

Stereo Vision II

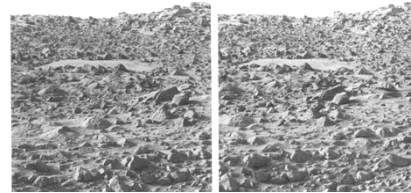
Introduction to Computer Vision
CSE 152
Lecture 14

CSE152, Spring 2011

Intro Computer Vision

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

CSE152, Spring 2011

Intro Computer Vision

Stereo Vision Outline

- Offline: Calibrate cameras & determine “epipolar geometry”
- Online
 1. Acquire stereo images
 2. Rectify images to convenient epipolar geometry
 3. Establish correspondence
 4. Estimate depth

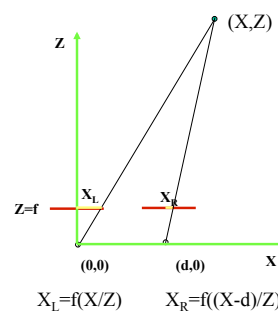
CSE152, Spring 2011

Intro Computer Vision

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: $(X_L - X_R)$

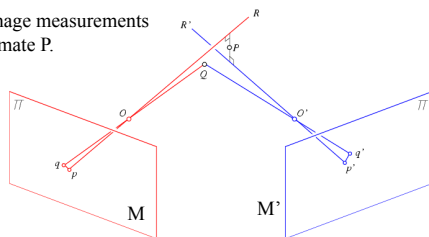
(Adapted from Hager)

CSE152, Spring 2011

Intro Computer Vision

Reconstruction: General 3-D case

Given two image measurements p and p' , estimate P .



• Linear Method: find P such that

$$\begin{cases} p \times M P = 0 \\ p' \times M' P = 0 \end{cases} \iff \begin{bmatrix} p \times M \\ p' \times M' \end{bmatrix} P = 0$$

Where M is camera matrix

• Non-Linear Method: find Q minimizing $d^2(p, q) + d^2(p', q')$
where $q = MQ$ and $q' = M'Q$

CSE152, Spring 2011

Intro Computer Vision

The search for correspondence: Where do you look?



CSE152, Spring 2011

Intro Computer Vision

Two Approaches

1. Feature-Based

- From each image, process “monocular” image to obtain image features or cues(e.g., corners, lines).
- Establish correspondence between the detected features.

2. Area-Based

- Directly compare image regions between the two images.

CSE152, Spring 2011

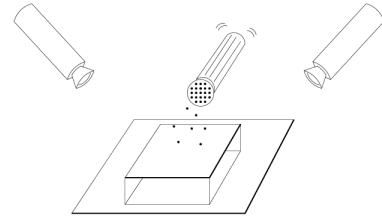
Intro Computer Vision

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).

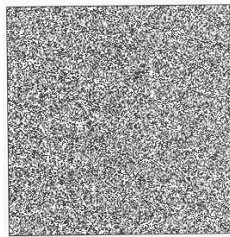


- Random dot stereograms provide an objective answer

CSE152, Spring 2011

Intro Computer Vision

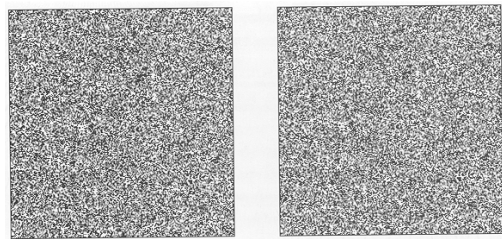
Random Dot Stereograms



CSE152, Spring 2011

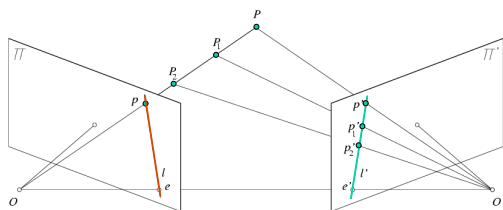
Intro Computer Vision

Random Dot Stereograms



CSE152, Spring 2011

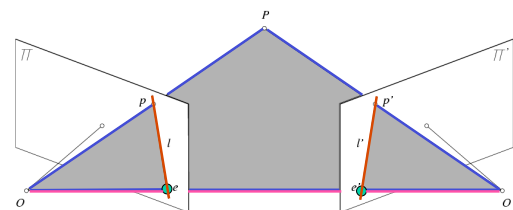
Intro Computer Vision



- Potential matches for p have to lie on the corresponding epipolar line l' .
- Potential matches for p' have to lie on the corresponding epipolar line l .

