **Non è adatta a applicazioni real-time**

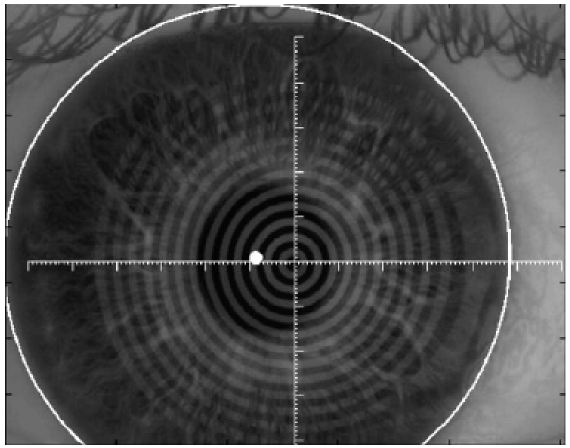
Geometric Fitting ai minimi quadrati

Fitting di curve ellittiche









Formalizzazione del problema

$$\boxed{F(x,y) = ax^2 + bxy + cy^2 + dx + ey + f = 0}$$

$$\mathbf{a} = [a,b,c,d,e,f]^T$$

$$\mathbf{x} = [x^2, xy, y^2, x, y, 1]$$

$$\boxed{F_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a} = 0}$$

Least Square Fitting

Opera minimizzando la somma dei quadrati di *un opportuno termine di errore* $\boldsymbol{e_i}$ valutato in ciascun punto dell'insieme dato

$$T = \{(x_i, y_i) : i = 1 \ldots N\}$$

$$\min_{\mathbf{a}} \left(\sum_{i=1}^{N} e_i^2 \right)$$

Identificazione funzione di errore

- **Distanza euclidea** dei punti dalla curva teorica
  - Risoluzione di un'equazione quartica
  - Per ogni soluzione calcolare la distanza minima

- Approssimare con misure più semplici la distanza euclidea
  - Distanza algebrica
  - Distanza algebrica pesata dal gradiente
  - etc

Equazione del vincolo

$$T = \{(x_i, y_i) : i = 1 \ldots N\}$$

Algebraic distance $E(\mathbf{a}; \mathbf{p}_i)$

$$\boxed{\min_{\mathbf{a}} \left(\sum_{i=1}^{N} E(\mathbf{a}; \mathbf{p}_i) \right) = \min_{\mathbf{a}} \left(\sum_{i=1}^{N} (\mathbf{p}_i \cdot \mathbf{a})^2 \right)}$$

Introdurre un vincolo per evitare la soluzione triviale
(es. $||\mathbf{a}||^2 = 1$ oppure $a + c = 1$ o anche $f = 1$)

Se vogliamo trovare solo curve ellittiche il vincolo deve essere

$$b^2 - 4ac < 0$$

$$\boxed{\begin{array}{llll} \forall \alpha \neq 0 \quad \text{if} & F_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a} = 0 & \Rightarrow & \alpha \cdot F_{\mathbf{a}}(\mathbf{x}) = \alpha \cdot \mathbf{x} \cdot \mathbf{a} = 0 \\ \text{therefore} & \exists \tilde{\alpha} & \text{such that} & 4ac - b^2 = 1 \end{array}}$$

Risoluzione

$$\min_{\mathbf{a}}\left(\sum_{i=1}^{N} E(\mathbf{a};\mathbf{p}_i)\right)=\min_{\mathbf{a}}\left(\sum_{i=1}^{N}(\mathbf{p}_i\cdot\mathbf{a})^2\right)$$

$$4ac-b^2=1$$

Design matrix

$$\mathbf{D}=\begin{pmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1\\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1\\ . & . & . & . & . & .\\ x_N^2 & x_Ny_N & y_N^2 & x_N & y_N & 1 \end{pmatrix}$$

Scatter matrix

$$\mathbf{S}=\mathbf{D}^T\mathbf{D}$$

Constraint matrix

$$\mathbf{C}=\begin{pmatrix} 0 & 0 & 2 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0 & 0\\ 2 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{cases}\mathbf{Sa}=\lambda\mathbf{Ca}\\ \mathbf{a}^T\mathbf{Ca}=1\end{cases}$$

TEOREMA La soluzione del problema vincolato precedente ammette esattamente una soluzione ellittica corrispondente all'unico autovalore positivo del problema di autovalori generalizzato

Attenzione al condizionamento di S

Re-centering and scaling

$$\mathbf{S}=\mathbf{D}^T\mathbf{D}=\begin{pmatrix}
\sum_{i=1}^{N} x_i^4 & \sum_{i=1}^{N} x_i^3 y_i & \sum_{i=1}^{N} x_i^2 y_i^2 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^2 y_i & \sum_{i=1}^{N} x_i^2 \\
\sum_{i=1}^{N} x_i^3 y_i & \sum_{i=1}^{N} x_i^2 y_i^2 & \sum_{i=1}^{N} x_i^3 y_i & \sum_{i=1}^{N} x_i^2 y_i & \sum_{i=1}^{N} x_i y_i^2 & \sum_{i=1}^{N} x_i y_i \\
\sum_{i=1}^{N} x_i^2 y_i^2 & \sum_{i=1}^{N} x_i^3 y_i & \sum_{i=1}^{N} y_i^4 & \sum_{i=1}^{N} x_i y_i^2 & \sum_{i=1}^{N} y_i^3 & \sum_{i=1}^{N} y_i^2 \\
\sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^2 y_i & \sum_{i=1}^{N} x_i y_i^2 & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i y_i & \sum_{i=1}^{N} x_i \\
\sum_{i=1}^{N} x_i^2 y_i & \sum_{i=1}^{N} x_i y_i^2 & \sum_{i=1}^{N} y_i^3 & \sum_{i=1}^{N} x_i y_i & \sum_{i=1}^{N} y_i^2 & \sum_{i=1}^{N} y_i \\
\sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i y_i & \sum_{i=1}^{N} y_i^2 & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} y_i & \sum_{i=1}^{N} i
\end{pmatrix}$$

N $(10^3)^4$ = N 10^{12}

N

Normalizzazione

Centering factors: $x_m = \min_{i=1}^{N}\{x_i\}$ $\quad y_m = \min_{i=1}^{N}\{y_i\}$

Scaling factors: $s_x = \dfrac{\max_{i=1}^{N}\{x_i\} - \min_{i=1}^{N}\{x_i\}}{2}$ $\quad s_y = \dfrac{\max_{i=1}^{N}\{y_i\} - \min_{i=1}^{N}\{y_i\}}{2}$

Affine transformation $\boxed{\hat{x} = \dfrac{x - x_m}{s_x} - 1 \qquad \hat{y} = \dfrac{y - y_m}{s_y} - 1}$

$$\mathbf{S} = \mathbf{D}^T\mathbf{D} = \begin{pmatrix}
\sum_{i=1}^N x_i^4 & \sum_{i=1}^N x_i^3 y_i & \sum_{i=1}^N x_i^2 y_i^2 & \sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^2 y_i & \sum_{i=1}^N x_i^2 \\
\sum_{i=1}^N x_i^3 y_i & \sum_{i=1}^N x_i^2 y_i^2 & \sum_{i=1}^N x_i^3 y_i & \sum_{i=1}^N x_i^2 y_i & \sum_{i=1}^N x_i y_i^2 & \sum_{i=1}^N x_i y_i \\
\sum_{i=1}^N x_i^2 y_i^2 & \sum_{i=1}^N x_i^3 y_i & \sum_{i=1}^N y_i^4 & \sum_{i=1}^N x_i y_i^2 & \sum_{i=1}^N y_i^3 & \sum_{i=1}^N y_i^2 \\
\sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^2 y_i & \sum_{i=1}^N x_i y_i^2 & \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i y_i & \sum_{i=1}^N x_i \\
\sum_{i=1}^N x_i^2 y_i & \sum_{i=1}^N x_i y_i^2 & \sum_{i=1}^N y_i^3 & \sum_{i=1}^N x_i y_i & \sum_{i=1}^N y_i^2 & \sum_{i=1}^N y_i \\
\sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i y_i & \sum_{i=1}^N y_i^2 & \sum_{i=1}^N x_i & \sum_{i=1}^N y_i & \sum_{i=1}^N i
\end{pmatrix}$$

