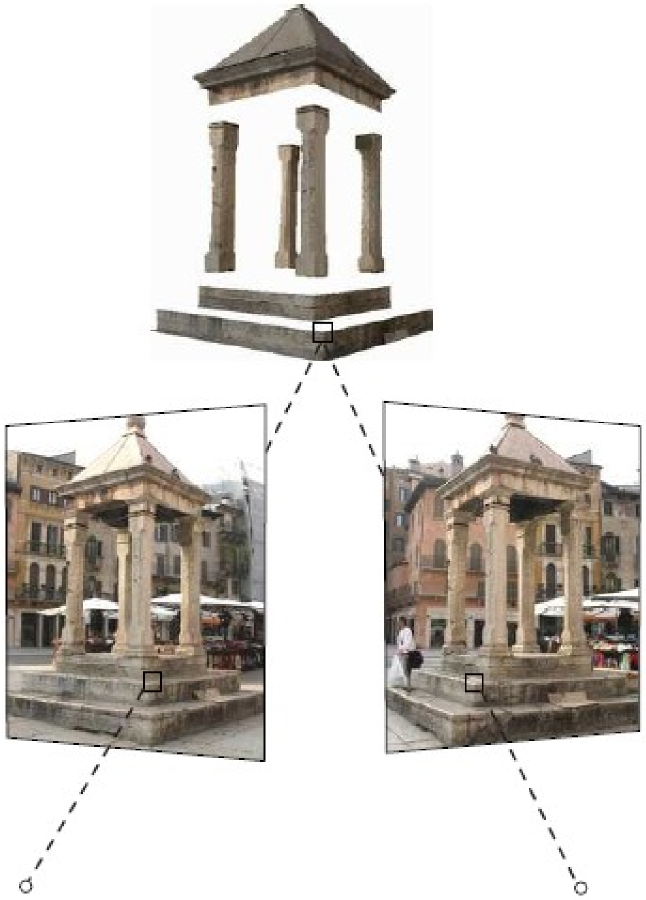
Corso di Visione Artificiale

Stereopsi

Samuel Rota Bulò

Introduzione

- La stereopsi è il processo di inferenza della struttura 3D da una coppia di immagini di una stessa scena catturate da posizioni diverse.
- Due sottoproblemi:
  - calcolo delle corrispondenze
  - triangolazione



