CAP 5415 Computer Vision
Fall 2005

Dr. Alper Yilmaz
Univ. of Central Florida
www.cs.ucf.edu/courses/cap5415/fall2005
Office: CSB 250

Recap
3D Motion

- Displacement model

$$X' = X - \alpha Y + \beta Z + T_X$$
$$Y' = \alpha X + Y - \gamma Z + T_Y$$
$$Z' = -\beta X + \gamma Y + Z + T_Z$$

- Velocity model

$$\dot{X} = \Omega_2 Z - \Omega_3 Y + V_1$$
$$\dot{Y} = \Omega_3 X - \Omega_1 Z + V_2$$
$$\dot{Z} = \Omega_1 Y - \Omega_2 X + V_3$$

- 2D optical flow

orthographic projection

$$u = \Omega_2 Z - \Omega_3 y + V_1$$
$$v = \Omega_3 x - \Omega_1 Z + V_2$$

perspective projection

$$u = f(\frac{V_1}{Z} + \Omega_2) - \frac{V_3}{Z}x - \Omega_3 y - \frac{\Omega_1}{f}xy + \frac{\Omega_2}{f}x^2$$
$$v = f(\frac{V_2}{Z} - \Omega_1) + \Omega_3 x - \frac{V_3}{Z}y + \frac{\Omega_2}{f}xy - \frac{\Omega_1}{f}y^2$$

$$x' = x - \alpha y + \beta Z + T_X$$
$$y' = \alpha x + y - \gamma Z + T_Y$$

Recap
SFM (Orthographic Displacement)

- Two step simple approach
  - Assume depth Z is known compute motion parameters α, β, γ, T_x, T_y
  - Using new motion parameters refine depth Z

$$\begin{bmatrix} x' - x \\ y' - y \end{bmatrix} = \begin{bmatrix} -y & Z & 0 & 1 & 0 \\ x & 0 & -Z & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ T_x \\ T_y \end{bmatrix}$$

Step 1

$$\begin{bmatrix} \beta \\ -\gamma \end{bmatrix} [Z] = \begin{bmatrix} x' - x - \alpha y - T_x \\ y' - y - \alpha x - T_y \end{bmatrix}$$

Step 2

$$u = -(\frac{V_1}{Z} + \Omega_2) + \frac{V_3}{Z} x + \Omega_3 y + \Omega_1 xy - \Omega_2 x^2$$
$$v = -(\frac{V_2}{Z} - \Omega_1) - \Omega_3 x + \frac{V_3}{Z} y - \Omega_2 xy + \Omega_1 y^2$$

Recap
SFM (Perspective Motion)

- Find translation using search.
- Find rotation using least squares fit.
- Find Depth.

$$\Theta = \mathbf{C}(\mathbf{V})\mathbf{q} \qquad E(\mathbf{V},\mathbf{q}) = \|\Theta - \mathbf{C}(\mathbf{V})\mathbf{q}\|^2$$

$\mathbf{q}$ → 1/Z and rotations

Θ → optical flow

$\mathbf{C}(\mathbf{V})$ → matrix composed of translation and observations

$$E(\mathbf{V},\mathbf{q}) = \|\Theta - \mathbf{C}(\mathbf{V})\mathbf{q}\|^2$$

Recap
SFM (Perspective Motion)

- Translation search
  - Find derivative wrt. **q** and replace in *E*
  - Search for possible translations minimizing *E*
- Find Rotation
  - Define a vector perpendicular to ***p*A**
  - Least squares solution to Ω

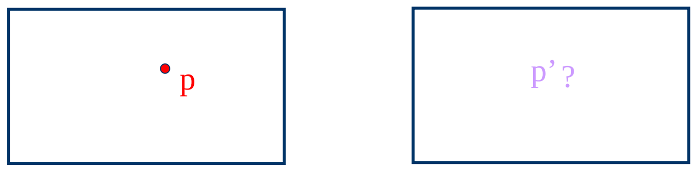
$$\overbrace{d^T(x,y,V)\Theta(\mathbf{x},\mathbf{y})}^{D} = \overbrace{d^T(x,y,V)\mathbf{B}(x,y)}^{E}\Omega$$

- Find depth by putting all values in

$$u = f(\frac{V_1}{Z} + \Omega_2) - \frac{V_3}{Z}x - \Omega_3 y - \frac{\Omega_1}{f}xy + \frac{\Omega_2}{f}x^2 \qquad v = f(\frac{V_2}{Z} - \Omega_1) + \Omega_3 x - \frac{V_3}{Z}y + \frac{\Omega_2}{f}xy - \frac{\Omega_1}{f}y^2$$