N $(1)^4$ = N

N

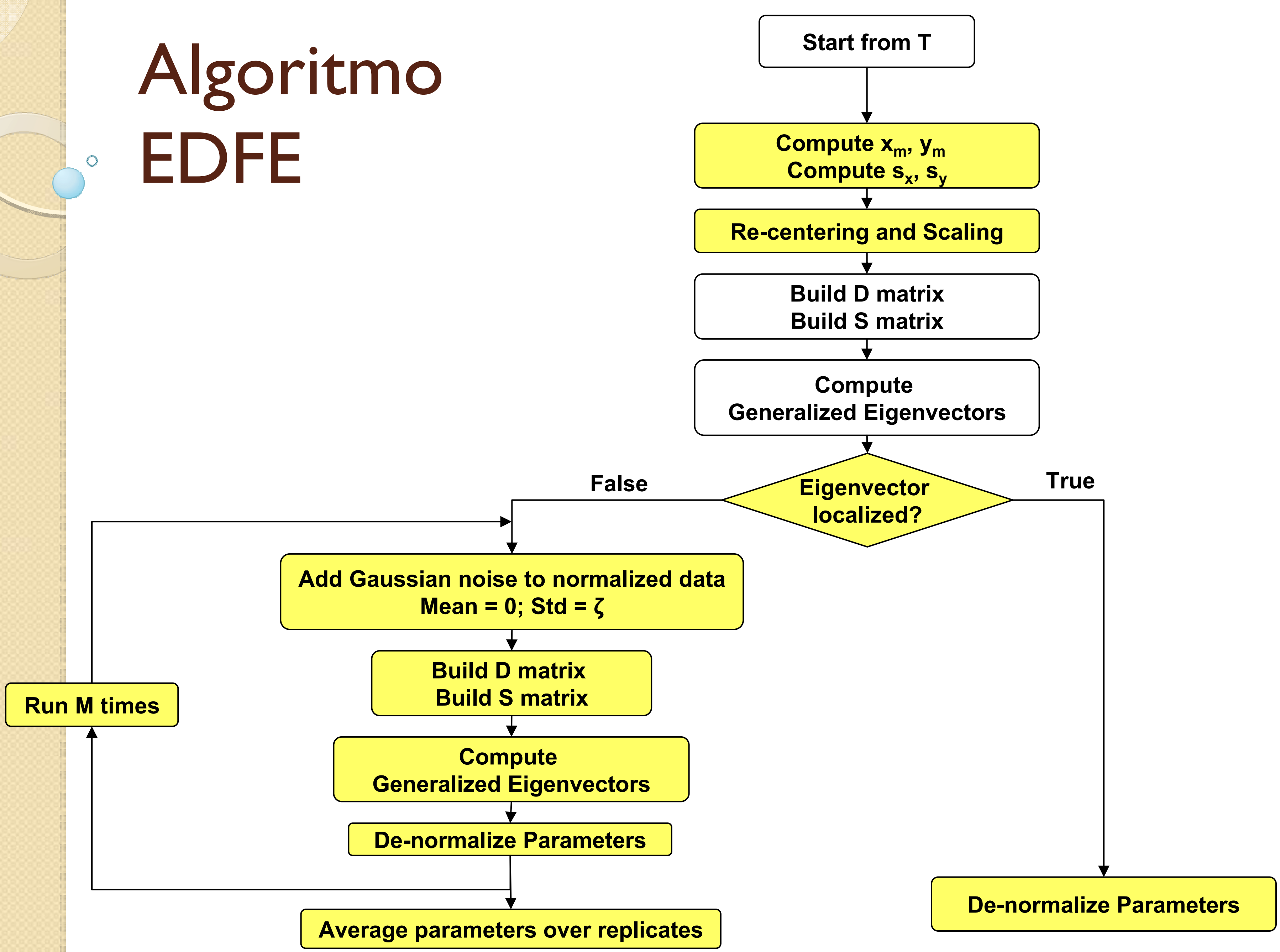
Attenzione alla soluzione "ideale"

- Punti esattamente sull'ellisse $\Rightarrow$ determinante di S nullo

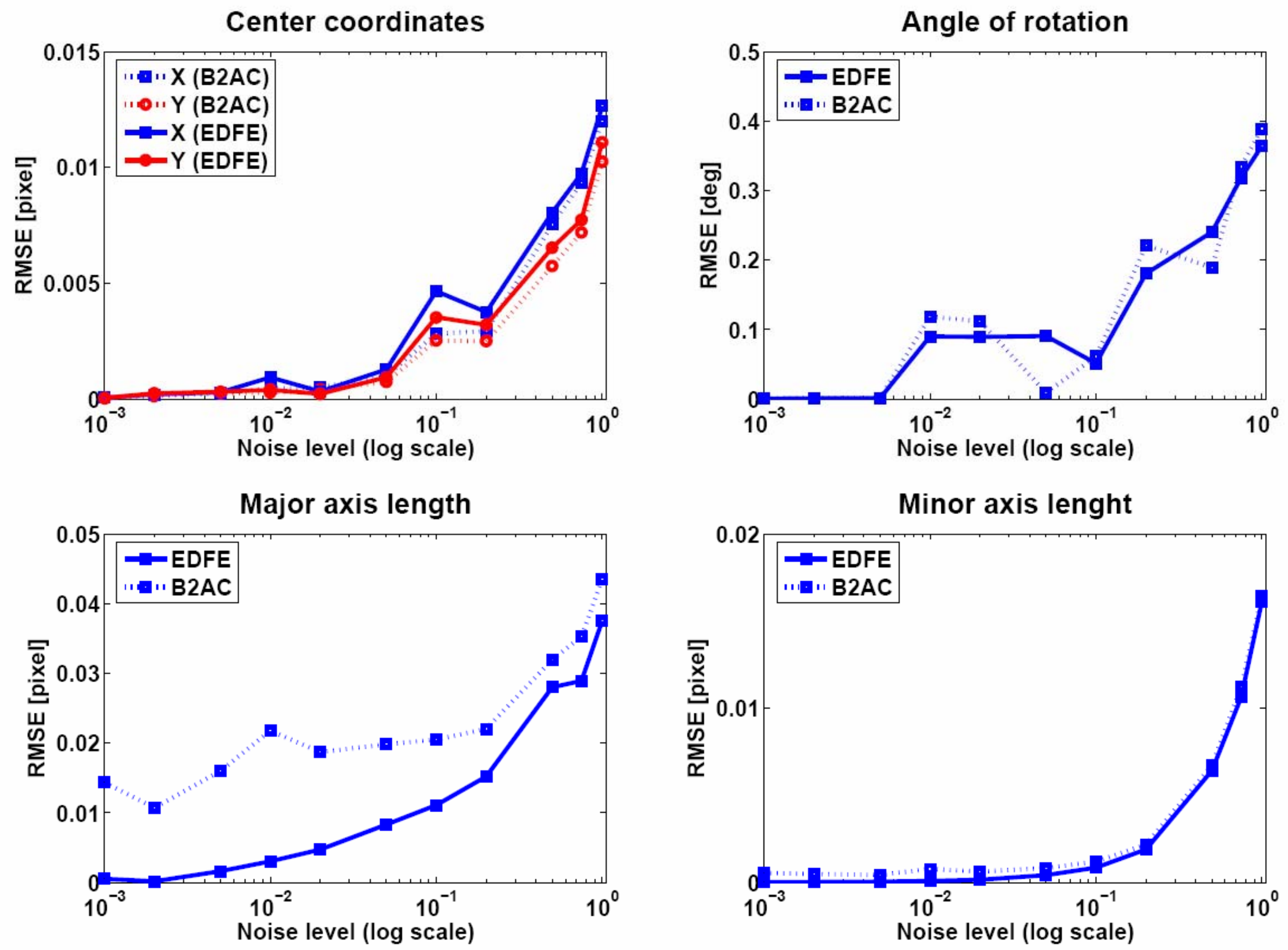
Soluzione

- Se necessario introduco piccole quote di rumore su dati e ripeto più volte il calcolo della soluzione
- La soluzione finale è la media delle soluzioni trovate in tutti i passi precedenti

