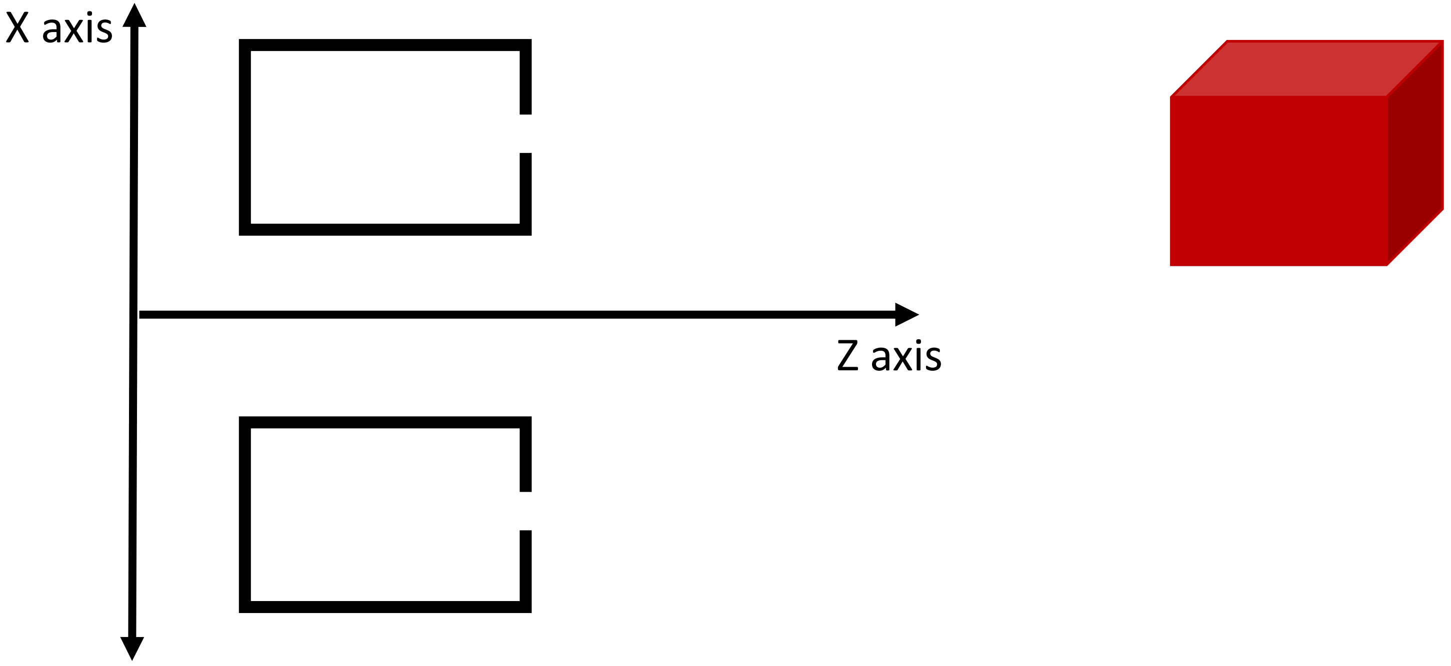
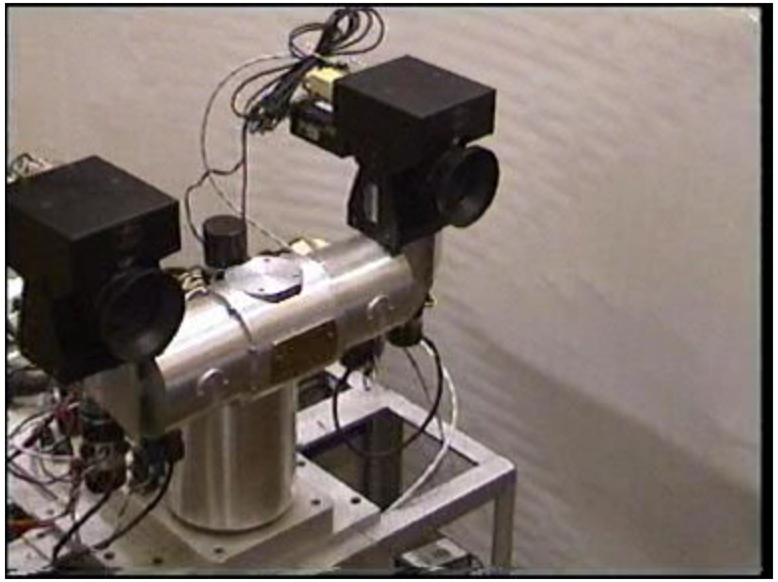
Epipolar geometry

