



Il flusso ottico

Alberto Borghese
Laboratory of Motion Analysis, Virtual Reality (MAVR)



Il flusso ottico

Origina dalla psicologia sperimentale della visione (1950-1960).

PROBLEMA:

- Input al sistema visivo: una matrice bidimensionale di punti luminosi.
- Operazioni localizzate.
- Output sono unità: oggetti, persone nello spazio tridimensionale.

SOLUZIONI:

- Gestalt (Koffka, 1935) raggruppamento in una percezione unitaria di elementi che hanno caratteristiche simili.
- Flusso ottico (Gibson, 1950).

RICADUTE TECNOLOGICHE OGGIGIORNO:

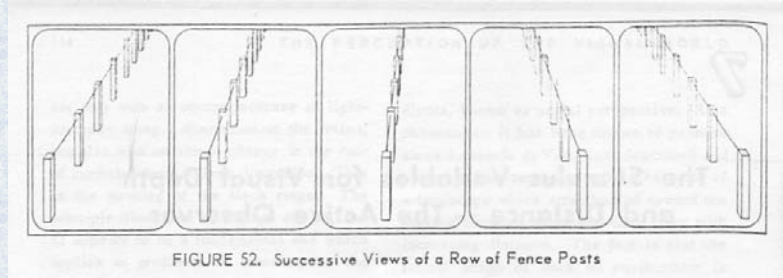
- Compressori MPEG.
- Sistemi di Visione Artificiale.



Da cosa deriva il flusso ottico?



- Variazioni di luminosità in immagini statiche => bordi => oggetti + posizione 3D (stereo).
- Variazioni di luminosità si possono ottenere anche se ci spostiamo noi davanti ad uno scenario fermo (variazioni temporali di luminosità).



In questo caso lo spostamento dell'immagine sulla retina è inversamente proporzionale alla distanza dei pali.

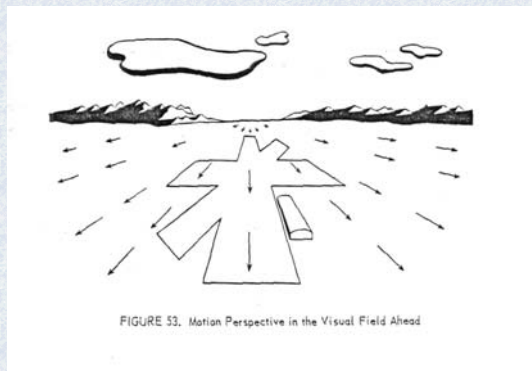


Che cos'è il flusso ottico?



L'intuizione di Gibson è stata che per ciascun punto degli oggetti nello spazio 3D si può ottenere questa misura => campo di movimento.

- **FLUSSO OTTICO: il movimento apparente del pattern di luminanza dell'immagine.**



- Movimenti omogenei sono generati dallo stesso oggetto => segmentazione, struttura e posizione 3D. Nessuna conoscenza 3D. Approccio bottom-up.



Come rappresentiamo il flusso ottico

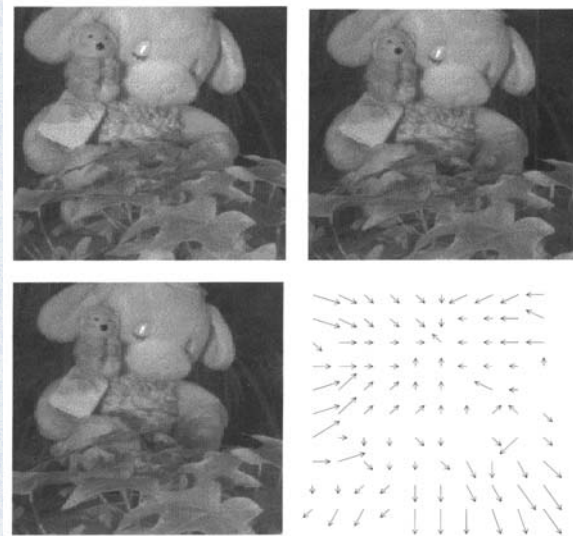


Figure 8.3 Three frames from a long image sequence (left to right and top to bottom) and the optical flow computed from the sequence, showing that the plant in the foreground is moving towards the camera, and the soft toys away from it.



Come calcoliamo il flusso ottico?



