Stereo

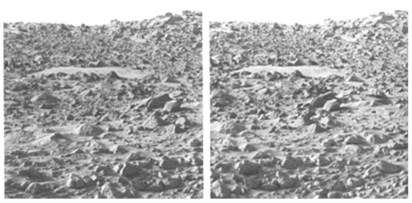
Introduction to Computer Vision
CSE 152
Lecture 8

Announcements

- Homework 2 is due today, 11:59 PM
- Homework 3 will be assigned today
- Reading:
  - Chapter 7: Stereopsis

Binocular Stereopsis: Mars

Given two images of a scene where relative locations of cameras are known, estimate depth of all common scene points.



Two images of Mars (Viking Lander)

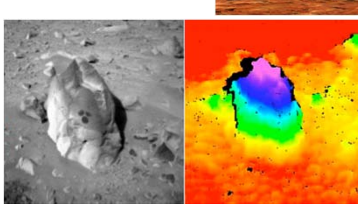
An Application: Mobile Robot Navigation

The Stanford Cart,
H. Moravec, 1979.

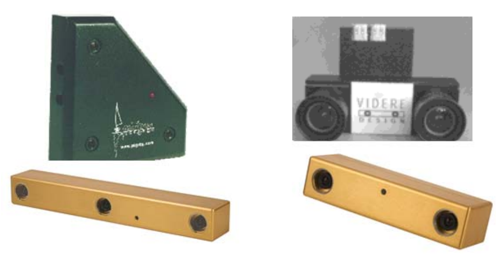
The INRIA Mobile Robot, 1990.

Courtesy O. Faugeras and H. Moravec.

Mars Exploratory Rovers:
Spirit and Opportunity



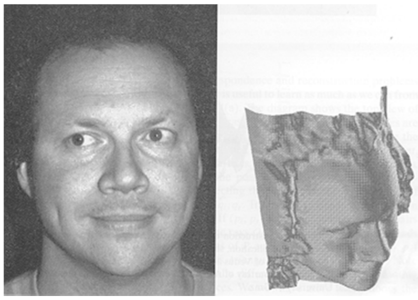
Commercial Stereo Heads



Trinocular stereo

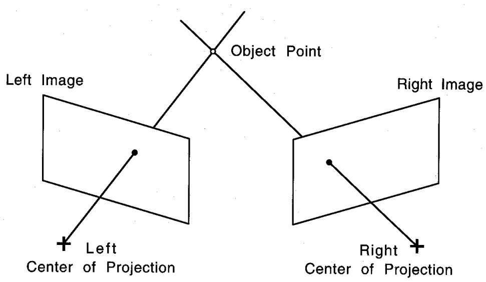
Binocular stereo

Stereo can work well



Need for correspondence

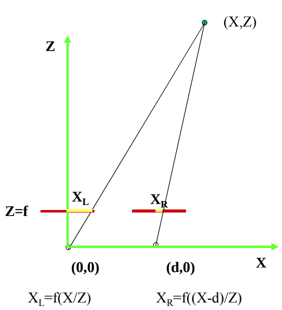
Triangulation



Stereo Vision Outline

- Offline: Calibrate cameras & determine
B "epipolar geometry"
- Online
  1. Acquire stereo images
  C 2. Rectify images to convenient epipolar geometry
  D 3. Establish correspondence
  A 4. Estimate depth

BINOCULAR STEREO SYSTEM
Estimating Depth\
2D world with 1-D image plane



Two measurements: X_L, X_R
Two unknowns: X,Z

Constants:
Baseline: d
Focal length: f

$$X = \frac{d\,X_L}{(X_L - X_R)}$$

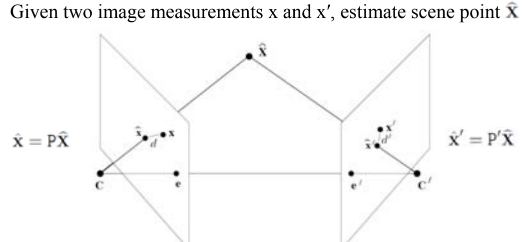
$$Z = \frac{d\,f}{(X_L - X_R)}$$

Disparity: $(X_L - X_R)$

(Adapted from Hager)