Binocular stereo

- Special case: cameras are parallel to each other and translated along X axis



Stereo head



Kinect / depth cameras

Perspective projection in rectified cameras

- Without loss of generality, assume origin is at pinhole of 1st camera

X coordinate differs by t_x/Z

$$x_1 = \frac{X}{Z} \qquad x_2 = \frac{X + t_x}{Z}$$

$$y_1 = \frac{Y}{Z} \qquad y_2 = \frac{Y}{Z}$$

Y coordinate is the same!

Perspective projection in rectified cameras

- X coordinate differs by t_x/Z
- That is, difference in X coordinate is *inversely proportional to depth*
- Difference in X coordinate is called *disparity*
- Translation between cameras (tx) is called *baseline*

- *disparity = baseline / depth*

The NCC cost volume

- Consider M x N image
- Suppose there are D possible disparities.
- For every pixel, D possible scores
- Can be written as an M x N x D array
- To get disparity, take max along 3rd axis

Computing the NCC volume

1. For every pixel (x, y)
    1. For every disparity d
        1. Get normalized patch from image 1 at (x, y)
        2. Get normalized patch from image 2 at (x + d, y)
        3. Compute NCC

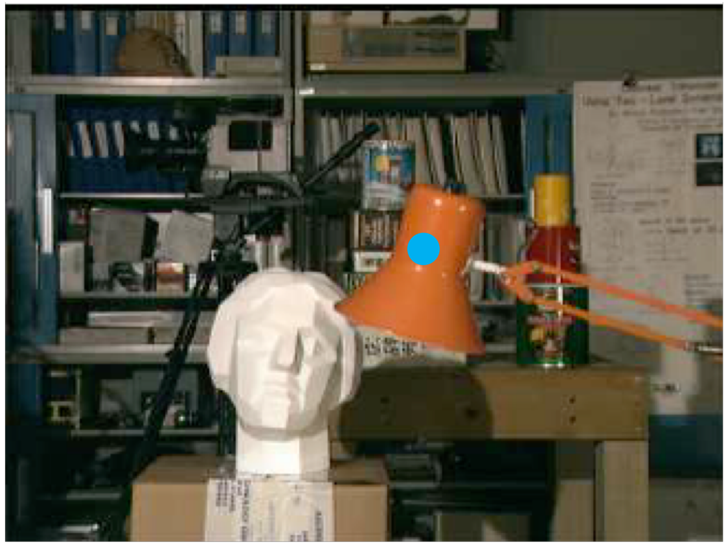
Computing the NCC volume

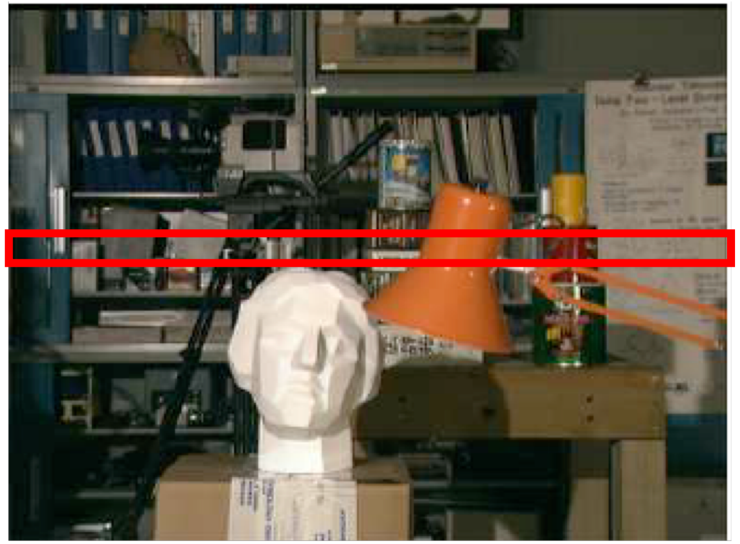
1. For every disparity d
    1. For every pixel (x, y)
        1. Get normalized patch from image 1 at (x, y)
        2. Get normalized patch from image 2 at (x + d, y)
        3. Compute NCC

Assume all pixels lie at same disparity d (i.e., lie on same plane) and compute cost for each