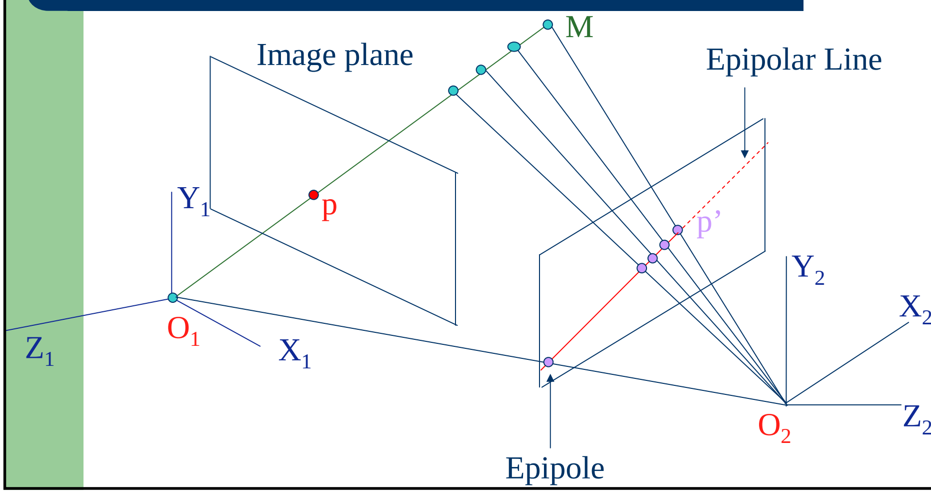
Multi-View Geometry

Stereo Constraints



Given p in the left image, where can the corresponding point p' be in the right image?