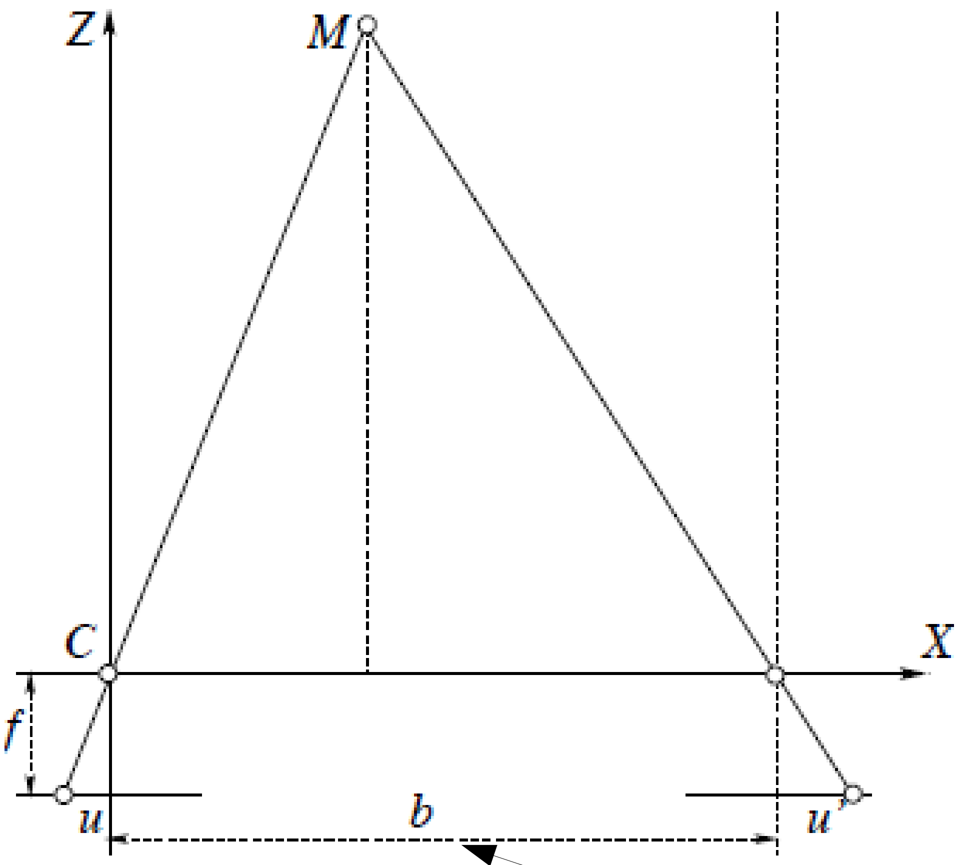
Caso di triangolazione semplice

- Camere parallele e allineate
- Sistema di riferimento solidale con la camera sinistra

$$M = (x, y, z)^\top$$

$$\begin{cases} \dfrac{f}{z} = \dfrac{-u}{x} \\ \dfrac{f}{z} = \dfrac{-u'}{x - b} \end{cases}$$

$$z = \frac{bf}{u' - u}$$



linea di base legata alla disparità

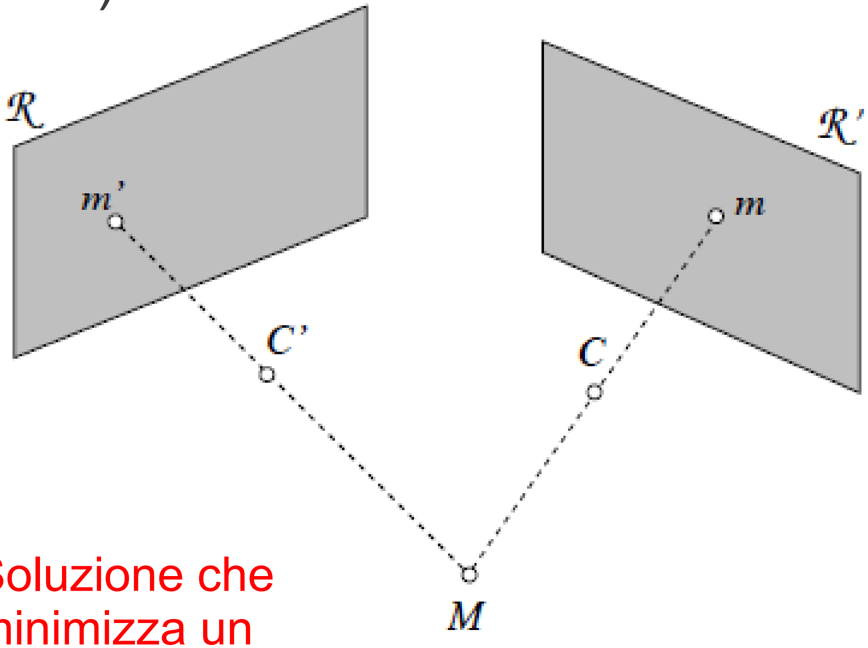
Triangolazione

- Consideriamo la triangolazione nel caso in cui conosciamo:
    - coordinate di 2 punti coniugati (generati dallo stesso punto 3D)
    - matrici di proiezione prospettica (MPP) delle due fotocamere.

$$\mathbf{m} \simeq P\mathbf{M} \qquad \mathbf{m} = \left[u, v, 1\right]^\top$$

$$\mathbf{m}' \simeq P'\mathbf{M} \qquad \mathbf{m}' = \left[u', v', 1\right]^\top$$

$$\begin{bmatrix} (\mathbf{p}_1 - u\mathbf{p}3)^\top \\ (\mathbf{p}_2 - v\mathbf{p}_3)^\top \\ (\mathbf{p}'_1 - u'\mathbf{p}'_3)^\top \\ (\mathbf{p}'_1 - u\mathbf{p}'_3)^\top \end{bmatrix} \mathbf{M} = \mathbf{0}_{4\times 1}$$



Soluzione che
minimizza un
errore algebrico !

- Possiamo trovare **M** a meno di un fattore di scala, trovando il nucleo della matrice 4x4 dei coefficienti, o imponendo vincoli per evitare la soluzione banale **M=0**.