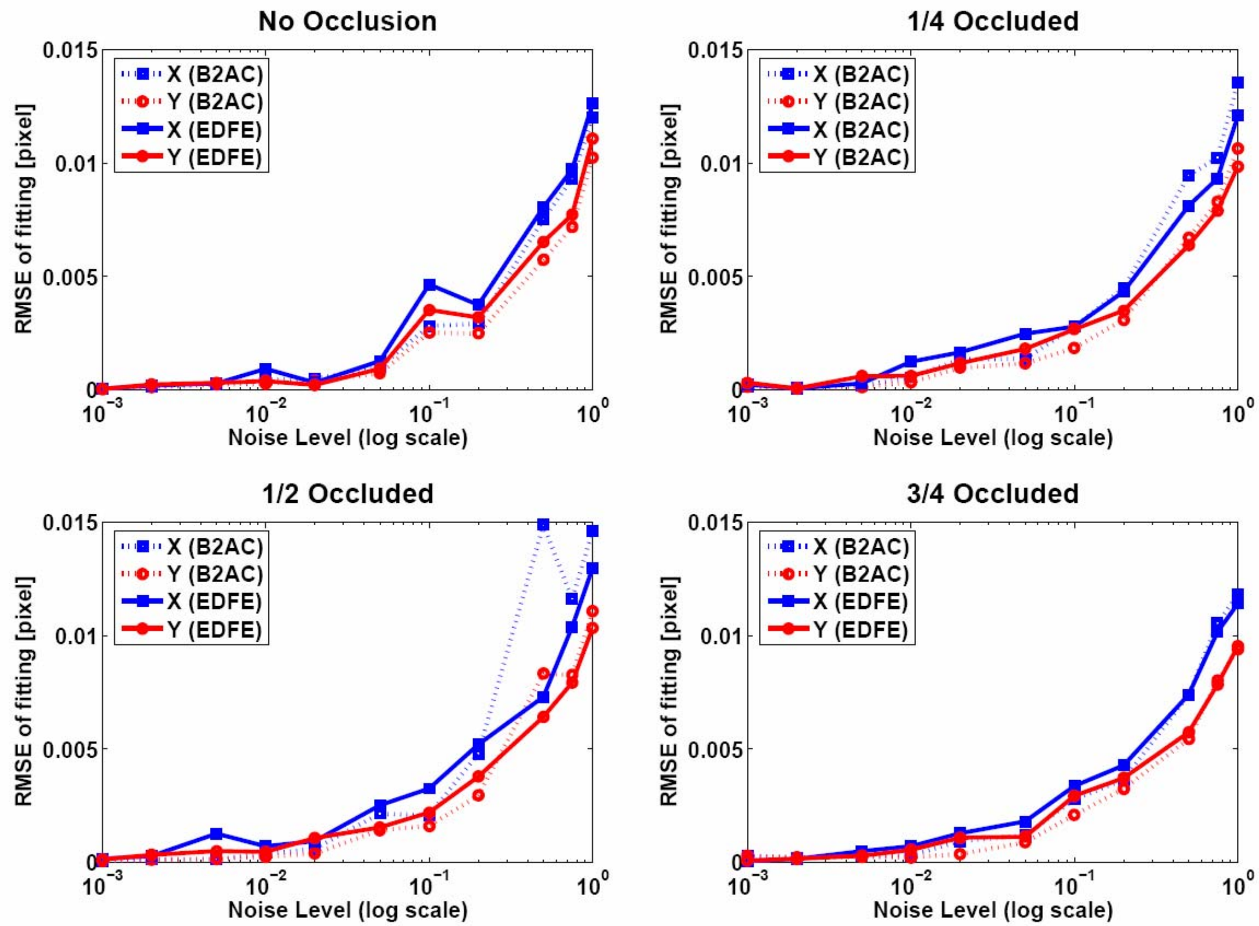
Algoritmo EDFE



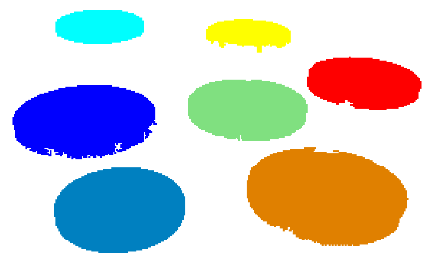
Risultati dati simulati



Risultati dati simulati



Risultati scena reale



#1
Center position [x,y] : **[69.9849 , 108.5173]**
Principal axis rotation (deg) : **-3.9001**
Major axis length (pixel) : **48.9353**
Minor axis length (pixel) : **24.1162**

#2
Center position [x,y] :**[94.5038 , 172.0267]**
Principal axis rotation (deg) : **-5.0783**
Major axis length (pixel) : **46.9288**
Minor axis length (pixel) : **29.5247**

#3
Center position [x,y] : **[80.4714 , 43.3702]**
Principal axis rotation (deg) : **-1.0271**
Major axis length (pixel) : **32.2435**
Minor axis length (pixel) : **12.6023**

#4
Center position [x,y] : **[184.4672 , 101.6561]**
Principal axis rotation (deg) : **3.1571**
Major axis length (pixel) : **43.3788**
Minor axis length (pixel) : **21.7222**

#5
Center position [x,y] : **[185.5677 , 49.3819]**
Principal axis rotation (deg) : **3.2145**
Major axis length (pixel) : **30.5327**
Minor axis length (pixel) : **11.9083**

#6
Center position [x,y] : **[240.2718 , 162.1727]**
Principal axis rotation (deg) : **9.2962**
Major axis length (pixel) : **57.4923**
Minor axis length (pixel) : **34.0158**

#7
Center position [x,y] : **[266.3417 , 83.2969]**
Principal axis rotation (deg) : **5.7453**
Major axis length (pixel) : **41.4882**
Minor axis length (pixel) : **18.1594**

Maggiori dettagli

International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 20, No. 6 (2006) 939–953

World Scientific
www.worldscientific.com

ENHANCED DIRECT LEAST SQUARE FITTING OF ELLIPSES

ELISEO STEFANO MAINI

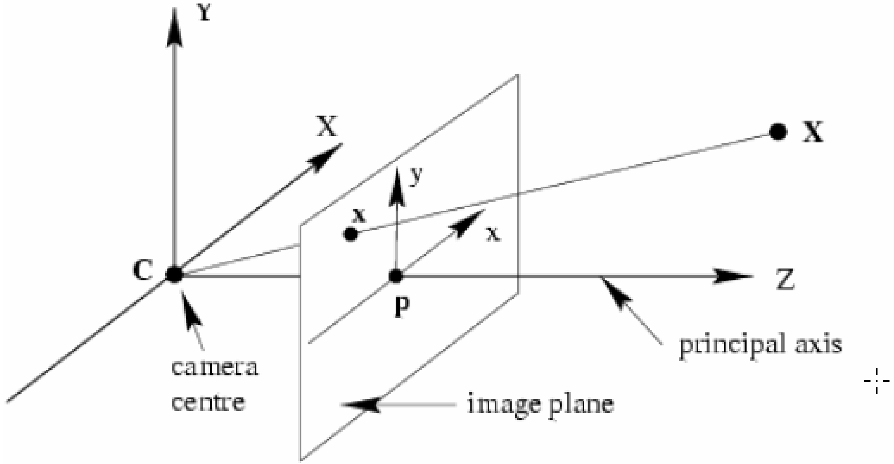
ARTS Lab — Scuola Superiore Sant' Anna, Polo Sant' Anna Valdera
Viale R. Piaggio, 34 — 56025 Pontedera, Italy
es.maini@ieee.org