CSE152, Spring 2011

Intro Computer Vision

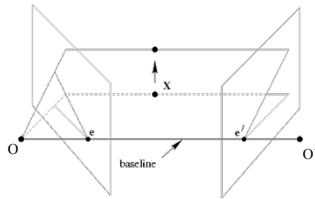


- Baseline
- Epipolar Plane
- Epipoles
- Epipolar Lines

CSE152, Spring 2011

Intro Computer Vision

Family of epipolar Planes



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

CSE152, Spring 2011

Intro Computer Vision

Skew Symmetric Matrix & Cross Product

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

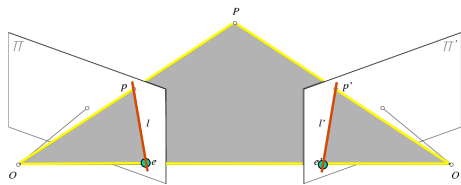
$$[\mathbf{a}_\times] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

- A matrix $\mathbf{S}$ is skew symmetric iff $\mathbf{S} = -\mathbf{S}^T$

CSE152, Spring 2011

Intro Computer Vision

Epipolar Constraint: Calibrated Case



$$\vec{Op} \cdot [\vec{Op} \times \vec{Op'}] = 0 \Rightarrow \mathbf{p} \cdot [\mathbf{t} \times (\mathbf{R}\mathbf{p}')] = 0 \text{ with } \begin{cases} \mathbf{p} = (u, v, 1)^T \\ \mathbf{p}' = (u', v', 1)^T \\ \mathbf{M} = (\mathbf{Id} \ \mathbf{0}) \\ \mathbf{M}' = (\mathbf{R}^T, -\mathbf{R}^T \mathbf{t}) \end{cases}$$

Note: Coordinates of $\mathbf{p}$ and $\mathbf{p}'$ are "calibrated" so that $(u, v, 1)$ is a 3-D vector pointing in the direction of $\mathbf{p}$.

Essential Matrix
(Longuet-Higgins, 1981)

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

CSE152, Spring 2011

Intro Computer Vision

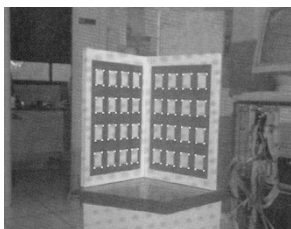
Two ways to estimate the Essential Matrix

1. Calibration-based
2. Eight-Point Algorithm

CSE152, Spring 2011

Intro Computer Vision

Calibration-Based Method



1. From image of known calibration fixture, determine intrinsic parameters and extrinsic relation of two cameras.
2. Compute the relative position and orientation of the two cameras from $\mathbf{R}_1, \mathbf{R}_2, \mathbf{t}_1, \mathbf{t}_2$
3. Compute the Essential Matrix $\mathbf{E} = [\mathbf{t}_\times] \mathbf{R}$

CSE152, Spring 2011

Intro Computer Vision

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

$$[u, v, 1] \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{bmatrix} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = 0 \Rightarrow [uu' \ uv' \ u \ uu' \ vv' \ v \ u' \ v' \ 1] \begin{bmatrix} E_{11} \\ E_{12} \\ E_{13} \\ E_{21} \\ E_{22} \\ E_{23} \\ E_{31} \\ E_{32} \\ E_{33} \end{bmatrix} = 0$$

- Set E_{33} to 1
- Consider 8 points $(u_i, v_i), (u'_i, v'_i), i=1..8$

$$\begin{bmatrix} u_1 u'_1 & u_1 v'_1 & u_1 & v_1 u'_1 & v_1 v'_1 & v_1 & u'_1 & v'_1 & 1 \\ u_2 u'_2 & u_2 v'_2 & u_2 & v_2 u'_2 & v_2 v'_2 & v_2 & u'_2 & v'_2 & 1 \\ u_3 u'_3 & u_3 v'_3 & u_3 & v_3 u'_3 & v_3 v'_3 & v_3 & u'_3 & v'_3 & 1 \\ u_4 u'_4 & u_4 v'_4 & u_4 & v_4 u'_4 & v_4 v'_4 & v_4 & u'_4 & v'_4 & 1 \\ u_5 u'_5 & u_5 v'_5 & u_5 & v_5 u'_5 & v_5 v'_5 & v_5 & u'_5 & v'_5 & 1 \\ u_6 u'_6 & u_6 v'_6 & u_6 & v_6 u'_6 & v_6 v'_6 & v_6 & u'_6 & v'_6 & 1 \\ u_7 u'_7 & u_7 v'_7 & u_7 & v_7 u'_7 & v_7 v'_7 & v_7 & u'_7 & v'_7 & 1 \\ u_8 u'_8 & u_8 v'_8 & u_8 & v_8 u'_8 & v_8 v'_8 & v_8 & u'_8 & v'_8 & 1 \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{12} \\ E_{13} \\ E_{21} \\ E_{22} \\ E_{23} \\ E_{31} \\ E_{32} \\ E_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Solve E_{11} to E_{32} --
The Essential Matrix

CSE152, Spring 2011

Intro Computer Vision

Properties of the Essential Matrix

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

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

CSE152, Spring 2011

Intro Computer Vision

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

Much more on multi-view in CSE252B!!

