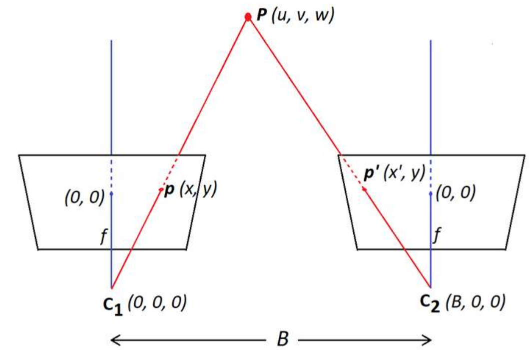
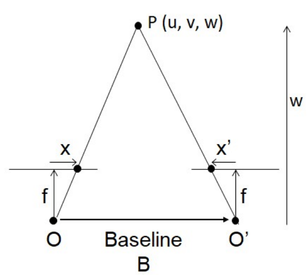
Computer Vision
and
Image Understanding
(Epipolar geometry)

Bhabatosh Chanda
chanda@isical.ac.in

Depth from disparity



Depth from disparity



$$depth = w = \frac{B \cdot f}{x - x'}$$

Depth w is inversely proportional to disparity $x - x'$!

World point to image point projection

- Intrinsic (parameter) matrix: $K = \begin{bmatrix} \varphi_x & 0 & d_x \\ 0 & \varphi_y & d_y \\ 0 & 0 & 1/f \end{bmatrix}$
- Extrinsic (parameter) matrix:

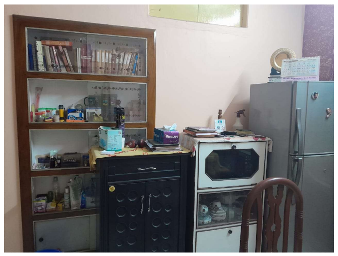
$$[R|t] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}$$

- Mapping world coordinate to image coordinate:

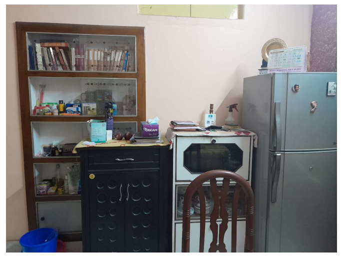
$$P^I_{3\times1} = K_{3\times3}[R_{3\times3}|t_{3\times1}]P^W_{4\times1} = A_{3\times4}P^W_{4\times1}$$

Binary stereo vision

Left image

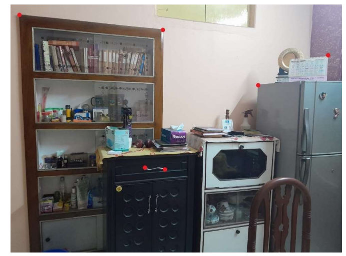


Right image

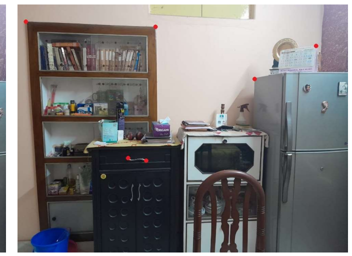


Binary stereo vision

Left image



Right image



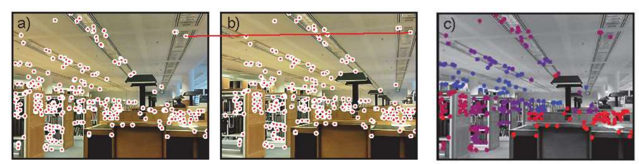
Binary stereo vision

Left image

Right image

Correspondence problem

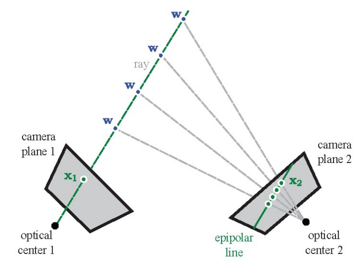
- How to detect points in the scene (object) whose coordinates need to be determined.
- How to establish correspondence between points (in different camera frames) which are images of same scene point.
- How to perform reliable and efficient search.