Plane sweep stereo

Perspective projection in rectified cameras





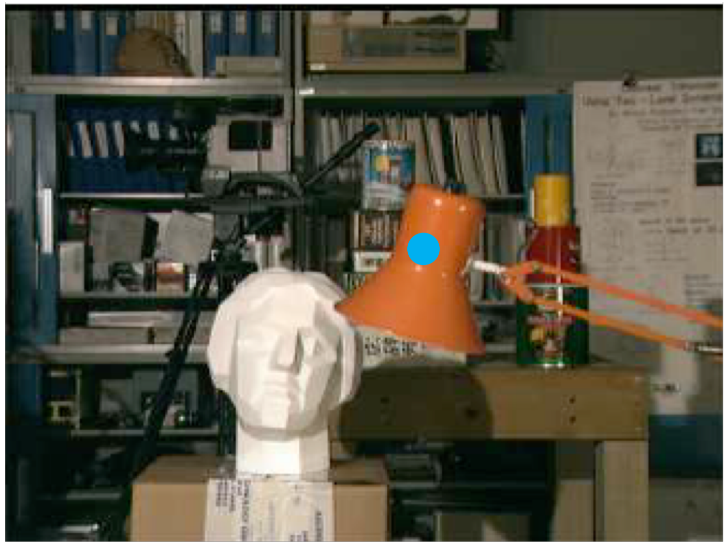
- For rectified cameras, correspondence problem is easier
- Only requires searching along a particular *row.*

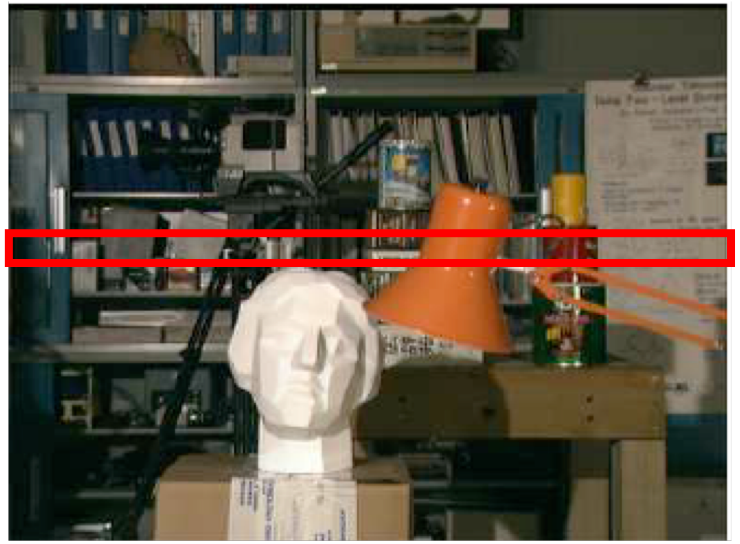
Perspective projection in rectified cameras

What about non-rectified cameras? Is there an equivalent?

- For re… problem… easi…
- Only requi… c… along … particular *row.*

Epipolar constraint





- *Reduces 2D search problem to search along a particular line: epipolar line*

Epipolar constraint

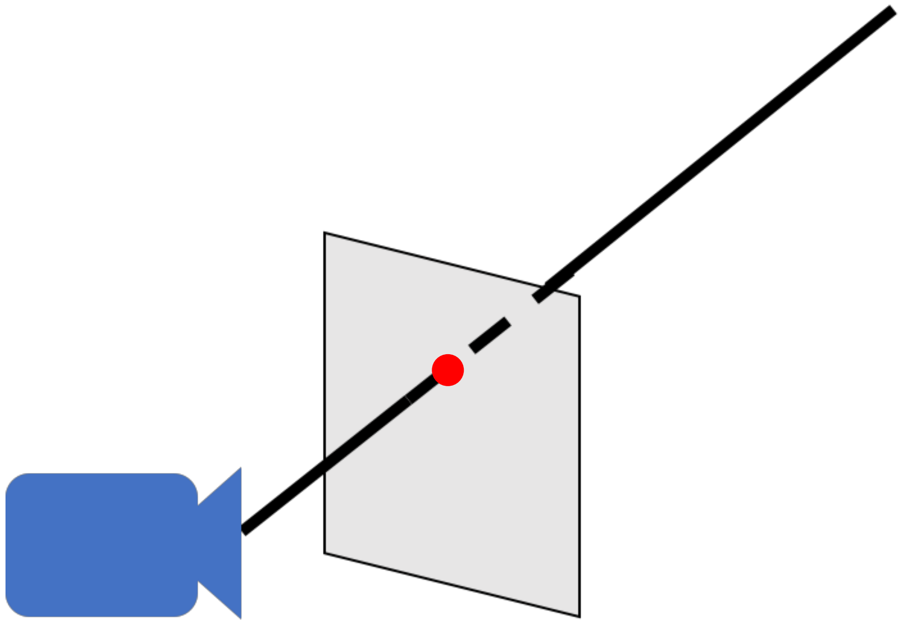
True in general!