Triangolazione

- La soluzione che minimizza l'errore algebrico è facile da calcolare (lineare), ma non ha un significato geometrico.
- Una maggior accuratezza si ottiene minimizzando l'errore geometrico:

$$\varepsilon(\mathbf{M}) = \left\| \begin{bmatrix} u \\ v \end{bmatrix} - \begin{bmatrix} \frac{\mathbf{p}_1^\top \mathbf{M}}{\mathbf{p}_3^\top \mathbf{M}} \\ \frac{\mathbf{p}_2^\top \mathbf{M}}{\mathbf{p}_3^\top \mathbf{M}} \end{bmatrix} \right\|^2 + \left\| \begin{bmatrix} u' \\ v' \end{bmatrix} - \begin{bmatrix} \frac{\mathbf{p}_1'^\top \mathbf{M}}{\mathbf{p}_3'^\top \mathbf{M}} \\ \frac{\mathbf{p}_2'^\top \mathbf{M}}{\mathbf{p}_3'^\top \mathbf{M}} \end{bmatrix} \right\|^2$$

- Tuttavia questo problema è difficile in quanto non lineare e non convesso (possibili minimi locali).

Quanto visto si generalizza al caso di più fotocamere.

Geometria epipolare

- Studiamo la relazione tra la posizione di un punto in un'immagine e la posizione che un suo coniugato può avere nella seconda.

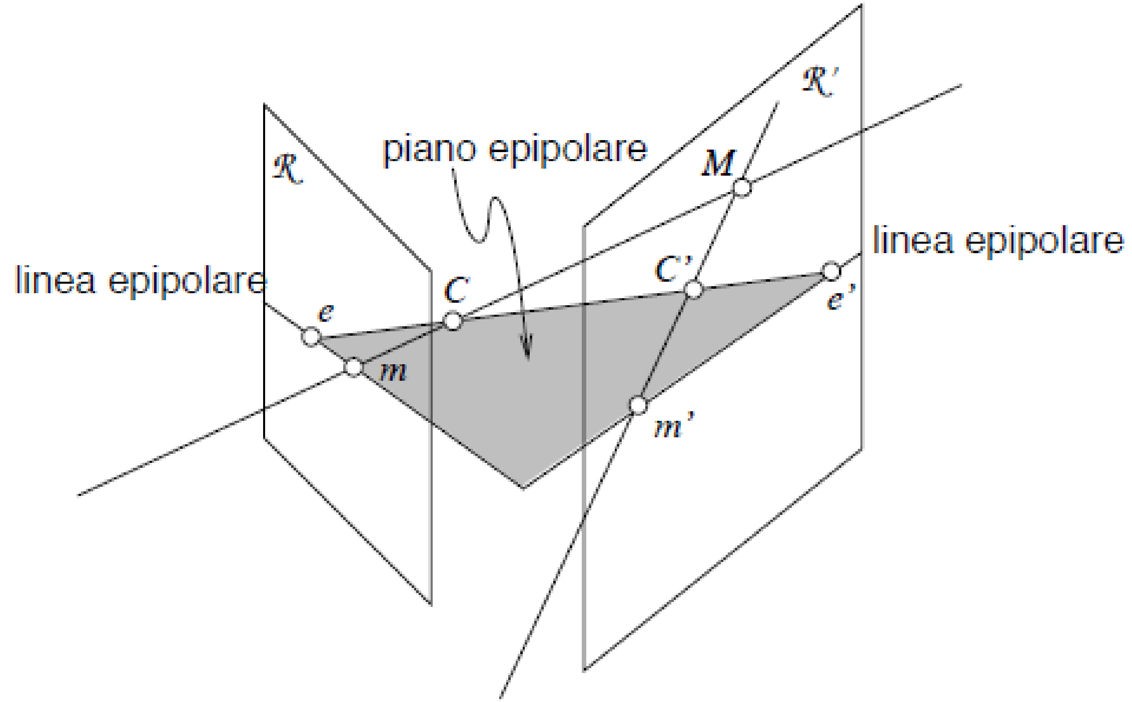
$$\begin{cases} \mathbf{m} \simeq P\mathbf{M} = [Q|\mathbf{q}]\mathbf{M} \\ \mathbf{m}' \simeq P'\mathbf{M} = [Q'|\mathbf{q}']\mathbf{M} \end{cases}$$

$$\mathbf{M} = \mathbf{C} + \lambda \begin{bmatrix} Q^{-1}\mathbf{m} \\ 0 \end{bmatrix}$$

epipolo

$$P'\mathbf{C} = P' \begin{bmatrix} -Q^{-1}\mathbf{q} \\ 1 \end{bmatrix} =$$

$$\mathbf{q}' - Q'Q^{-1}\mathbf{q} \triangleq \mathbf{e}'$$



Matrice fondamentale

- La matrice fondamentale contiene tutta l'informazione relativa alla geometria epipolare.