- Problema della corrispondenza in due istanti temporali successivi.
 - Campionamento temporale fitto \Rightarrow piccoli spostamenti.
 - Problema: stima del flusso ottico a partire dalle variazioni di luminosità
- $I = I(x(t), y(t), t)$. Luminosità è funzione dello spazio e del tempo. Spazio. I punti 3D possono spostarsi nel tempo, e, quindi, anche la loro proiezione.
Tempo. La luminosità del punto varia se varia la sorgente di illuminazione.
- Derivata totale di $I(\cdot)$
 - Deriviamo l'equazione della *costanza della luminosità*.

Ipotesi 1: Superfici Lambertiane

$L(\mathbf{P}, \mathbf{d}) = \rho \mathbf{I}^T \mathbf{n}$

La quantità di luce riflessa $L(\mathbf{P}, \mathbf{d}) = \rho \mathbf{I}^T \mathbf{n}$, dipende: dall'albedo in \mathbf{P} e dall'angolo tra la normale alla superficie in \mathbf{P} (\mathbf{n}) ed il raggio incidente \mathbf{I} .

testBy N.A. Borghese Università di Milano 19/03/2003 http://homes.dsi.unimi.it/~borghese 7/30

Ipotesi 2: Proiezione quasi-frontale

Luminosità della proiezione: $I(\mathbf{p}) = L(\mathbf{P}) * \pi/4 * (d/f)^2 * \cos^4(\alpha)$.

testBy N.A. Borghese Università di Milano 19/03/2003 http://homes.dsi.unimi.it/~borghese 8/30



Variazione locale di luminosità



• $I = I(\mathbf{p}(x,y),t)$. Luminosità è funzione della posizione sul piano immagine e del tempo $\implies dI/dt = 0$ - **Equazione della costanza della luminosità dell'immagine.**

• Calcoliamo la derivata totale rispetto al tempo:

$$\frac{dI}{dt} = \frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0$$

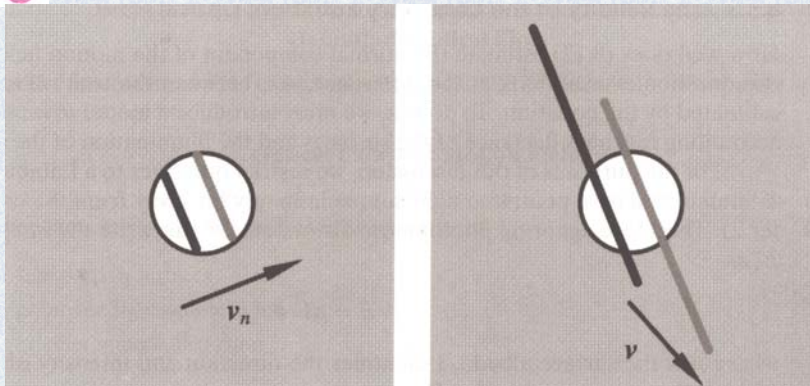
• Variazioni di luminosità si possono registrare se cambia il punto 3D che si proietta in un pixel. In particolare:

$$\nabla I \cdot \mathbf{v} = -\frac{\partial I}{\partial t}$$

Traduce variazioni temporali di luminosità in informazioni sul gradiente spaziale, ∇I , e sul campo di velocità sul piano immagine, \mathbf{v} (**flusso ottico**).



Problema dell'apertura



$$\frac{dI}{dt} = \frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0$$

$$\nabla I \cdot \mathbf{v} = -\frac{\partial I}{\partial t}$$

Ipotesi: conosciamo i bordi dell'immagine.

E' una equazione in 2 incognite: v_x e v_y . Solo la componente normale si può determinare.



Come determinare il flusso ottico?



$$\nabla I \cdot \mathbf{v} = -\frac{\partial I}{\partial t} \Rightarrow v_n = -\frac{\frac{\partial I}{\partial t}}{\|\nabla I\|}$$

La componente normale della velocità in un punto dipende dal gradiente spaziale (bordo) e dalla variazione di luminosità.

Soluzioni (Barron et al., 1994):

- Risolvere le equazioni differenziali per ogni punto dell'immagine.
- Calcolare le derivate seconde dell'equazione di costanza della luminosità e risolvere per ogni punto dell'immagine.
- Approssimazione lineare e stima locale ai minimi quadrati (soluzione semplice ed efficace).
- Estrazione di features significative e calcolo del loro spostamento.