Two view geometry

- An image point x_1 in camera-1 forms a ray passing through optical centre.
- This line in 3D projects to a straight line, called *epipolar line*, 2D image plane of camera-2.
- For any point on the image plane of Camera-1, corresponding point lies on its epipolar line on the image plane of Camera-2 and *vice versa*. This is *epipolar constraint*.

Epipolar geometry

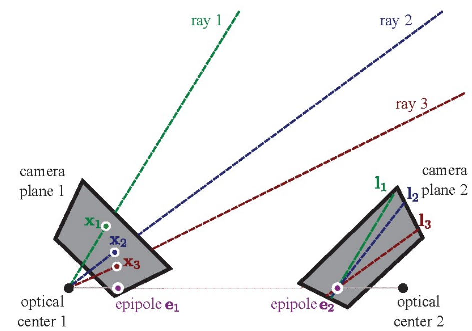


Epipolar constraint:
For any point in the first image, the corresponding point in the second image is constrained to lie on a line.

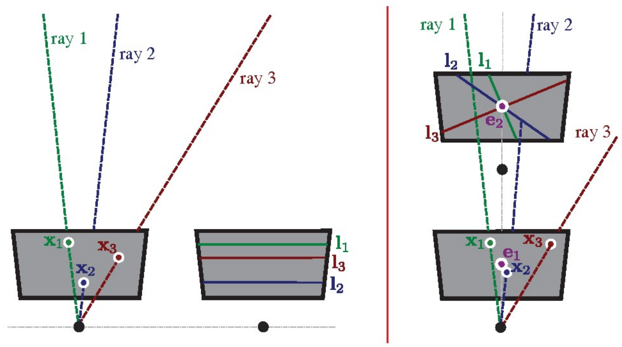
Epipolar geometry

- Epipolar constraint has two practical implications:
  - We can find correspondence easily.
  - Given intrinsic parameters, extrinsic parameters can be determined from correspondence pattern.
- All epipolar lines in image plane of Camera-2, due to image points in Camera-1, meet at a point.
  - This meeting points is called *epipole*.
- Epipole is the image of optical centre of Camera-1 on the image plane of Camera-2, and *vice versa*.
- Epipole *may not be observed* in the image plane.

Epipole



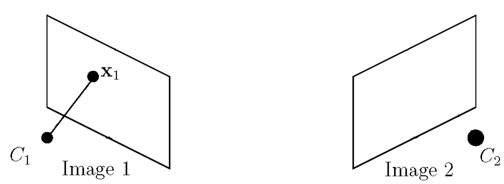
Epipolar lines



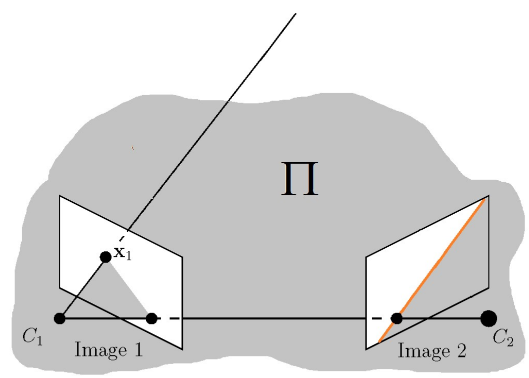
Determining epipolar line

- Consider a plane Π formed by
  - (i) an image point in Camera-1,
  - (ii) optical centre of Camera-1 and
  - (iii) optical centre of Camera-2.
- Epipolar line is formed by cross-section of plane Π and image plane of Camera-2.