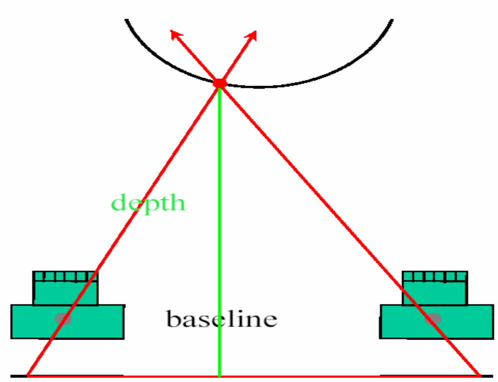
Stereo Constraints
Epipolar Geometry



Stereopsis

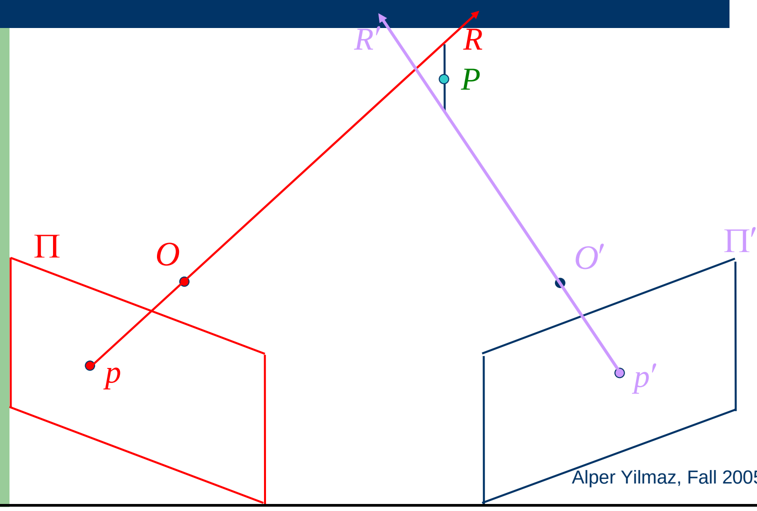
Alper Yilmaz, Fall 2005 UCF

Definition



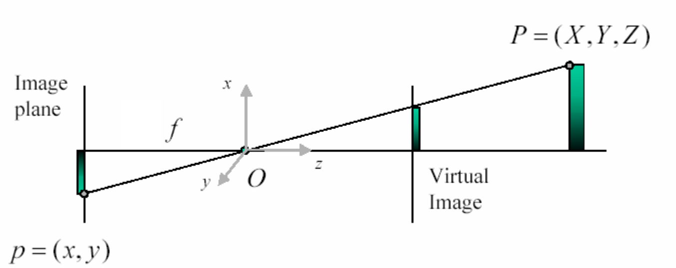
Alper Yilmaz, Fall 2005 UCF

Definition



Alper Yilmaz, Fall 2005 UCF

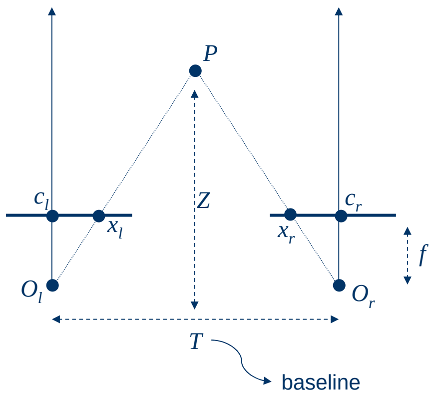
Pin Hole Camera Model



$$x = f\frac{X}{Z}$$

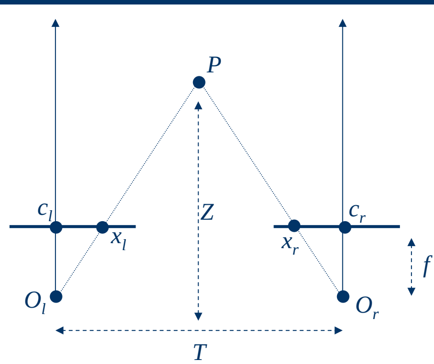
Alper Yilmaz, Fall 2005 UCF

Basic Stereo Derivations



Derive expression for Z as a function of x_l, x_r, f and T

Basic Stereo Derivations



$$x_l = f\frac{X}{Z} \qquad x_r = f\frac{X}{Z}$$

$$\frac{T + x_l - x_r}{Z - f} = \frac{T}{Z}$$

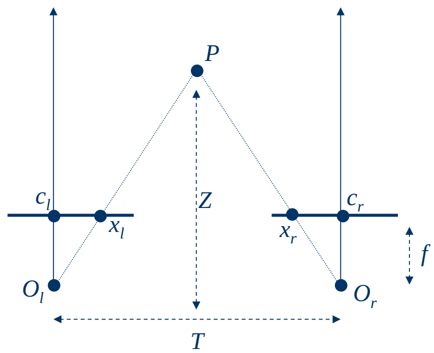
$$Z(T + x_l - x_r) = T(Z - f)$$

$$ZT + Zx_l - Zx_r = TZ - Tf$$

$$\boxed{Z = f\frac{T}{x_r - x_l}}$$

Alper Yilmaz, Fall 2005 UCF

Basic Stereo Derivations



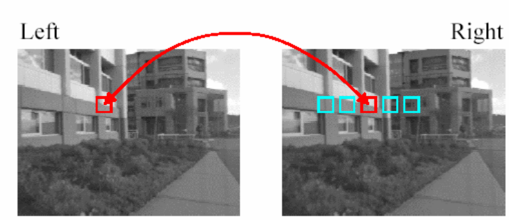
$$Z = f\frac{T}{x_r - x_l}$$

$$d = x_1 - x_2 \quad \text{disparity}$$

$$Z = f\frac{T}{d}$$

Token Based Stereo

- Detect token
  - Corners, interest point, edges
- Find correspondences
- Interpolate surface



Finding Correspondences

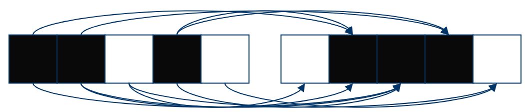
- Due to epilolar constraint stereo is a 1D search problem
  - Marr-Poggio
  - Correlation based stereo methods
  - Barnard's algorithm

Marr-Poggio Algorithm

- One of first stereo methods
- Applicable to image composed of pixels with randomly selected black or white colors
- Constraints:
  - **Compatibility:** white should match with a white
  - **Continuity:** neighbors should have similar disparity
  - **Uniqueness:** a pixel in left image matches with only one pixel on right image

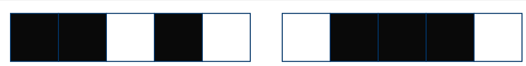
Marr-Poggio Algorithm

- Correspondence map
- Search is along the rows
- 1D case



	1	2	3	4	5
d=2	1	1	1		
d=1	1	1			
d=0		1		1	1
d=-1				1	
d=-2			1	1	

Initial correspondence array



Marr-Poggio Algorithm

- Iteratively update correspondence array using

1D Case

$$C^n(x,d)=\begin{cases}1 & \left\{\sum_{a=-w,a\neq 0}^{w} C^{n-1}(x+a,d) - \lambda \sum_{i=-D,i\neq 0}^{D} C^{n-1}(x\pm i, d+i) - C^0(x,d)\right\} > T \\ 0 & otherwise\end{cases}$$

excitatory term

max. allowed disparity

inhibitory term for continuity

2D Case

$$C^n(x,d)=\begin{cases}1 & \left\{\sum\sum_{|x-p|+|y-q|\leq w} C^{n-1}(x+p,y+q,d) - \lambda \sum_{i=-D,i\neq 0}^{D} C^{n-1}(x\pm i, y, d+i) - C^0(x,y,d)\right\} > T \\ 0 & otherwise\end{cases}$$