Dal flusso ottico al campo di moto



Per superfici Lambertiane e visione frontale dell'oggetto, la quantità di luce incidente il sensore, risulta in prima approssimazione: $I(\mathbf{P}, d) = \rho \mathbf{I}^T \mathbf{n}$.

$$\nabla I \cdot \mathbf{v} = -\frac{\partial I}{\partial t} \Rightarrow v_n = -\frac{\frac{\partial I}{\partial t}}{\|\nabla I\|}$$

La componente normale della velocità in un punto dipende dal gradiente spaziale (bordo) e dalla variazione di luminosità.

Quale altra condizione possiamo derivare dall'equazione di costanza della luminosità?



Condizione sul moto della superficie



Analizziamo n . Vediamo come varia la luminosità in \mathbf{p} quando \mathbf{n} varia.

Derivata rispetto al tempo di: $I(\mathbf{p}) = \rho \mathbf{I} \times \mathbf{n} (\pi/4)(d/f)^2 \cos^4 \alpha$.

$$\frac{d\mathbf{I}(\mathbf{p})}{dt} = k \mathbf{I}^T \times \frac{d\mathbf{n}}{dt} = k \mathbf{I}^T \times (\mathbf{w} \wedge \mathbf{n})$$

Cosa ci dice? La quantità è nulla se: $\mathbf{w} // \mathbf{n}$.

$$\text{Altrimenti: } \nabla I \times \mathbf{v} + \frac{\partial I}{\partial t} = k \mathbf{I}^T \times (\mathbf{w} \wedge \mathbf{n})$$

$$\text{L'errore su } \mathbf{v}, \text{ supponendo } I(\mathbf{p}) = \text{cost. è: } \Delta \mathbf{v} = k \frac{\mathbf{I}^T \times (\mathbf{w} \wedge \mathbf{n})}{\|\nabla I\|}$$

che decresce con il “crescere” del bordo.



Validità flusso ottico = campo di moto



Ipotesi sul processo di formazione dell'immagine:

- Superfici Lambertiane
- Sorgenti luminose all'infinito (raggi paralleli a sé stessi).
- Visione quasi-frontale.

Ipotesi sul tipo di movimento della superficie:

- Movimento traslatorio oppure
- Movimento rotatorio con asse di rotazione orientato nella direzione della sorgente luminosa.

L'errore sul campo di moto è inversamente proporzionale al gradiente spaziale di luminosità.



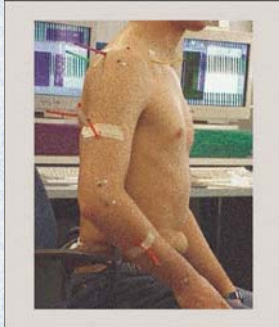
Algoritmo di Lucas e Kanade, 1991



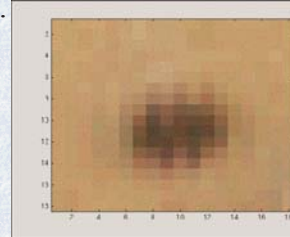
Ipotesi:

- L'equazione di costanza della luminosità, dà una buona stima di v_n .
- Il campo di moto può essere approssimato localmente (nello spazio e nel tempo) come costante.

Consideriamo una finestra $N \times N$ dell'immagine, W , in due istanti di tempo consecutivi. Tutti i calcoli saranno operati sui pixel di W .



Finestra $N \times N$ di un'immagine in 2 istanti successivi.



Scrittura di sistema lineare in v_p



- Fitting di un campo di velocità costante su regioni dell'immagine.
- Sfruttiamo l'elevata correlazione spazio-temporale di due immagini

$$\nabla I_x v_x + \nabla I_y v_y = -\frac{\partial I}{\partial t} \quad \forall p \in W$$

- v_x, v_y uguali in tutti i pixel di W .

$$a v_x + b v_y = c$$

- Abbiamo $N \times N$ equazioni in 2 incognite.