$$P'\mathbf{M} \simeq P'\mathbf{C} + \lambda P' \begin{bmatrix} Q^{-1}\mathbf{m} \\ 0 \end{bmatrix}$$

$$\mathbf{m}' \simeq \lambda Q' Q^{-1}\mathbf{m} + \mathbf{e}'$$

$$\mathbf{m}'^{\top}[\mathbf{e}']_{\times}\mathbf{m}' \simeq \lambda \mathbf{m}'^{\top}[\mathbf{e}']_{\times} Q' Q^{-1}\mathbf{m} + \mathbf{m}'^{\top}[\mathbf{e}']_{\times}\mathbf{e}'$$

$$0 = \mathbf{m}'^{\top}[\mathbf{e}']_{\times} Q' Q^{-1}\mathbf{m}$$

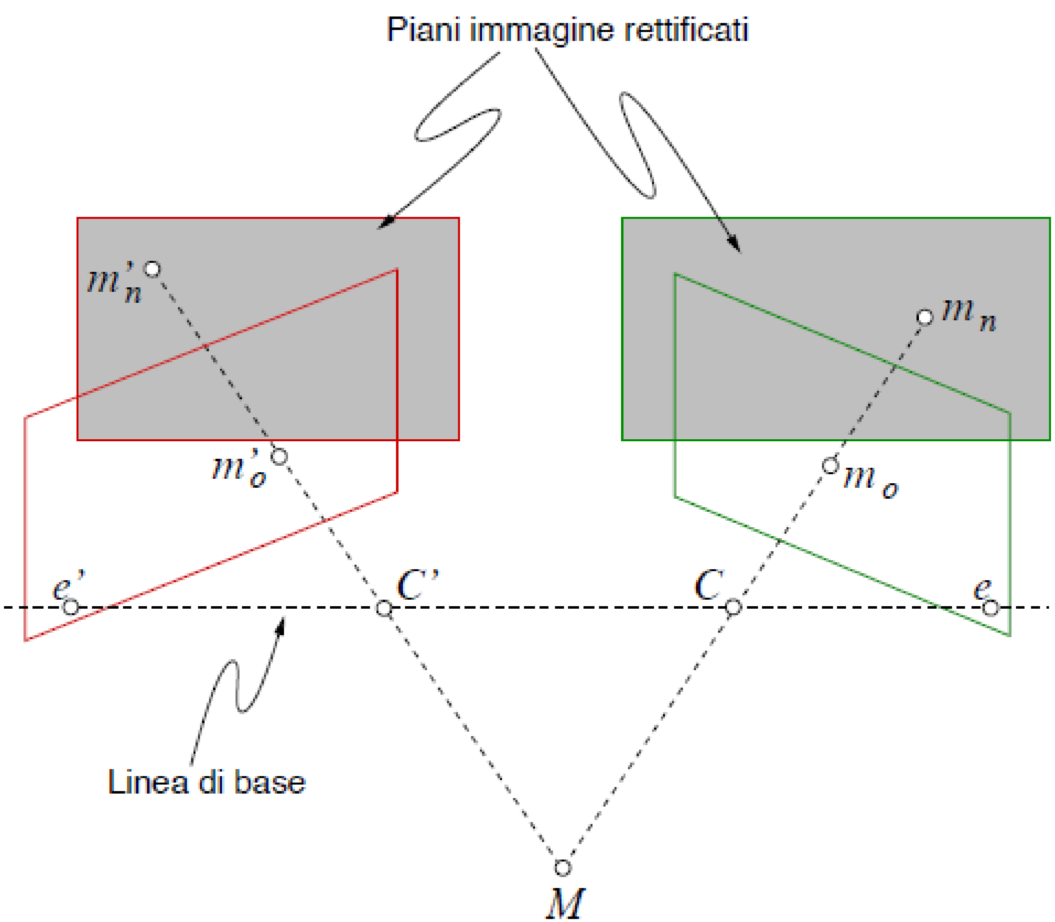
Equazione di Longuet-Higgins

$$\mathbf{m}'^{\top} F \mathbf{m} = 0$$

$$F = [\mathbf{e}']_{\times} Q' Q^{-1}$$

Rettificazione delle MPP

- La rettificazione delle MPP è il processo che trasforma le MPP delle due fotocamere in modo da rendere i piani focali coplanari
- Vogliamo inoltre che punti coniugati abbiano le stesse coordinate verticali