- Given pixel (x,y) in one image, corresponding pixel in the other image must lie on a line
- Line function of (x,y) and parameters of camera
- These lines are called *epipolar line*

Epipolar geometry

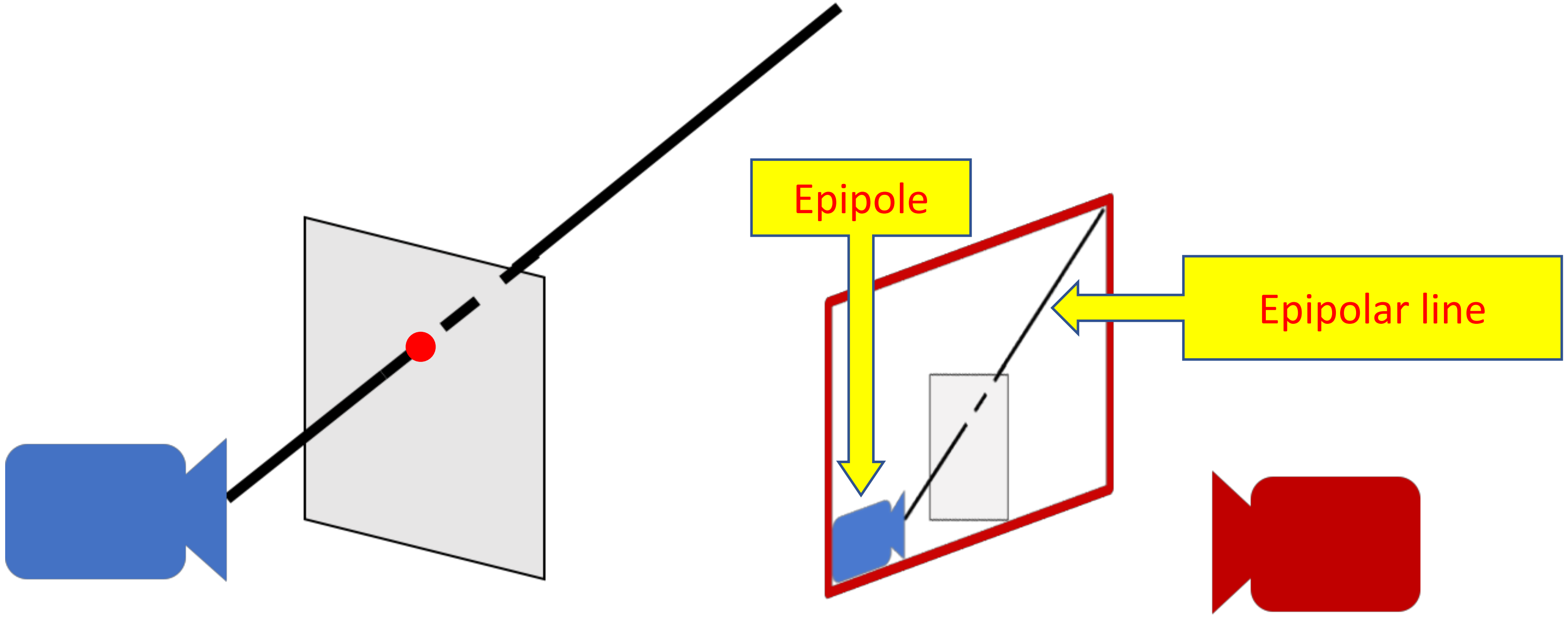
Epipolar geometry - why?