Soluzione sistema lineare in \mathbf{v}_p



$$\mathbf{A} = \begin{bmatrix} \nabla I_{p_1 x} & \nabla I_{p_1 y} \\ \nabla I_{p_2 x} & \nabla I_{p_2 y} \\ \vdots & \vdots \\ \nabla I_{p_{N^2} x} & \nabla I_{p_{N^2} y} \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} -\frac{\partial I}{\partial t} \Big|_{p_1} \\ -\frac{\partial I}{\partial t} \Big|_{p_2} \\ \vdots \\ -\frac{\partial I}{\partial t} \Big|_{p_{N^2}} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

$$\mathbf{Ax} = \mathbf{c} \quad \implies \quad \mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{c}$$

$\mathbf{x} = [v_x, v_y]$, \mathbf{v}_p calcolata al centro della finestra W .



Riassunto



• **Input:** una sequenza di immagini, $I(x,y,t)$ con $0 \leq x \leq X_{\max}$ e $0 \leq y \leq Y_{\max}$, $0 \leq t \leq T$. Dove con $[X_{\max}, Y_{\max}]$.

• **Filtriamo spazialmente** l'immagine ad esempio utilizzando un filtro Gaussiano ($\sigma_{\text{spaziale}} = 1.5$ pixels) per lasciare vivi solamente i contorni veri (la scala è legata alla scala dei bordi).

• **Filtriamo temporalmente le immagini** per eliminare piccole variazioni di illuminazione dovute a quantizzazione o variazioni di luminosità ($\sigma_{\text{temporale}} = 1.5$ frames; la scala è legata alla velocità del movimento).

• **Identifichiamo le regioni W** per ogni punto, \mathbf{p} . Per ogni regione:
Calcoliamo le matrici \mathbf{A} ed il vettore \mathbf{c} .
Risolviamo il sistema per determinare $\mathbf{v}(v_x, v_y)$.



Osservazioni



- Operazioni locali (poche operazioni sui pixel adiacenti).
- Nessuna iterazione.

Come mai se dall'equazione di costanza della luminosità derivo solamente la componente normale, con l'algoritmo di Lucas-Kanade, derivo la componente di velocità in direzione arbitraria?

Perché considero gradienti orientati in modo diverso all'interno della finestra W .

Come migliorare le prestazioni:

- Si può poi calcolare la somma vettoriale di v_p su più frame.
- Si può aumentare l'ampiezza della finestra (ma c'è un compromesso tra ampiezza della finestra di osservazione e validità dell'approssimazione velocità costante).



Come irrobustire l'algoritmo



Cosa succede se v_n non è uguale in tutti i p di W ?

Ci sarà un errore di stima: v_n non rappresenta il movimento di tutti i pixel in W . Come sarà distribuito questo errore?

Sarà massimo ai bordi della regione.

$$\nabla I_x v_x + \nabla I_y v_y = -\frac{\partial I}{\partial t} \quad \forall p \in W$$

Pesiamo ciascuna delle $N \times N$ equazioni con un peso s :

$$s = |\mathbf{p} - \mathbf{p}_c|$$



Stima pesata



$$A = \begin{bmatrix} \nabla I_{p_1 x} & \nabla I_{p_1 y} \\ \nabla I_{p_2 x} & \nabla I_{p_2 y} \\ \vdots & \vdots \\ \nabla I_{p_{N^2} x} & \nabla I_{p_{N^2} y} \end{bmatrix} \quad c = \begin{bmatrix} -\frac{\partial I}{\partial t} \Big|_{p_1} \\ -\frac{\partial I}{\partial t} \Big|_{p_2} \\ \vdots \\ -\frac{\partial I}{\partial t} \Big|_{p_{N^2}} \end{bmatrix} \quad x = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

$$Ax = c \implies ASx = c \quad S = \text{diag}[|p_{11}-p_c|; |p_{12}-p_c|; \dots; |p_{NN}-p_c|]$$

$$x = (A^T S^2 A)^{-1} A^T S c$$



Osservazione



Quando il sistema analizzato non è risolubile?

$$A = \begin{bmatrix} \nabla I_{p_1 x} & \nabla I_{p_1 y} \\ \nabla I_{p_2 x} & \nabla I_{p_2 y} \\ \vdots & \vdots \\ \nabla I_{p_{N^2} x} & \nabla I_{p_{N^2} y} \end{bmatrix}$$

$$A^T A = \begin{bmatrix} \sum_k \nabla I_{p_k x} \nabla I_{p_k x} & \sum_k \nabla I_{p_k x} \nabla I_{p_1 y} \\ \sum_k \nabla I_{p_k x} \nabla I_{p_1 y} & \sum_k \nabla I_{p_k y} \nabla I_{p_k y} \end{bmatrix}_{2 \times 2}$$



Condizione di risolubilità



$$A^T A = \begin{bmatrix} \sum_k (\nabla I_{p_k x})^2 & \sum_k \nabla I_{p_k x} \nabla I_{p_1 y} \\ \sum_k \nabla I_{p_k x} \nabla I_{p_1 y} & \sum_k (\nabla I_{p_k y})^2 \end{bmatrix}$$

Dipende solamente da informazioni spaziali (gradiente).