Correlation Based Stereo Methods

- Depth is computed only at tokens and interpolated/extrapolated to remaining pixel
- Disparity map is constructed based on a correlation measure

SSD $= \sum\sum(I_{t+1} - I_t)^2$

AD $= \sum\sum|(I_{t+1} - I_t)|$

CC $= \sum\sum I_{t+1} I_t$

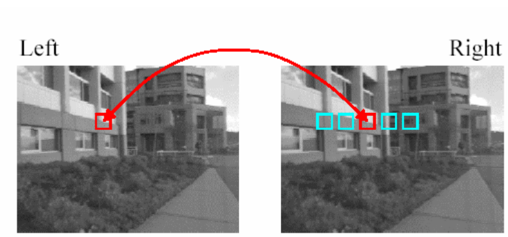
NC $= \frac{\sum\sum(I_{t+1}.I_t)}{\sqrt{\sum\sum I_t.I_t}}$

MC $= \frac{1}{64\sigma_{t+1}\sigma_t}\sum\sum(I_{t+1} - \mu_{t+1})(I_t - \mu_t)$

Alper Yilmaz, Fall 2005 UCF

Correlation Based Stereo Methods

- Once disparity is available compute depth using $Z = f\frac{T}{d}$

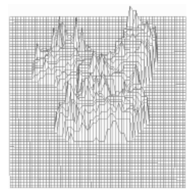


Alper Yilmaz, Fall 2005 UCF

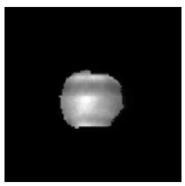
Examples

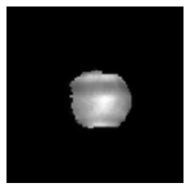
Left Image

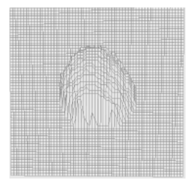
Right Image



Depth Map







Alper Yilmaz, Fall 2005 UCF

Barnard's Stereo Method

- Similar intensity
  - Similar to brightness constraint
- Smoothness of disparity

$$\mathrm{E} = \sum_{i=-1}^{1}\sum_{j=-1}^{1}\left\| I_{left}(x+i, y+j) - I_{right}(x+i+D(x,y), y+j)\right\| + \lambda \left\| \nabla D(x,y) \right\|$$

$$\nabla D(x,y) = \sum_{i=-1}^{1}\sum_{j=-1}^{1}\left| D(x+i, y+j) - D(x,y) \right|$$

Alper Yilmaz, Fall 2005 UCF

Barnard's Stereo Method

- Energy can be minimized using brute force search
  - Let max allowed disparity is 10 pixels
  - For every point we have to search a *circle* of radius 10
    - Aprox. 80 possible associations per pixel
    - For 128x128 image we will have $2^{(7+7+6)}=2^{20}$ associations
- We can select any minimization technique
  - Barnard choose simulated annealing

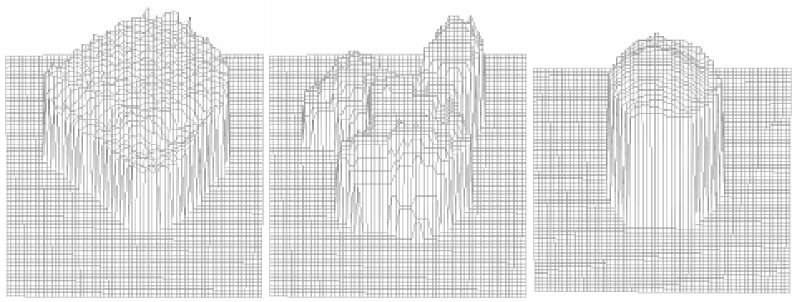
Alper Yilmaz, Fall 2005 UCF

Simulated Annealing

- Select a random state S (disparities)
- Select a high temperature
  - Select random S'
  - Compute $\Delta E=E(S')-E(S)$
  - If $(\Delta E<0)$ $S \leftarrow S'$
  - Else
    - $P \leftarrow \exp(-\Delta E/T)$
    - $X \leftarrow$ random(0,1)
      - If $X<P$ then $S \leftarrow S'$
    - If no decrease in several iterations lower T

Alper Yilmaz, Fall 2005 UCF

Examples



bread toy apple

Reading

- Trucco & Verri
  - Section 7.1 (pages 140-150)
- Lecture notes
  - Chapter 6

Alper Yilmaz, Fall 2005 UCF