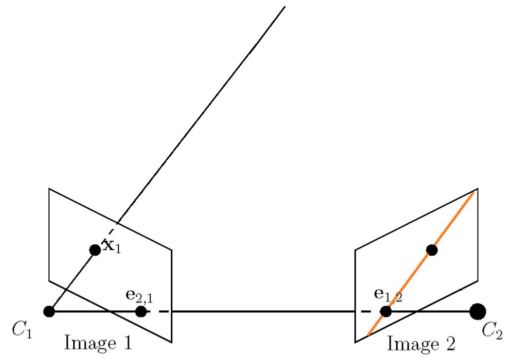
Determining epipolar line



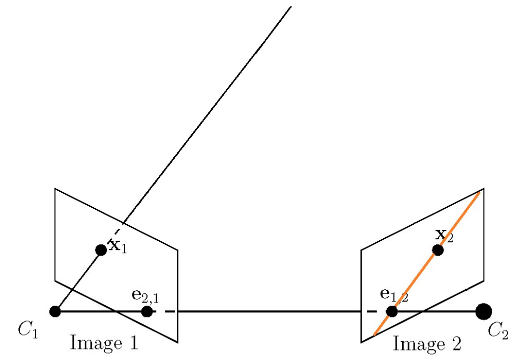
Determining epipolar line



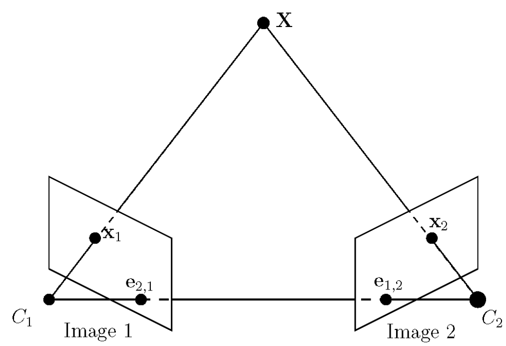
Determining epipolar line



Determining epipolar line



Determining epipolar line



Determining epipolar line

- Consider a plane Π formed by
  - (i) an image point in Camera-1,
  - (ii) optical centre of Camera-1 and
  - (iii) optical centre of Camera-2.
- Epipolar line is formed by cross-section of plane Π and image plane of Camera-2.
- A simpler way to determine epipolar line is by means of *essential matrix*.

Normalized camera

- Focal length parameter *f=1*
- Drift of origin $(d_x, d_y) = (0,0)$
- Pixel (sensor) density $\varphi_x = \varphi_y = 1$
- Intrinsic matrix becomes

$$K = \begin{bmatrix} \varphi_x & 0 & d_x/f \\ 0 & \varphi_y & d_y/f \\ 0 & 0 & 1/f \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Normalized camera point

- Mapping world point to real camera point:

$$P_h^I = K[R|t]P_h^W \qquad \text{where } P_h^I = \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \text{ and } P_h^W = \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}$$

- Normalized camera point may be obtained as

$$\hat{P}_h^I = K^{-1}P_h^I = [R|t]P_h^W$$

- This leads to

$$\hat{P}_h^I = [R|t]P_h^W \;\Rightarrow\; \hat{P}_h^I = RP^W + t$$

Epipolar line

- For Camera-1 (at origin of world coordinate and axes aligned with world coordinate axes):

$$\hat{P}_h{}^{I_1} = IP^W + 0 = P^W$$

- For Camera-2 (at any arbitrary location and direction with respect to amera-1):

$$\hat{P}_h{}^{I_2} = RP^W + t \qquad \Rightarrow \hat{P}_h{}^{I_2} = R\hat{P}_h{}^{I_1} + t$$

- Taking cross-product with *t*:

$$t \times \hat{P}_h{}^{I_2} = t \times R\hat{P}_h{}^{I_1}$$

Essential matrix

- Taking cross-product with *t*:

$$t \times \hat{P}_h{}^{I_2} = t \times R\hat{P}_h{}^{I_1}$$

- Taking dot-product with $\hat{P}_h^{I_2}$:

$$\hat{P}_h{}^{I_2} \cdot (t \times \hat{P}_h{}^{I_2}) = \hat{P}_h{}^{I_2} \cdot (t \times R\hat{P}_h{}^{I_1})$$

$$(\hat{P}_h{}^{I_2})^T (t \times R\hat{P}_h{}^{I_1}) = 0$$

- '$t \times$' is equivalent to multiplication with matrix T' , i.e.,

$$T' = \begin{bmatrix} 0 & -t_w & t_v \\ t_w & 0 & -t_u \\ -t_v & t_u & 0 \end{bmatrix}$$