Reconstruction: General 3D case

Given two image measurements x and x′, estimate scene point $\hat{X}$



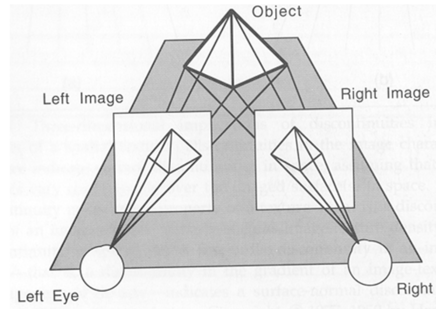
- Linear Method: $x \times (P\hat{X}) = 0$, $x' \times (P'\hat{X}) = 0$ Solve system of linear equations
- Non-Linear Method: Estimate $\hat{X}$ that minimizes $d(x,\hat{x})^2 + d(x',\hat{x}')^2$

Two Approaches

1. Feature-Based
   – From each image, process “monocular” image to obtain cues (e.g., corners, lines).
   – Establish correspondence between the two images.

2. Area-Based
   – Directly compare image regions between the two images.

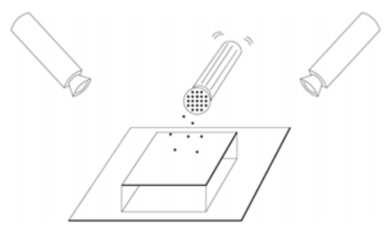
Human Stereopsis



Human Stereopsis: Binocular Fusion

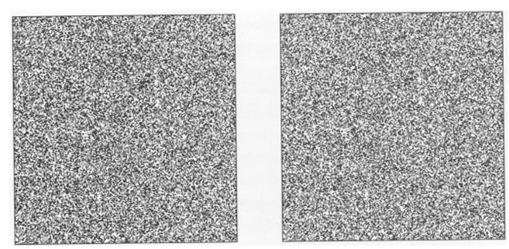
How are the correspondences established?

Julesz (1971): Is the mechanism for binocular fusion a monocular process or a binocular one??
• There is anecdotal evidence for the latter (camouflage).

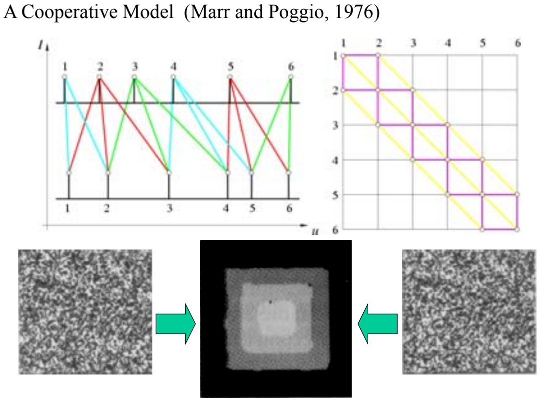


• Random dot stereograms provide an objective answer

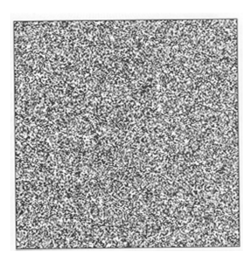
Random Dot Stereograms



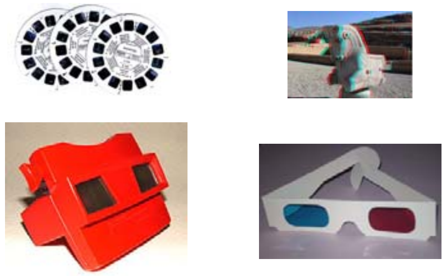
A Cooperative Model (Marr and Poggio, 1976)



Random Dot Stereograms



Stereoscopic 3D



CSE 152, Spring 2017 Introduction to Computer Vision

Stereoscopic 3D

CSE 152, Spring 2017 Introduction to Computer Vision

Was Rembrandt Stereo Blind?

- Detail of a 1639 etching.

CSE 152, Spring 2017 Introduction to Computer Vision

- In Rembrandt's painted self-portraits (left panel) in which the eyes are clearly visible, his left eye frequently looks straight out and the right off to the side. It is the opposite in his etchings (right panel).

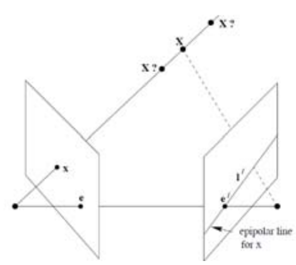
CSE 152, Spring 2017 Introduction to Computer Vision

Need for correspondence

CSE 152, Spring 2017 Introduction to Computer Vision

Point correspondence geometry

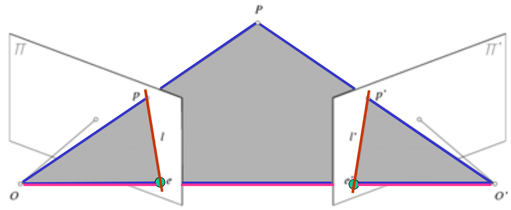
Where does a point **x** in the left image match a point in the right image?