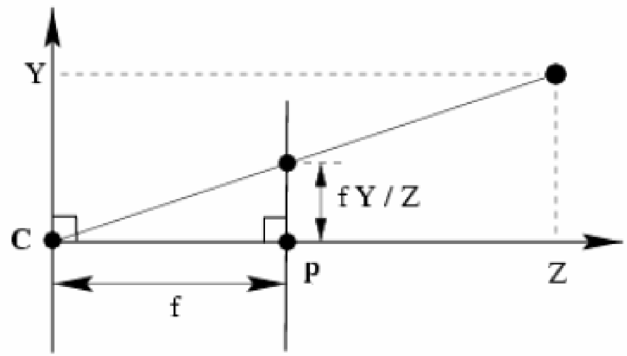
Visione 3D

Spazio 3D e piano immagine

- Proiezione prospettica (già vista!)

$$\begin{cases} x' = f'\dfrac{x}{z} \\ y' = f'\dfrac{y}{z} \end{cases}$$



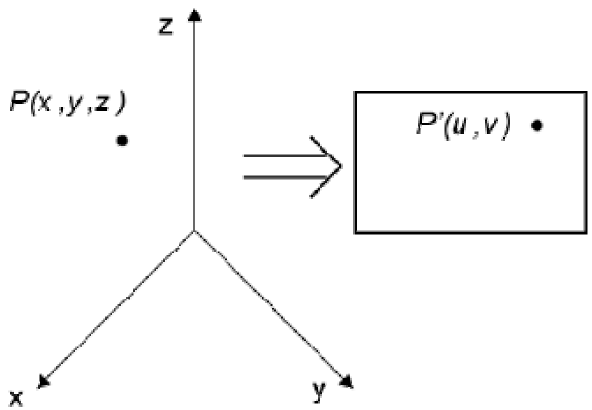


Un punto nello spazio **X**=(x,y,z) è mappato in un punto sul piano dell'immagine corrispondente all'intersezione della retta passante per il punto **X** ed il centro di proiezione con il piano dell'immagine.

$$(X,Y,Z)^T \text{ a } (fX/Z, fY/Z)$$

Mapping punti scena – punti immagine

- Ci si chiede se sia possibile operare una trasformazione inversa, ovvero a partire dai punti di un'immagine, ricavare informazioni sui punti corrispondenti nello spazio.



Risposta SI …. però:

- La proiezione operata da una telecamera comporta un'inevitabile perdita d'informazione: infatti il passaggio da un punto dello spazio tridimensionale ad un punto dell'immagine bidimensionale, prevede la perdita di una coordinata, quella che tiene conto della profondità.
- Per cui in un ipotetico passaggio inverso si ha il problema di recuperare l'informazione relativa a questa coordinata "smarrita".
- Questo problema viene superato da una tecnica nota come *stereopsi*.

Ricostruzione 3D

$$x_1 = \frac{(x-d)f}{f-z} \qquad (f-z)x_1 = (x-d)f$$

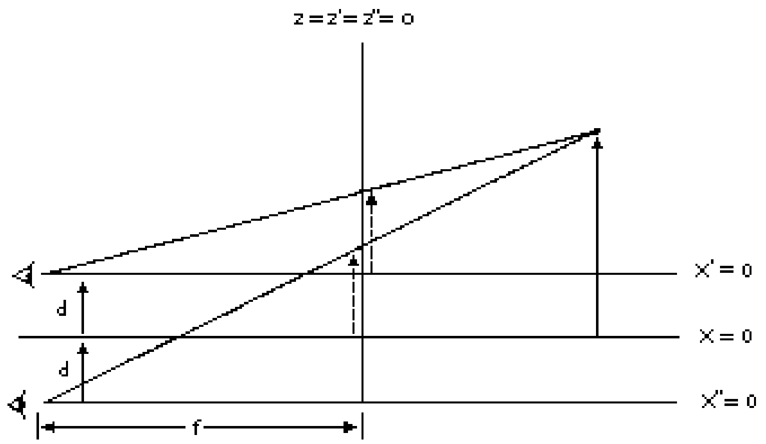
$$x_2 = \frac{(x+d)f}{f-z} \qquad (f-z)x_2 = (x-d)f$$

$$(f-z)(x_2 - x_1) = 2df$$

$$z = f - \frac{2df}{x_2 - x_1}$$

Disparità

Sistema binoculare non convergente

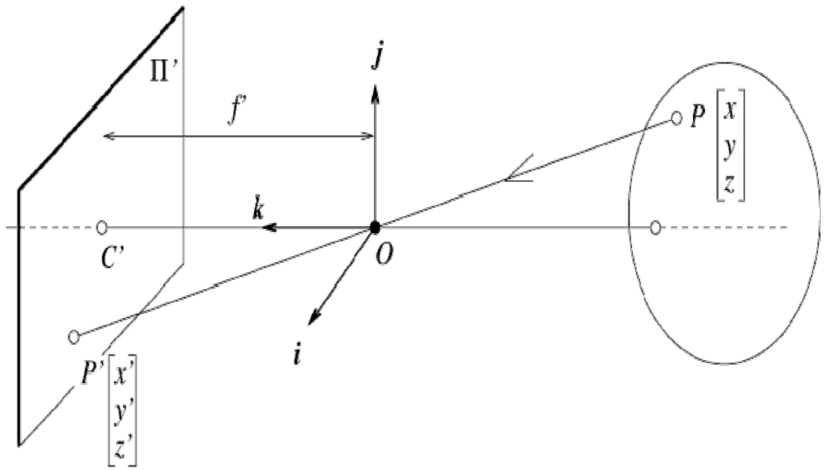


- la calibrazione è il punto di partenza obbligato, e quanto più preciso sarà il processo che ricava i parametri della telecamera, quanto più robusto ed efficiente sarà il funzionamento del sistema finale.

La calibrazione della telecamera

Tre diversi sistemi di riferimento:

1. il sistema di riferimento 3d detto anche *sistema mondo*;
2. il sistema di riferimento standard 3d della telecamera, centrato in *C*;
3. il sistema di riferimento 2d per l'immagine.



- X = (*X*; *Y*; *Z*) coordinate del punto *P* nel sistema mondo;
- XC = (*XC*; *YC*; *ZC*) coordinate del punto *P* nel sistema della camera;
- x = (*x*; *y*) coordinate del punto *P* nel sistema dell'immagine;
- w = (*u*; *v*) coordinate del punto *P* nel sistema dell'immagine digitale.

Passi fondamentali

- Identificare la trasformazione:

  Sistema mondo $\Rightarrow$ sistema della telecamera (rototraslazione)

- Identificare la trasformazione:

  Sistema di della telecamera $\Rightarrow$ sistema dell'immagine (eq. prospettica)

- Identificare la trasformazione:

  Coordinate dell'immagine $(x; y)$ $\Rightarrow$ coordinate in pixel $(u; v)$

Analiticamente

- Passo 1 $\hat{\mathbf{X}}_C = \mathbf{P}_R\hat{\mathbf{X}}$

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Passo 2 $\hat{\mathbf{x}} = \mathbf{P}_P\hat{\mathbf{X}}_C$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-f}{Z_C} & 0 & 0 & 0 \\ 0 & \frac{-f}{Z_C} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix}$$

- Passo 3 $\hat{\mathbf{w}} = \mathbf{P}_C\hat{\mathbf{x}}$

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} sx \\ sy \\ s \end{bmatrix}$$

- X = (*X; Y; Z*) coordinate del punto *P* nel sistema mondo;
- XC = (*XC; YC; ZC*) coordinate del punto *P* nel sistema della camera;
- x = (*x; y*) coordinate del punto *P* nel sistema dell'immagine;
- w = (*u; v*) coordinate del punto *P* nel sistema dell'immagine digitale.

$$\hat{\mathbf{w}} = \mathbf{P}_C\mathbf{P}_P\mathbf{P}_R\hat{\mathbf{X}}$$

In pratica

$\hat{\mathrm{w}} = \mathbf{P}_C \mathbf{P}_P \mathbf{P}_R \hat{\mathrm{X}}$ Fattorizzazione $P_C P_P$

$$\hat{\mathrm{w}} = \mathbf{P}_{CP} \mathbf{P}_R \hat{\mathrm{X}} = \begin{bmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \hat{\mathrm{X}}$$