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

- Following the text, 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

$$\begin{aligned} &\text{Minimize:} \\ &\sum_{i=1}^n (\mathbf{p}_i^T \mathcal{E} \mathbf{p}_i)^2 \\ &\text{under the constraint} \\ &|\mathcal{E}|^2 = 1. \end{aligned}$$

CSE152, Spring 2011

Intro Computer Vision

The Fundamental Matrix

The epipolar constraint is given by: $\mathbf{p}^T \mathcal{E} \mathbf{p}' = 0$ with $\mathcal{E} = [\mathbf{t}_\times] \mathbf{R}$ where $\mathbf{p}$ and $\mathbf{p}'$ are 3-D coordinates of the image coordinates of points in the two images.

Without calibration, we can still identify corresponding points in two images, but we can't convert to 3-D coordinates. However, the relationship between the calibrated coordinates ($\mathbf{p}, \mathbf{p}'$) and uncalibrated image coordinates ($\mathbf{q}, \mathbf{q}'$) can be expressed as $\mathbf{p} = \mathbf{A} \mathbf{q}$, and $\mathbf{p}' = \mathbf{A}' \mathbf{q}'$

Therefore, we can express the epipolar constraint as:

$$(\mathbf{A} \mathbf{q})^T \mathcal{E} (\mathbf{A}' \mathbf{q}') = \mathbf{q}^T (\mathbf{A}^T \mathcal{E} \mathbf{A}') \mathbf{q}' = \mathbf{q}^T \mathbf{F} \mathbf{q}' = 0$$

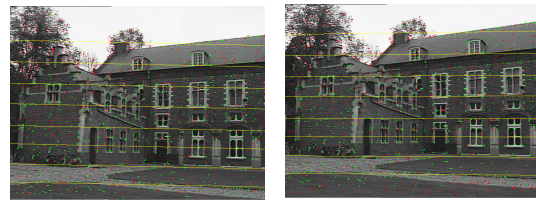
where $\mathbf{F}$ is called the Fundamental Matrix.

Can estimate $\mathbf{F}$ using 8 point algorithm WITHOUT CALIBRATION

CSE152, Spring 2011

Intro Computer Vision

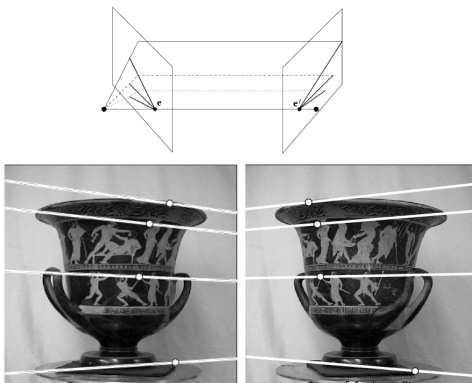
Epipolar geometry example



CSE152, Spring 2011

Intro Computer Vision

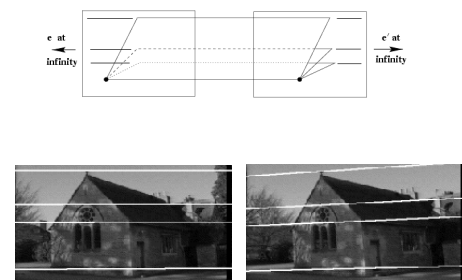
Example: converging cameras



CSE152, Spring 2011

courtesy of Andrew Zisserman

Example: motion parallel with image plane

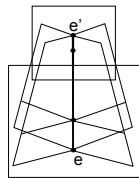
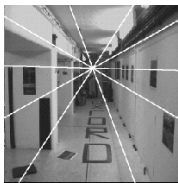
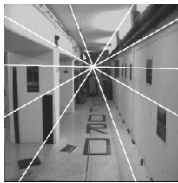


(simple for stereo → rectification)

courtesy of Andrew Zisserman

CSE152, Spring 2011

Example: forward motion



CSE152, Spring 2011

courtesy of Andrew Zisserman