Rettificazione delle MPP

MPP originali: P_o, P'_o

$$P_n = K[R| - R\tilde{\mathbf{C}}]$$

$$P'_n = K[R| - R\tilde{\mathbf{C}}']$$

$$R = \begin{bmatrix} \mathbf{r}_1{}^\top \\ \mathbf{r}_2{}^\top \\ \mathbf{r}_3{}^\top \end{bmatrix}$$

- Dopo la rettificazione è come se avessimo immagini scattate da 2 fotocamere con stessi parametri intrinseci e orientamento. L'unica differenza è la posizione dei centri ottici che rimangono gli stessi delle vecchie camere.
- La matrice dei parametri intrinseci viene fissata arbitrariamente

asse X parallelo alla linea di base

$$\mathbf{r}_1 = \frac{\tilde{\mathbf{C}}' - \tilde{\mathbf{C}}}{\|\tilde{\mathbf{C}}' - \tilde{\mathbf{C}}\|}$$

asse Y perpendicolare all'asse X e a un versore k (utilizziamo l'asse Z della vecchia fotocamera)

$$\mathbf{r}_2 = \mathbf{k} \times \mathbf{r}_1$$

asse Z perpendicolare al piano XY

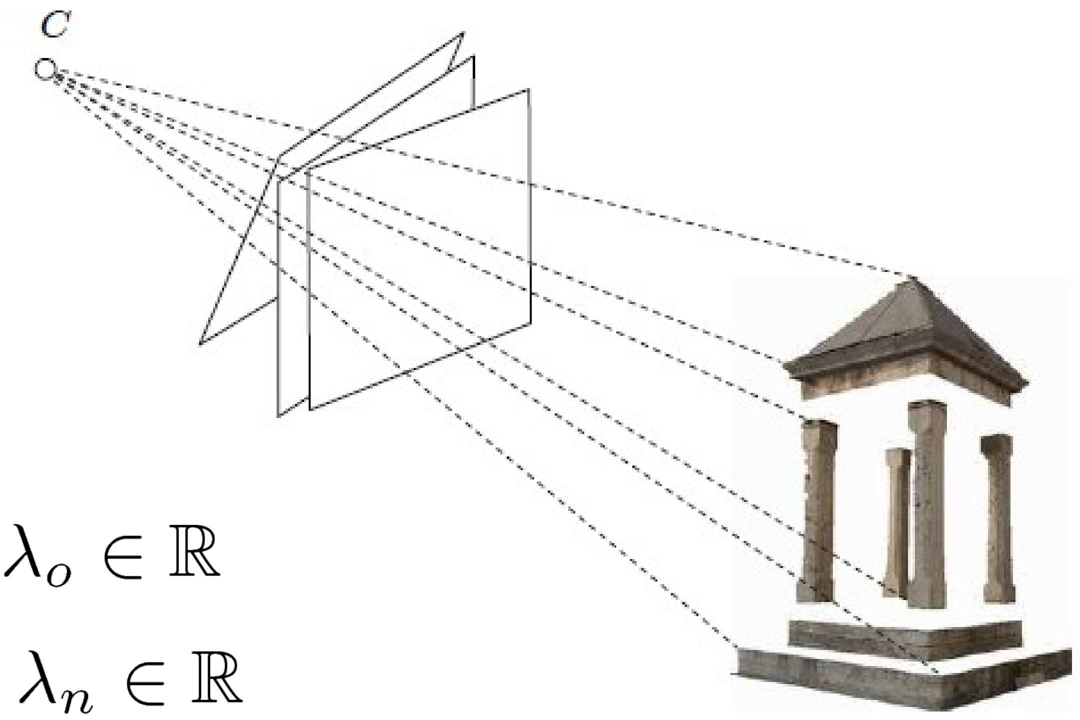
$$\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$$

Rettificazione delle MPP

$$P_o = [Q_o | \mathbf{q}_o]$$

$$P_n = [Q_n | \mathbf{q}_n]$$

$$\begin{cases} \mathbf{m}_o \simeq P_o \mathbf{M} \\ \mathbf{m}_n \simeq P_n \mathbf{M} \end{cases}$$



$$\begin{cases} \tilde{\mathbf{M}} = \tilde{\mathbf{C}} + \lambda_o Q_o^{-1} \mathbf{m}_o & \lambda_o \in \mathbb{R} \\ \tilde{\mathbf{M}} = \tilde{\mathbf{C}} + \lambda_n Q_n^{-1} \mathbf{m}_n & \lambda_n \in \mathbb{R} \end{cases}$$