- For a single camera, pixel in image plane must correspond to point somewhere along a ray



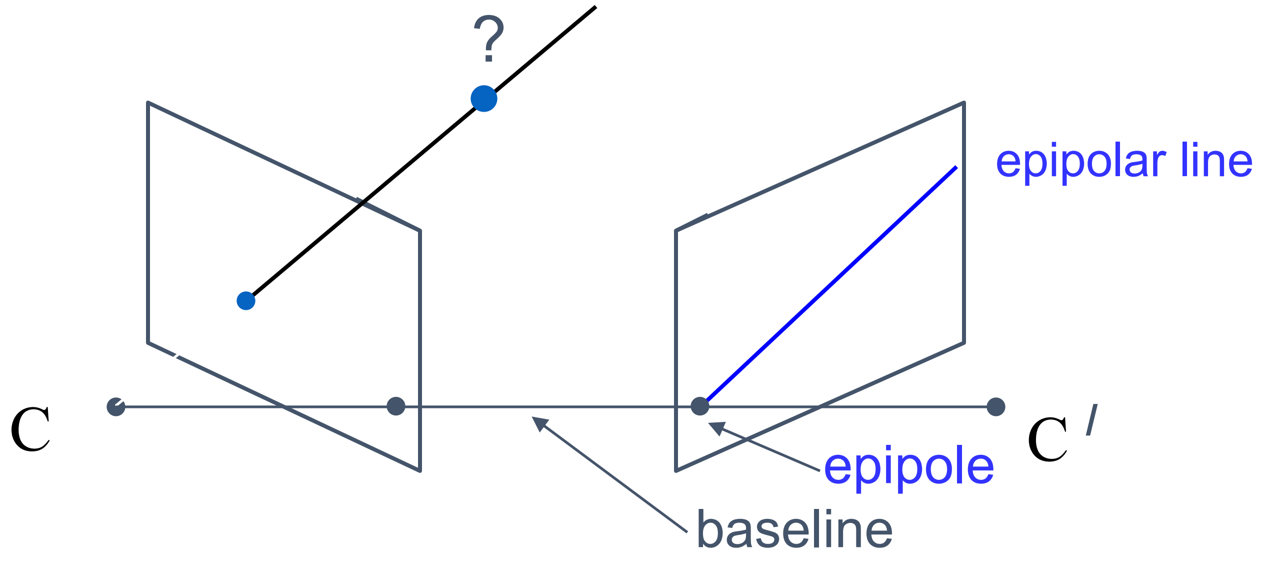
Epipolar geometry

- When viewed in second image, this ray looks like a line: *epipolar line*
- The epipolar line must pass through image of the first camera in the second image - *epipole*



Epipolar geometry

Given an image point in one view, where is the corresponding point in the other view?



- A point in one view "generates" an epipolar line in the other view
- The corresponding point lies on this line