Essential matrix and Epipolar line

- '$t \times$' is equivalent to multiplication with matrix T'

$$T' = \begin{bmatrix} 0 & -t_w & t_v \\ t_w & 0 & -t_u \\ -t_v & t_u & 0 \end{bmatrix}$$

- This converts $(\hat{P}_h{}^{I_2})^T (t \times R) \hat{P}_h{}^{I_1} = 0$ to

$$\left(\hat{P}_h{}^{I_2}\right)^T E \left(\hat{P}_h{}^{I_1}\right) = 0$$

- The matrix $E = T'R$ is called *Essential matrix*.

Essential matrix and Epipolar line

- So we have essential matrix $E = T'R$

  where $(\hat{P}_h{}^{I_2})^T E \hat{P}_h{}^{I_1} = 0$

- Epipolar lines:

$$l_1 \equiv (\hat{P}_h{}^{I_2})^T E \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \text{ and}$$

$$l_2 \equiv (\hat{P}_h{}^{I_1})^T E^T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0$$

Epipoles

- The epipoles can be retrieved by computing the singular value decomposition of the essential matrix

$$E_{3\times3} = U_{3\times3} L_{3\times3} V_{3\times3}^T$$

- Epipole in image-1 is the last column of *V*

$$e_1 = (v_{13} \quad v_{23} \quad v_{33})$$

- Epipole in image-2 is the last row of *U*

$$e_2 = (u_{31} \quad u_{32} \quad u_{33})$$

Epipolar lines and Fundamental matrix

- Real camera point and normalized camera point are related by

$$\hat{P}_h^I = K^{-1} P_h^I$$

- Normalized Camera-1 and camera-2 points are related by essential matrix *E* as

$$(\hat{P}_h{}^{I_2})^T E \hat{P}_h{}^{I_1} = 0$$

- Combining them:

$$(K_2^{-1} P_h{}^{I_2})^T E K_1^{-1} P_h{}^{I_1} = 0$$

$$(P_h{}^{I_2})^T (K_2^{-1})^T E K_1^{-1} P_h{}^{I_1} = 0 \quad \equiv (P_h{}^{I_2})^T F P_h{}^{I_1} = 0$$

Epipolar lines and Fundamental matrix

- From $(P_h^{I_2})^T (K_2^{-1})^T E K_1^{-1} P_h^{I_1} = 0$

we simplify to $(P_h^{I_2})^T F P_h^{I_1} = 0$

- The matrix $F = (K_2^{-1})^T E K_1^{-1}$ is *Fundamental matrix*.
- Epipolar lines:

$$l_1 \equiv (P_h^{I_2})^T F \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \qquad l_2 \equiv (P_h^{I_1})^T F^T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0$$

Fundamental matrix

- Fundamental matrix: $F = (K_2^{-1})^T E K_1^{-1}$
- Each pair of points (one from camera-1 and other from camera-2) produces one equation:

$$(P_h^{I_2})^T F P_h^{I_1} = 0$$

- So 8-pairs of correspondence points are sufficient to find fundamental matrix *F* between two cameras.
- This is known as *8-point algorithm*.
- If K_1 and K_2 are known *Essential matrix* E can be estimated from Fundamental matrix as $E = (K_2)^T F K_1$

Binocular stereo reconstruction

1. Compute image features.
2. Compute feature descriptors.
3. Find initial matches.
4. Compute fundamental matrix.
5. Refine matches.
6. Estimate essential matrix.
7. Decompose essential matrix.
8. Estimate 3D points.

Binocular stereo reconstruction

1. Compute image features.
2. Compute feature descriptors.
3. Find initial matches.
4. Compute fundamental matrix.
5. Refine matches.
6. Estimate essential matrix.
7. Decompose essential matrix.
8. Estimate 3D points.

Thank you !

Any question?