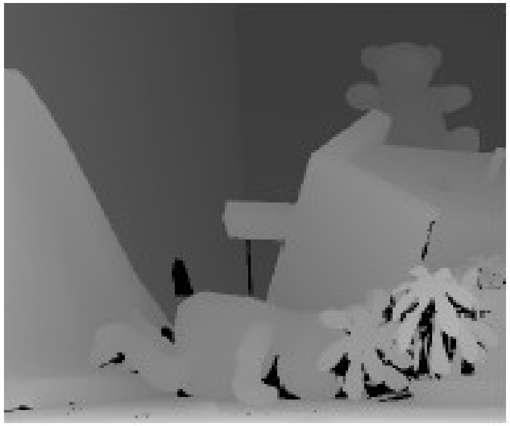
$$\mathbf{m}_n \simeq Q_n Q_o^{-1} \mathbf{m}_o$$

Rettificazione

Calcolo delle corrispondenze

- Il calcolo delle corrispondenze consiste nel trovare coppie di **punti coniugati** in due immagini, ovvero che sono proiezione dello stesso punto della scena.
- La disparità tra due punti coniugati è la loro differenza (vettore) se le due immagini fossero sovrapposte.
- L'assunzione che facciamo è che le due immagini non siano troppo diverse, quindi vi siano molte parti simili nelle due immagini.



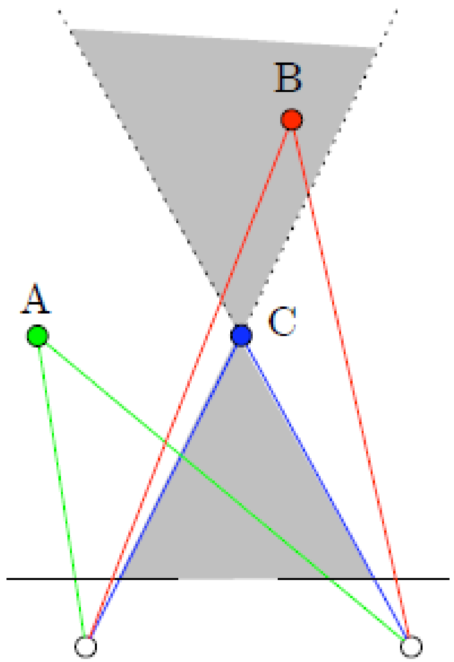


Difficoltà

- **False corrispondenze**: un punto dell'immagine può essere messo in corrispondenza a molti punti dell'altra immagine.
- **Occlusioni**: per un punto dell'immagine può non esistere il coniugato nell'altra.
- **Distorsione radiometrica**: intensità percepite dalle due fotocamere diverse a causa di superfici non perfettamente Lambertiane.
- **Distorsione prospettica**: a causa della proiezione prospettica, gli oggetti hanno forme diverse.

Vincoli

- **Somiglianza**: alcuni punti appaiono simili nelle due immagini.
- **Geometria epipolare**: il punto coniugato giace su una retta (epipolare).
- **Lisciezza**: lontano dai bordi la variazione in profondità varia lentamente.
- **Unicità**: un punto dell'immagine ha al più un punto coniugato.
- **Ordinamento monotono**: se m1 e m'1 sono coniugati allora se m2 giace a destra (sinistra) di m1, anche il coniugato m'2 di m2 giace a destra (sinistra) di m'1. (Non vale se M2 è nel cono d'ombra di M1)



Metodi di block matching

- Metodi locali. Consente ricostruzione densa.
- Astraggono un'immagine con piccoli blocchi rettangolari e cercano blocchi simili nell'altra mediante misure di somiglianza.
- Sfruttano il vincolo epipolare per restringere la ricerca per blocco alle rette epipolari.
- Problemi in corrispondenza di discontinuità di profondità. Possibili soluzioni sono
  - finestre adattive/eccentriche
  - metodi multirisoluzione
- Problemi in caso di occlusioni. Possibili soluzioni:
  - sfruttare vincolo di ordinamento
  - sfruttare vincolo destra-sinistra, ovvero se p viene accoppiato a p' effettuando la ricerca da I1 a I2, allora p' deve essere accoppiato a p effettuando la ricerca da I2 a I1.

Metodi basati su features

- Estraggono punti salienti dalle due immagini.
- Cercano di accoppiare le features sfruttando i loro descrittori per valutarne la similarità e i vincoli per escludere falsi accoppiamenti.
- Non consentono ricostruzione densa, ma sparsa. Vanno eventualmente combinati con metodi di ricostruzione densa.