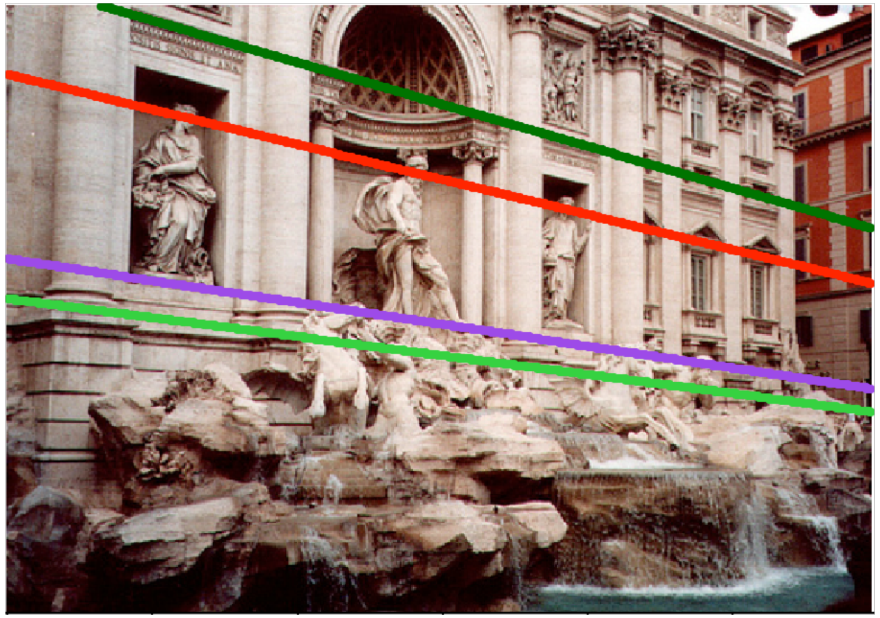
Epipolar line

Epipolar constraint

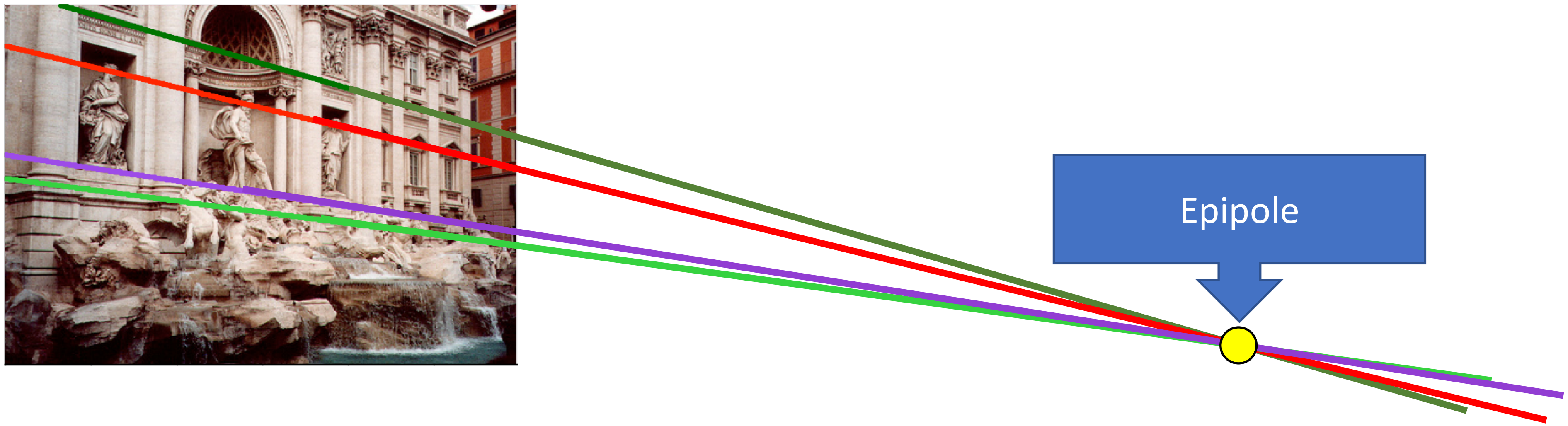
- Reduces correspondence problem to 1D search along an epipolar line