Pongo inoltre

$$\mathbf{P}_{CP} = \mathbf{K}\,[\mathbf{I} \mid \mathbf{0}] \qquad \text{con } \mathbf{K} = \begin{bmatrix} -fk_u & 0 & u_0 \\ 0 & -fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{P}_R = \begin{bmatrix} \mathrm{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \qquad \text{con } \mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \qquad \text{e } \mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

Allora risulta

$$\hat{\mathrm{w}} = \mathbf{K}\,[\mathbf{I} \mid \mathbf{0}] \begin{bmatrix} \mathrm{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \hat{\mathrm{X}} = \mathbf{K}\,[\mathrm{R} \mid \mathbf{t}]\,\hat{\mathrm{X}} = \mathbf{P}\hat{\mathrm{X}}$$

P è detta

MATRICE di PROIEZIONE PROSPETTICA

Matrice di Proiezione Prospettica (MPP)

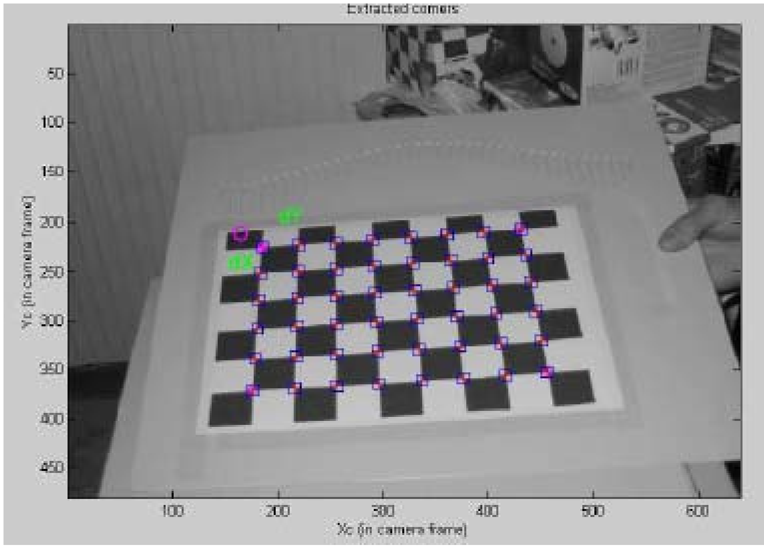
- *MPP* rappresenta il modello geometrico della telecamera.
- E' una matrice 3x4, ma dipende da 11 parametri indipendenti.
- La matrice [R|t] incorpora sia la rotazione che la traslazione del sistema mondo rispetto a quello della telecamera. Si dice che tale matrice contiene i ***parametri estrinseci*** della telecamera.
- Nella matrice K si puμo porre $a_u = -f\,k_u$ e $a_v = -f\,k_v$. Si tratta della **lunghezza focale espressa in pixel** orizzontali e verticali.
- Queste due grandezze, assieme a u_0 e v_0 rappresentano i ***parametri intrinseci*** della telecamera.
- Questi due insiemi di parametri caratterizzano completamente la telecamera.

La *calibrazione* è quel processo che permette di determinare le matrici dei parametri intrinseci ed estrinseci di una telecamera.

Calibrazione pratica

- E' necessario munirsi di un oggetto di calibrazione (geometria nota)
- una buona scelta è rappresentata da una scacchiera.
  - quadrati di lato noto
  - colore alterno (estrazione dei contorni semplice)

- In teoria, dato che *MPP* dipende da 11 parametri, sarebbero necessari 6 soli punti per poter calibrare la telecamera,
  - nella pratica ne sono disponibili molti di più.



Calibrazione pratica

1. Ottenere le coordinate dei corner nell'immagine digitale
2. Creare una tabella di corrispondenze

Punto	Sistema mondo	Sistema immagine
1	(X_1, Y_1, Z_1)	(u_1, v_1)
2	(X_2, Y_2, Z_2)	(u_2, v_2)
...	...	...
n	(X_n, Y_n, Z_n)	(u_n, v_n)

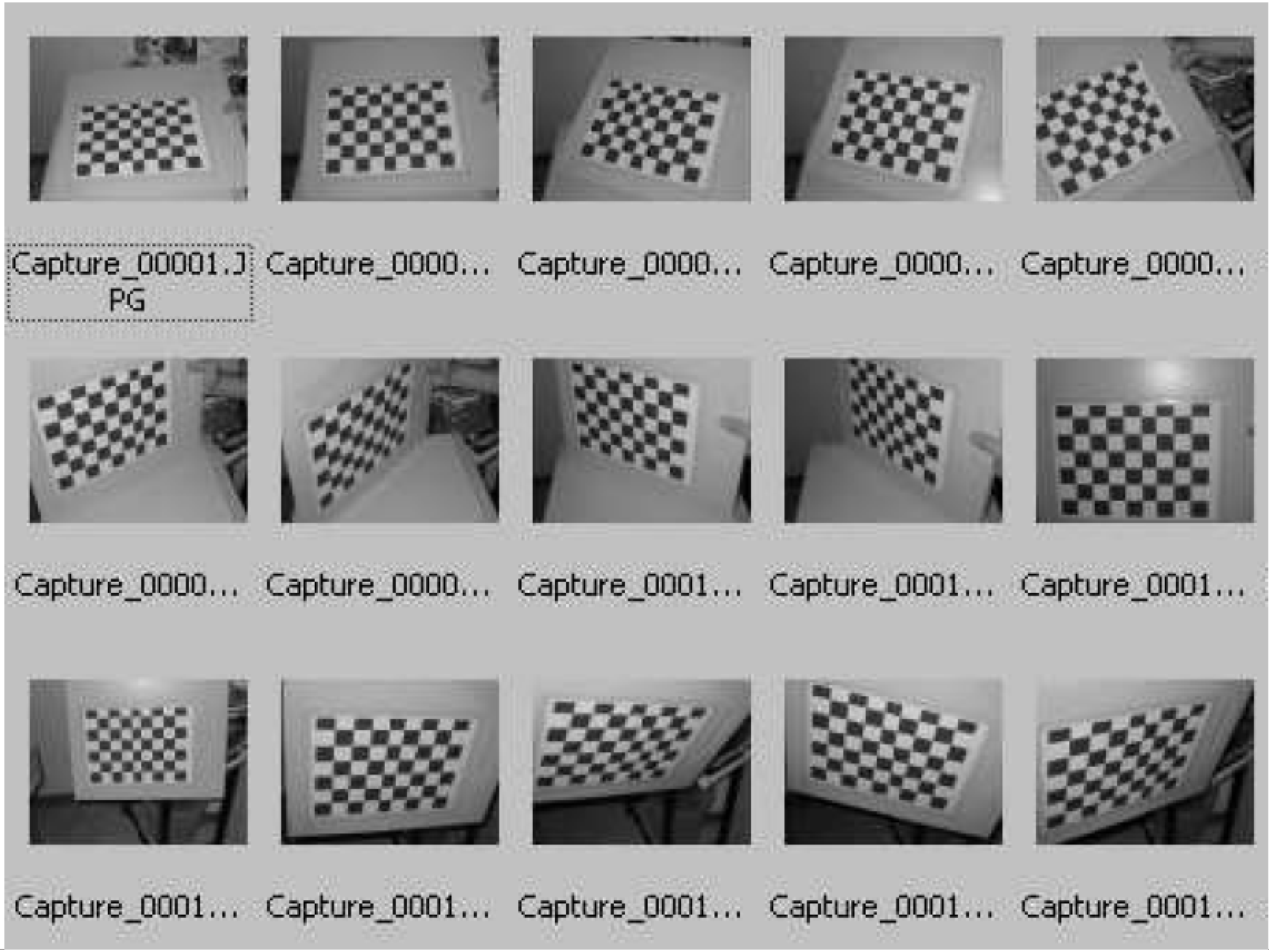
3. Si sfrutta la relazione $\hat{\mathbf{w}}_i = \mathbf{P}\hat{\mathbf{X}}_i = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \hat{\mathbf{X}}_i$