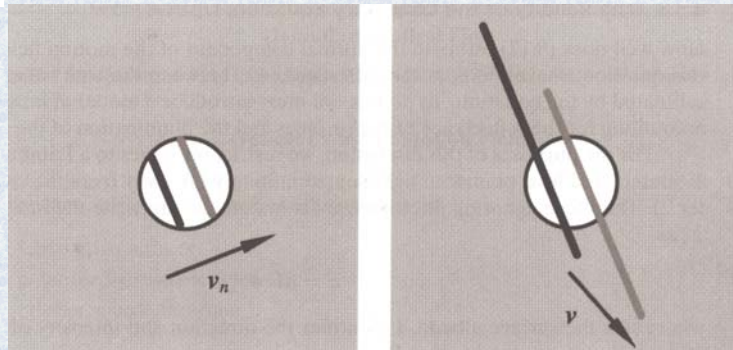
La matrice $A^T A$ è singolare, quando:

- I gradienti sono nulli (non ci sono bordi: rango 0).
- I gradienti sono orientati in una sola direzione (bordi paralleli): rango 1.

Nel secondo caso cosa riesco a vedere? Solamente la direzione normale della velocità, è il *problema dell'apertura*. Occorre quindi che i bordi non siano paralleli tra loro.



Problema dell'apertura



$$\nabla I \cdot \mathbf{v} + \frac{\partial I}{\partial t} = 0$$

$$\nabla I \cdot \mathbf{v} + \frac{\partial I}{\partial t} = 0$$

$$\nabla I_x v_x + \nabla I_y v_y = -\frac{\partial I}{\partial t}$$

- Se il bordo è orientato come l'asse x, $\nabla I_x = 0$
 v_x non può essere osservato.
- Se il bordo è orientato come l'asse y, $\nabla I_y = 0$
 v_y non può essere osservato.



Flusso ottico mediante features



Le stesse features vengono identificate in una sequenza di immagini.

$$A^T A = \begin{bmatrix} \sum_k (\nabla I_{p_k x})^2 & \sum_k \nabla I_{p_k x} \nabla I_{p_k y} \\ \sum_k \nabla I_{p_k x} \nabla I_{p_k y} & \sum_k (\nabla I_{p_k y})^2 \end{bmatrix}$$

Come possiamo identificare una feature, analizzando i gradienti spaziali?

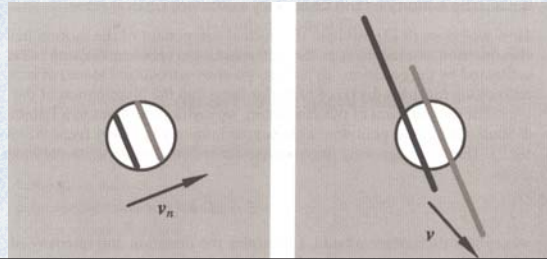
Possiamo ad esempio trovare gli angoli.

Utilizzando la svd:

$$A^T A = U W V$$

$$W = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \lambda_1 \geq \lambda_2 \geq 0.$$

testBy N.A. Borghese Università di Milano 19/03/2003



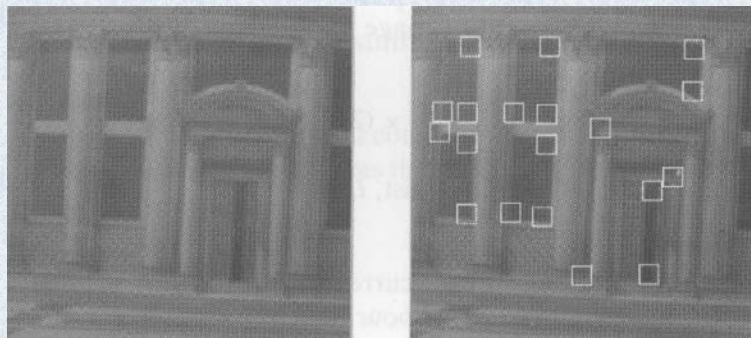
Analisi degli autovalori



$\lambda_1 \geq \lambda_2 \geq 0$. λ_2 è il punto critico.