Epipolar lines

Epipolar lines

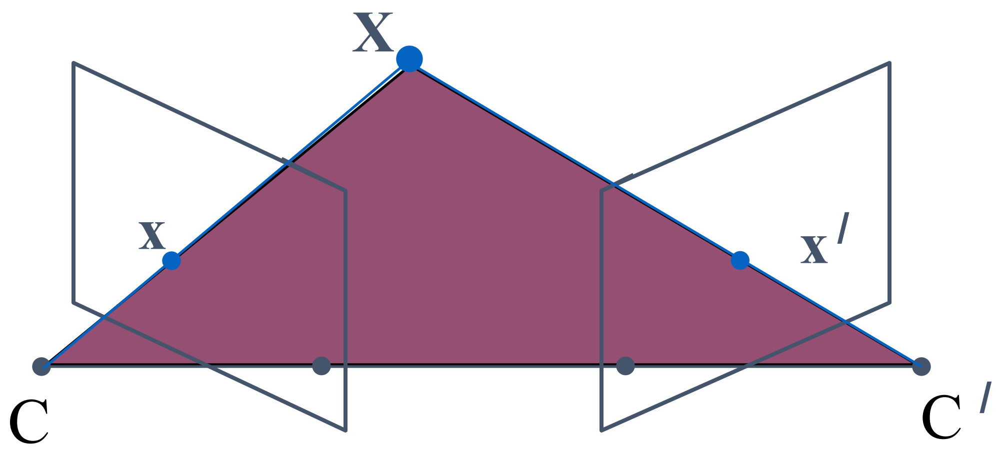


Epipolar lines



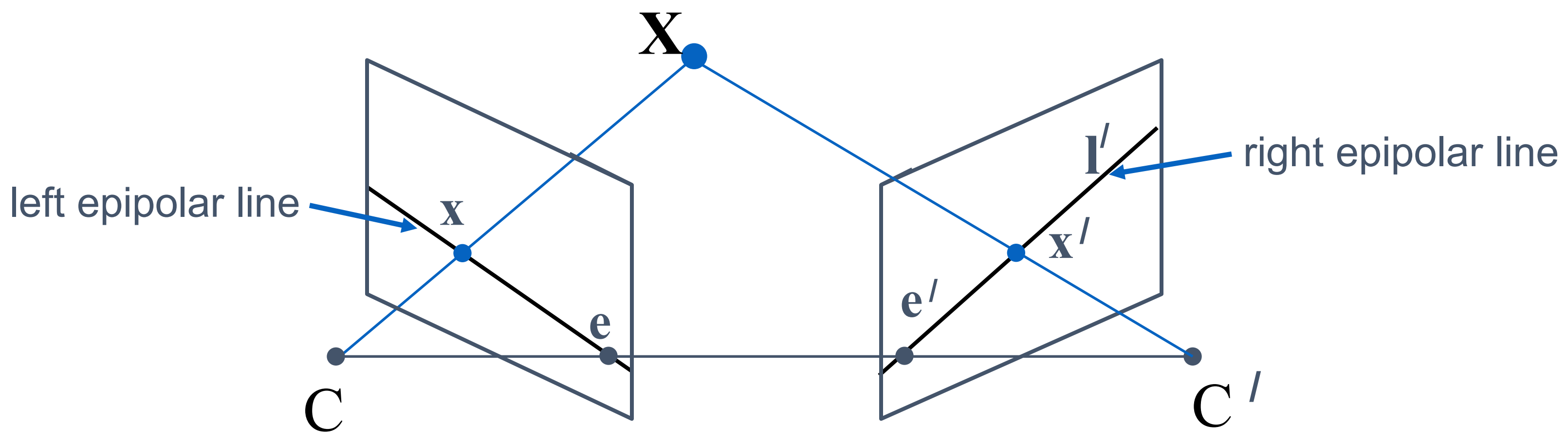
Epipolar geometry continued

Epipolar geometry is a consequence of the coplanarity of the camera centres and scene point



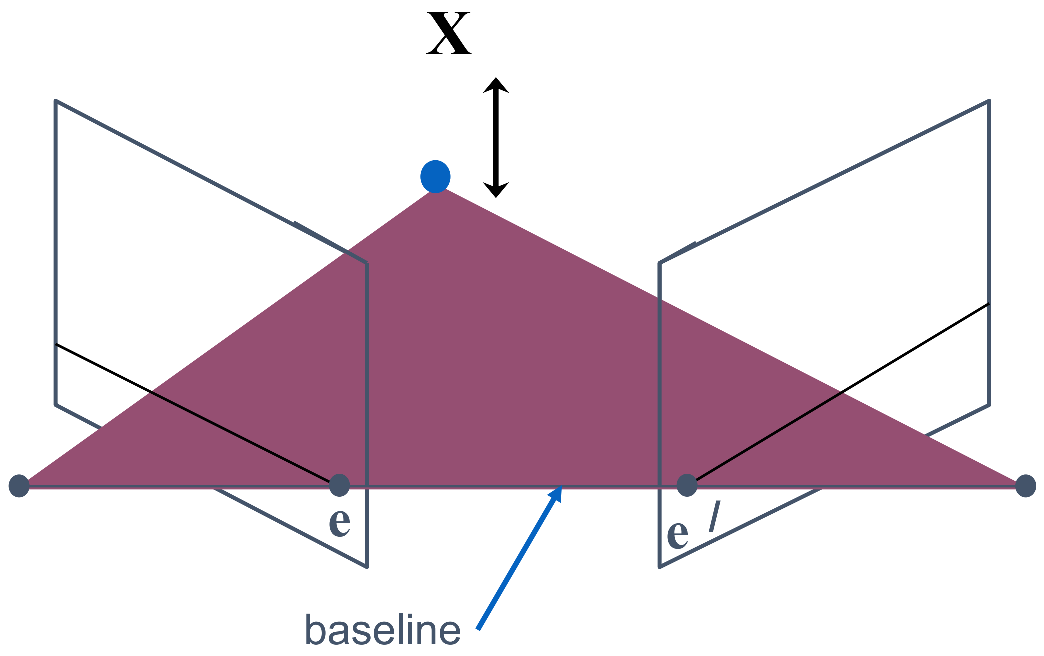
The camera centres, corresponding points and scene point lie in a single plane, known as the epipolar plane