- Potential matches for **x** have to lie on the corresponding epipolar line **l′**

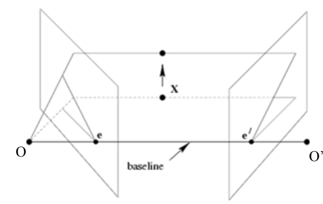
CSE 152, Spring 2017 Introduction to Computer Vision

Epipolar geometry



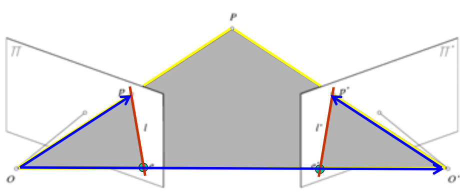
- Baseline
- Epipoles
- Epipolar Plane
- Epipolar Lines

Family of epipolar Planes



Family of planes π and lines l and l'
Intersection in e and e'

Epipolar Constraint: Calibrated Case



The vectors Op, OO', and O'p' are coplanar

$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0 \quad \Rightarrow \quad \boldsymbol{p} \cdot [\boldsymbol{t} \times (\mathcal{R}\boldsymbol{p}')] = 0 \quad \text{with} \begin{cases} \boldsymbol{p} = (u, v, 1)^T \\ \boldsymbol{p}' = (u', v', 1)^T \\ \mathcal{M} = (\text{Id} \;\; \boldsymbol{0}) \\ \mathcal{M}' = (\mathcal{R}^T, -\mathcal{R}^T\boldsymbol{t}) \end{cases}$$

$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0 \quad \text{with} \quad \mathcal{E} = [\boldsymbol{t}_\times]\mathcal{R}$$

Essential Matrix
(Longuet-Higgins, 1981)

Skew Symmetric Matrix & Cross Product

- The cross product ***a x b*** of two vectors **a** and **b** can be expressed a matrix vector product $[\mathbf{a}_x]\mathbf{b}$ where $[\mathbf{a}_x]$ is the skew symmetric matrix:

$$[\mathbf{a}_\times] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

- A matrix ***S*** is skew symmetric if and only if

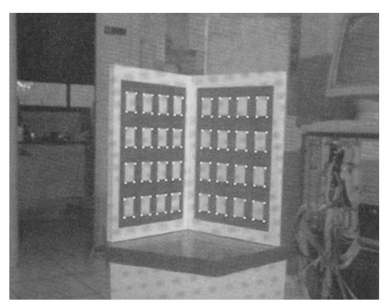
$$\boldsymbol{S} = -\boldsymbol{S}^T$$

Properties of the Essential Matrix

$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0 \quad \text{with} \quad \mathcal{E} = [\boldsymbol{t}_\times]\mathcal{R}$$

- $\mathcal{E}$ p' is the epipolar line associated with p'.
- $\mathcal{E}^T$p is the epipolar line associated with p.
- $\mathcal{E}$ e'=0 and $\mathcal{E}^T$e=0.
- $\mathcal{E}$ is singular (rank 2).
- $\mathcal{E}$ has two equal non-zero singular values (Huang and Faugeras, 1989).

Calibration



Determine intrinsic parameters and extrinsic relation of two cameras

The Eight-Point Algorithm (Longuet-Higgins, 1981)

$$(u, v, 1)\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \;\Rightarrow\; (uu', uv', u, vu', vv', v, u', v', 1)\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0$$

- Consider 8 points (u_i,v_i), (u_i',v_i')
- Set F_{33} to 1

$$\begin{pmatrix} u_1u_1' & u_1v_1' & u_1 & v_1u_1' & v_1v_1' & v_1 & u_1' & v_1' \\ u_2u_2' & u_2v_2' & u_2 & v_2u_2' & v_2v_2' & v_2 & u_2' & v_2' \\ u_3u_3' & u_3v_3' & u_3 & v_3u_3' & v_3v_3' & v_3 & u_3' & v_3' \\ u_4u_4' & u_4v_4' & u_4 & v_4u_4' & v_4v_4' & v_4 & u_4' & v_4' \\ u_5u_5' & u_5v_5' & u_5 & v_5u_5' & v_5v_5' & v_5 & u_5' & v_5' \\ u_6u_6' & u_6v_6' & u_6 & v_6u_6' & v_6v_6' & v_6 & u_6' & v_6' \\ u_7u_7' & u_7v_7' & u_7 & v_7u_7' & v_7v_7' & v_7 & u_7' & v_7' \\ u_8u_8' & u_8v_8' & u_8 & v_8u_8' & v_8v_8' & v_8 & u_8' & v_8' \end{pmatrix}\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix} = -\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Solve for F_{11} to F_{32}