   da cui

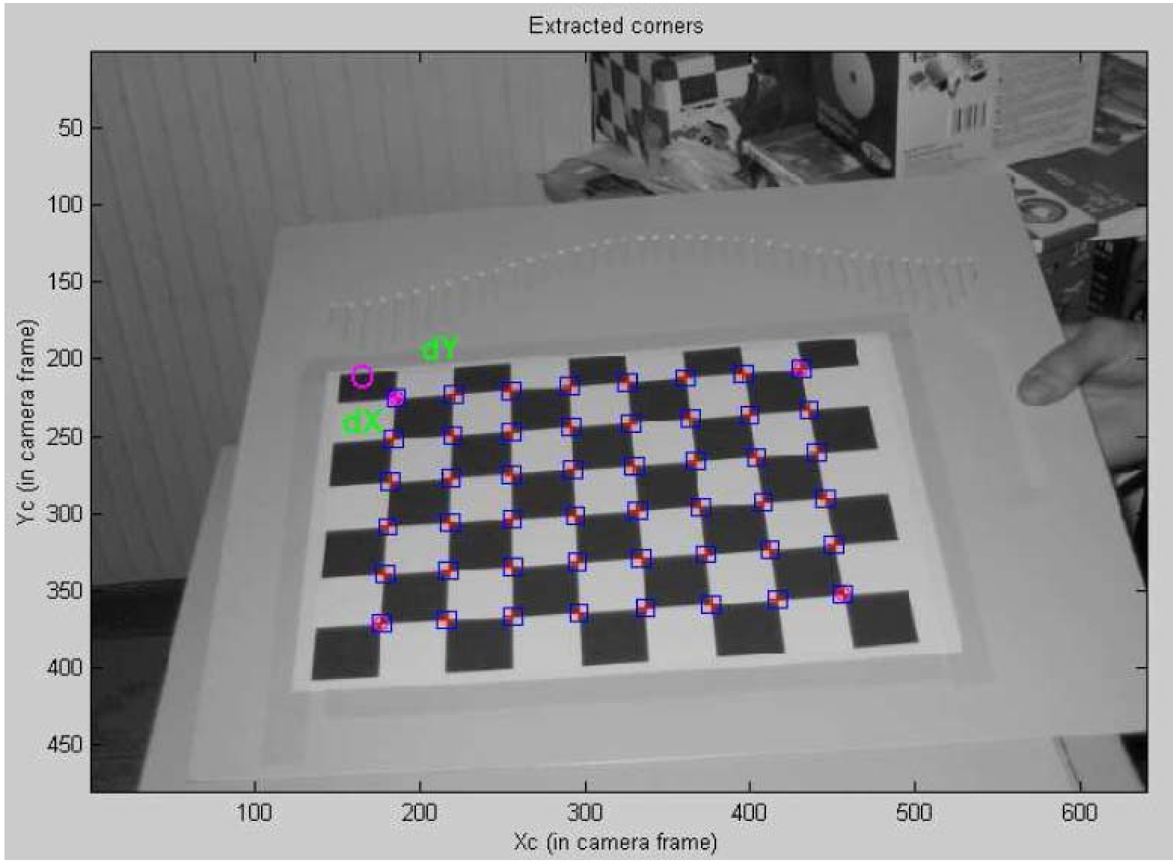
$$u_i = \frac{p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}} \qquad \text{e} \qquad v_i = \frac{p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}$$

4. Per ogni punto della tabella si scrive una coppia di equazioni.
5. Dato che i punti sono in numero maggiore di 6 ci si trova a dover risolvere un sistema lineare sovradeterminato, (le equazioni sono in numero maggiore delle incognite).
6. La soluzione può essere ricavata utilizzando il metodo dei *minimi quadrati*.
7. Metodi iterativi per il raffinamento della soluzione $\min_P \sum((u_i - \bar{u}_i)^2, (v_i - \bar{v}_i)^2)$
8. Decomposizione QR della matrice (intrinseci ed estrinseci)