- Gli autovettori codificano le *direzioni* prevalenti dei contorni.
- Gli autovalori codificano la *forza* dei contorni lungo le due direzioni principali.

•Su 4 pixel:



testBy N.A. Borghese Università di Milano 19/03/2003

<http://homes.dsi.unimi.it/~borghese> 26/30



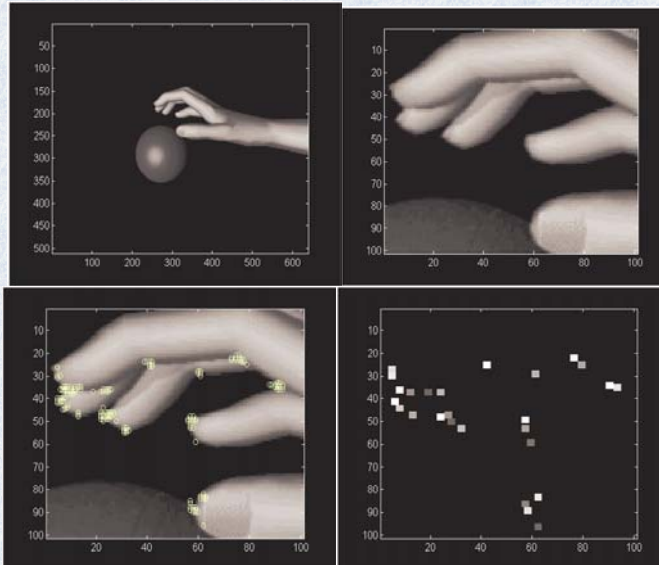
Features dal gradiente spaziale



- Partiamo da: (a) un'immagine $I(x,y)$; (b) una soglia sull'intensità minima del gradiente, $\lambda_{2 \min}$; (c) la dimensione della finestra di analisi, W .
- Per ogni finestra W , calcoliamo i gradienti spaziali.
- Per ciascuna finestra:
 - Calcoliamo λ_2 .
 - Se $\lambda_2 > \lambda_{2 \min}$, salviamo il centro della finestra nella lista delle features candidate.
- Ordiniamo la lista delle feature per valori di λ_2 decrescenti.
- Esaminiamo la lista. Per ogni feature, vengono eliminate tutte quelle altre feature che cadono all'interno di W .



Tracking di feature



Problemi su
superfici non
strutturate

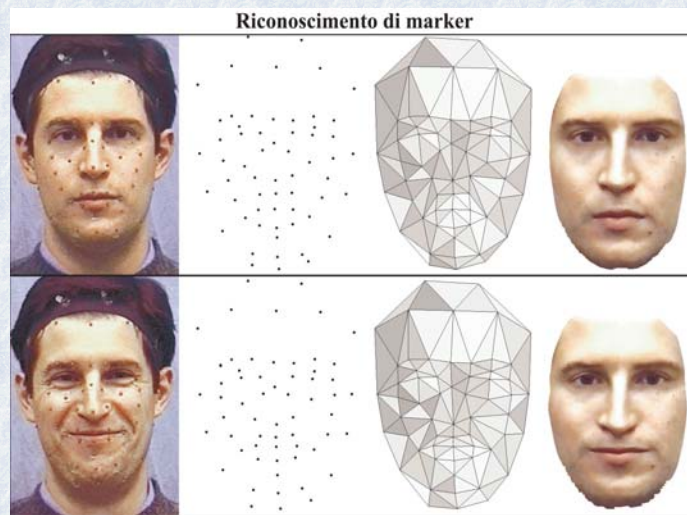


E per le superfici non strutturate?



Markers.

Shape from Shading.



Riassunto



- Determinazione del campo di moto 2D mediante tecniche differenziali o matching di feature.
- Flusso ottico: spostamento del pattern di luminosità di un'immagine indotto dal movimento rispetto alla scena.
- Sotto alcune ipotesi, il flusso ottico si può ricavare da un'analisi delle variazioni temporali di luminosità.
- Esistono algoritmi semplici (locali e non iterativi) per determinare una buona approssimazione del flusso ottico.
- Gli stessi algoritmi possono essere utilizzati per determinare feature (spigoli) nelle immagini.