For more than 8 points, solve using linear least squares

The Eight-Point Algorithm (Longuet-Higgins, 1981)

$$(u, v, 1)\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \;\Rightarrow\; (uu', uv', u, vu', vv', v, u', v', 1)\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0$$

- Alternatively, view this as system of homogenous equations in F_{11} to F_{33}
- Solve as Eigenvector corresponding to the smallest Eigenvalue of matrix created from the image data.

Equivalent to solving

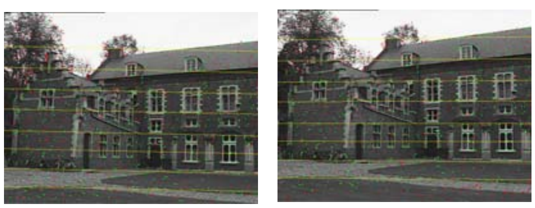
Minimize:

$$\sum_{i=1}^{n} (\boldsymbol{p}_i^T \mathcal{F} \boldsymbol{p}_i')^2$$

under the constraint

$|\mathcal{F}|^2 = 1.$

Epipolar geometry example



The Fundamental matrix

The epipolar constraint is given by:

where p and p′ are called homogeneous normalized image coordinates of points in the two images.

Without calibration, we can still identify corresponding points in two images, but we cannot convert to 3D coordinates. However, the relationship between the calibrated coordinates p and p′, and uncalibrated coordinates x and x′ can be expressed as $p = K^{-1} x$ and $p' = K'^{-1} x'$

Therefore, we can express the epipolar constraint as:

$$p'^T E p = (K'^{-1} x')^T E (K^{-1} x) = x'^T K'^{-T} E K^{-1} x = x'^T F x = 0$$

where $F = K'^{-T} E K^{-1}$ is called the Fundamental Matrix.

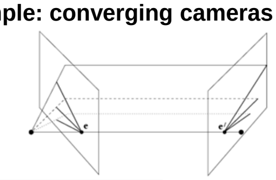
Can be solved using 8 point algorithm WITHOUT CALIBRATION

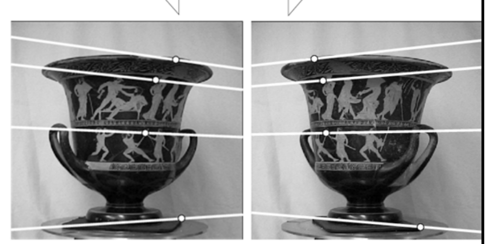
Two-View Geometry

Essential Matrix E
- Calibrated
- Normalized coordinates
- Rank 2
  - Two nonzero singular values are equal
- 5 degrees of freedom
  - Camera rotation
  - Direction of camera translation
- Similarity reconstruction

Fundamental Matrix F
- Uncalibrated
- Image coordinates
- Rank 2
- 7 degrees of freedom
  - Homogeneous matrix to scale
  - det F = 0
- Projective reconstruction

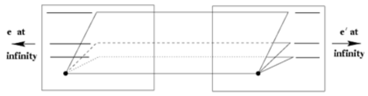
Example: converging cameras





courtesy of Andrew Zisserman

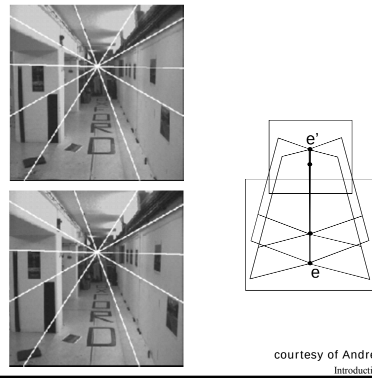
Example: motion parallel with image plane



(simple for stereo → rectification)

courtesy of Andrew Zisserman

Example: forward motion



courtesy of Andrew Zisserman

Next Lecture

- Early vision: multiple images
  - Stereo
- Reading:
  - Chapter 7: Stereopsis