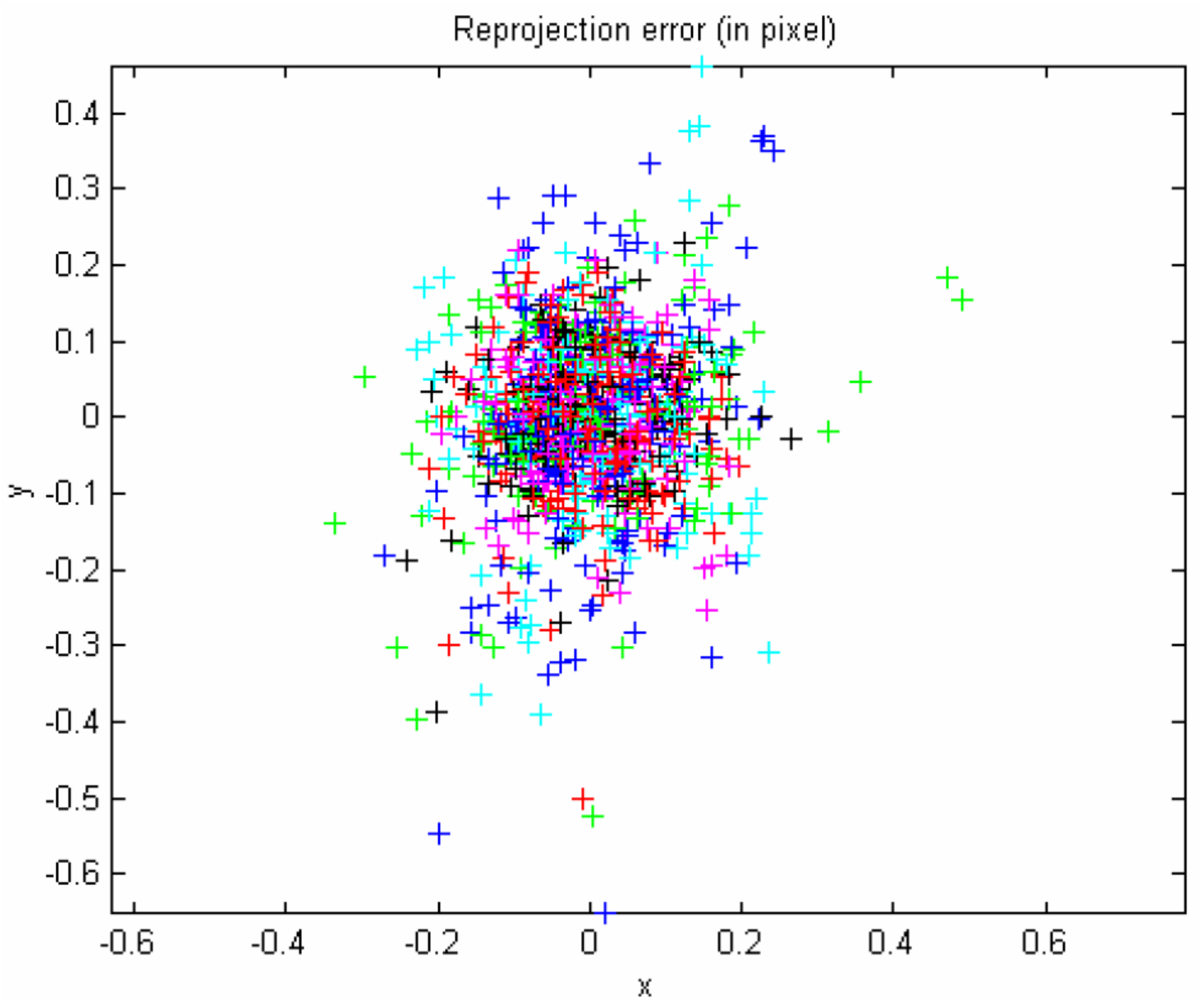
Sequenza immagini - Robustezza



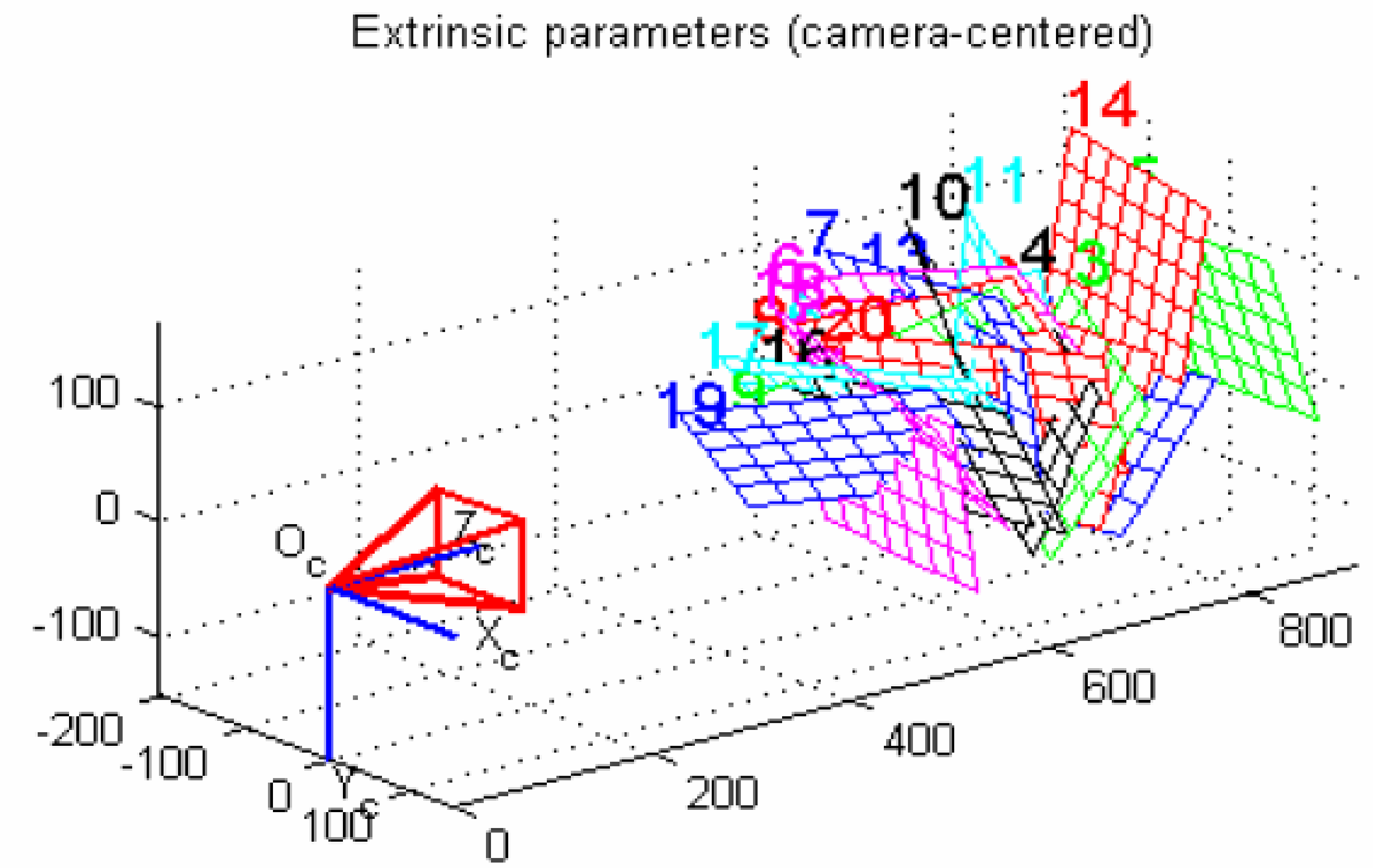
Estrazione contorni



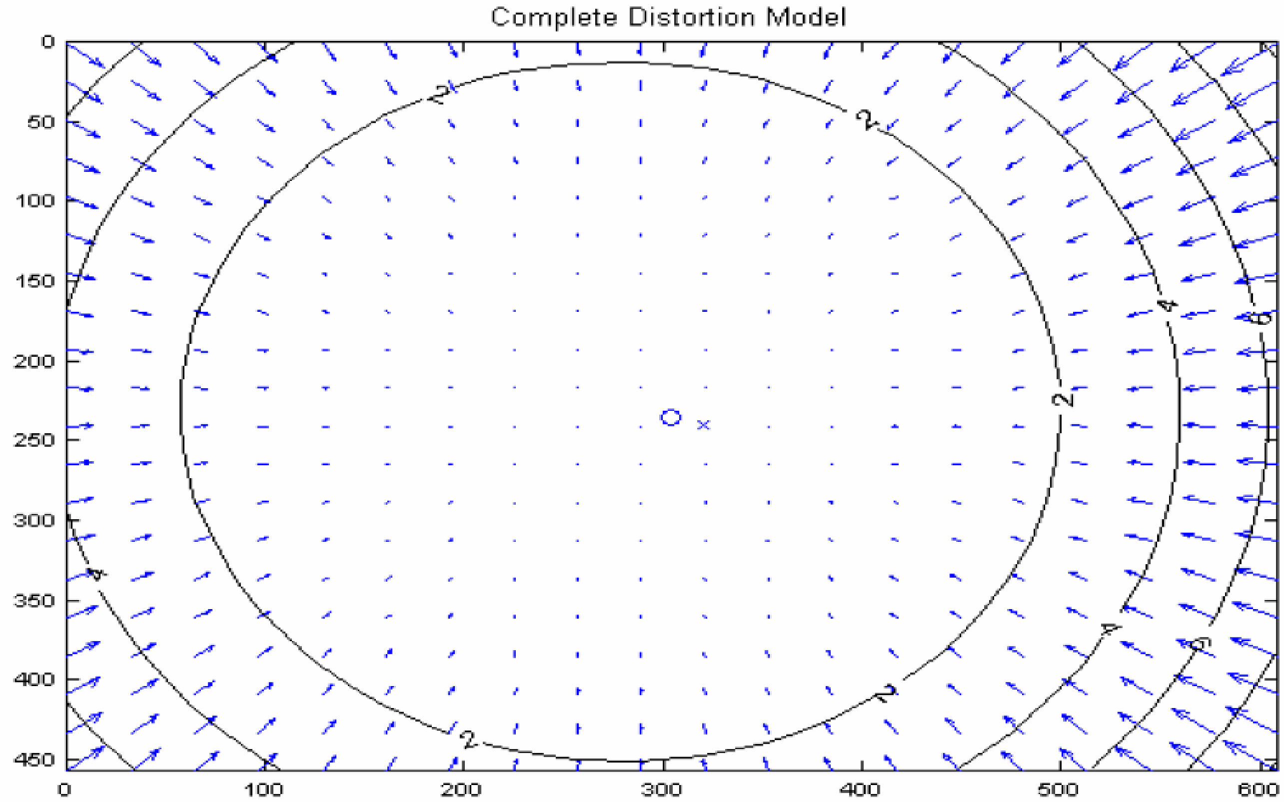
Errore di riproiezione



Parametri estrinseci



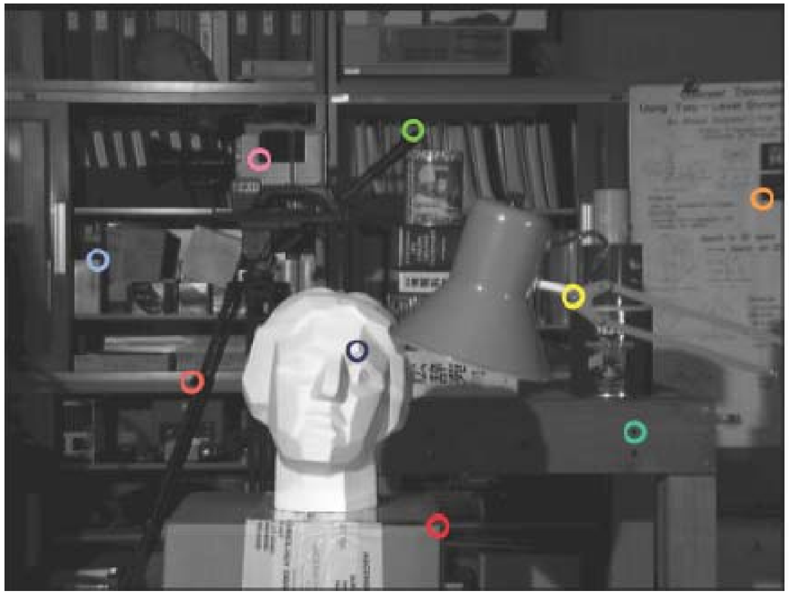
Stima distorsione

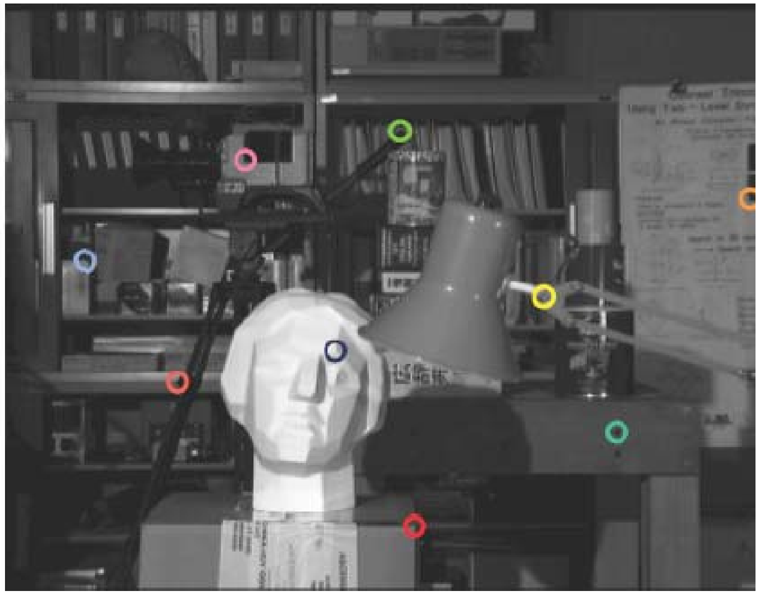


Pixel error	= [0.0976, 0.1119]	
Focal Length	= (948.228, 947.537)	+/- [1.91, 1.951]
Principal Point	= (302.509, 234.682)	+/- [3.906, 3.021]
Skew	= 0.001167	+/- 0.0005497
Radial coefficients	= (-0.1854, 0.263, 0)	+/- [0.01487, 0.175, 0]
Tangential coefficients	= (-0.0005125, -0.003946)	+/- [0.0006858, 0.000806]

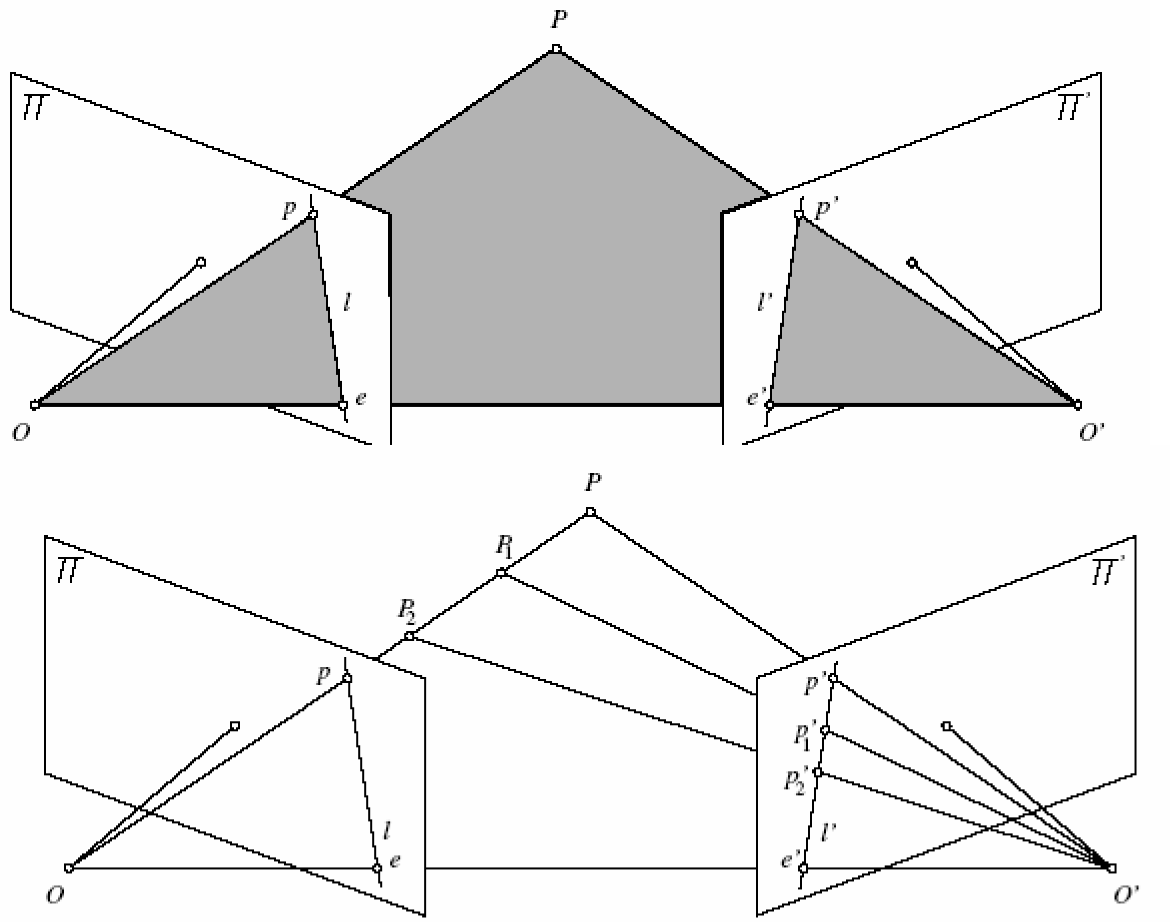
2 Telecamere : problema del matching

- Date due telecamere trovare la corrispondenza di punti omologhi su due diverse immagini



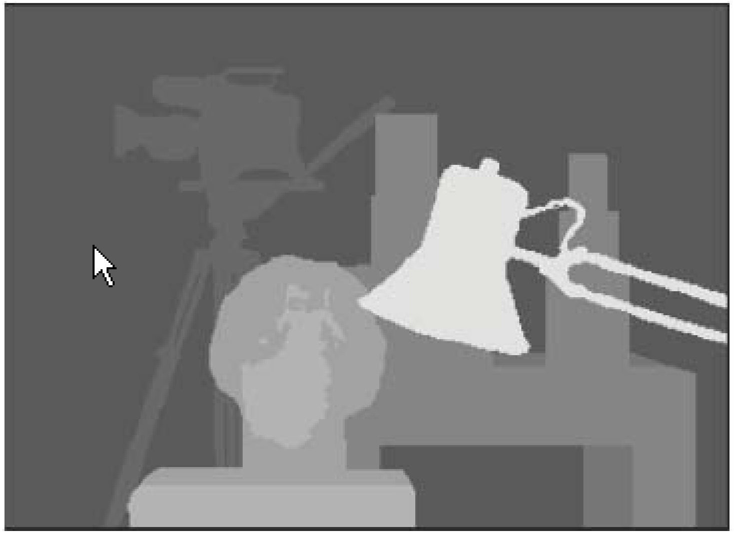


2 Telecamere – Geometria epipolare



Vincolo epipolare: dato un sistema stereo calibrato l'insieme dei punti di matching di p è vincolato a giacere sulla linea epipolare l'

Mappa delle disparità



Ricostruzione stereo

1. Calibrare la I telecamera
2. Calibrare la II telecamera
3. Calcolare la rototraslazione tra i sisteme di riferimento delle 2 telecamere
4. Determinare le corrispondenze tra punti (vincolo epipolare)
5. Costruire una mappa delle disparità
6. Ricostruire la posizione 3D dei punti di interesse usando la disparità e i parametri delle telecamere