Nomenclature



- The epipolar line $\mathbf{l}'$ is the image of the ray through $\mathbf{x}$
- The epipole $\mathbf{e}$ is the point of intersection of the line joining the camera centres with the image plane
  - this line is the baseline for a stereo rig, and
  - the translation vector for a moving camera
- The epipole is the image of the centre of the other camera: $\mathbf{e} = \mathrm{P}\mathbf{C}'$, $\mathbf{e}' = \mathrm{P}'\mathbf{C}$

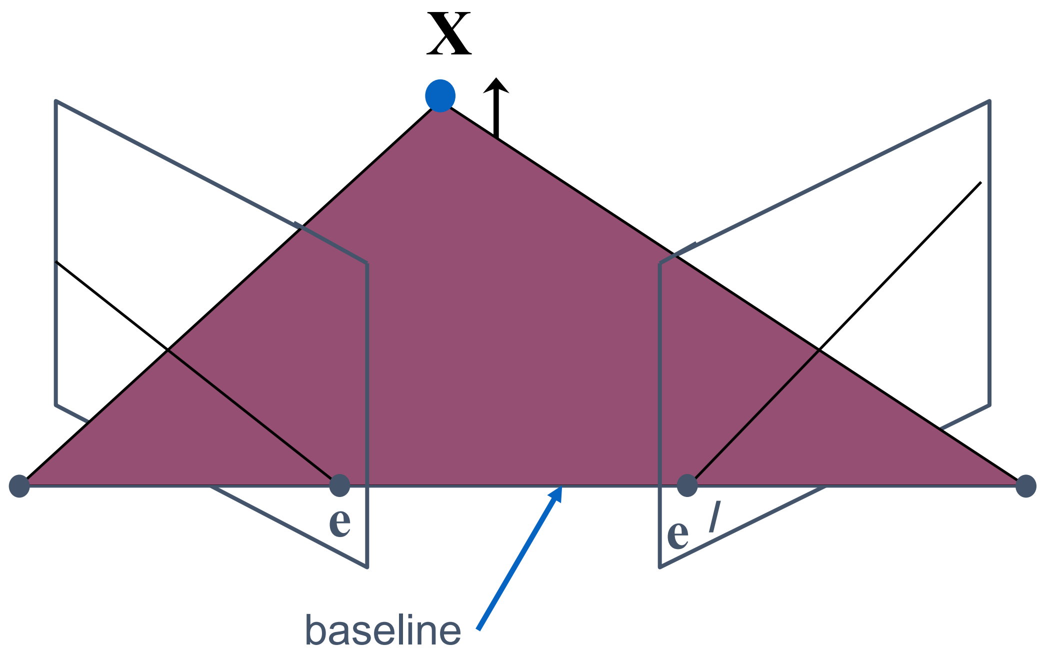
The epipolar pencil



As the position of the 3D point $\mathbf{X}$ varies, the epipolar planes “rotate” about the baseline. This family of planes is known as an epipolar pencil (a pencil is a one parameter family).

All epipolar lines intersect at the epipole.

The epipolar pencil



As the position of the 3D point $\mathbf{X}$ varies, the epipolar planes “rotate” about the baseline. This family of planes is known as an epipolar pencil (a pencil is a one parameter family).

All epipolar lines intersect at the epipole.

Epipolar geometry - the math

- Assume intrinsic parameters K are identity
- Assume world coordinate system is centered at 1st camera pinhole with Z along viewing direction

$$\vec{\mathbf{x}}_{img}^{(1)} \equiv K_1 \begin{bmatrix} R_1 & \mathbf{t}_1 \end{bmatrix} \vec{\mathbf{x}}_w$$

$$\vec{\mathbf{x}}_{img}^{(2)} \equiv K_2 \begin{bmatrix} R_2 & \mathbf{t}_2 \end{bmatrix} \vec{\mathbf{x}}_w$$

Epipolar geometry - the math

- Assume intrinsic parameters K are identity
- Assume world coordinate system is centered at 1st camera pinhole with Z along viewing direction

$$\vec{\mathbf{x}}^{(1)}_{img} \equiv \begin{bmatrix} I & 0 \end{bmatrix} \vec{\mathbf{x}}_w$$

$$\vec{\mathbf{x}}^{(2)}_{img} \equiv \begin{bmatrix} R & \mathbf{t} \end{bmatrix} \vec{\mathbf{x}}_w$$

Epipolar geometry - the math

- Assume intrinsic parameters K are identity
- Assume world coordinate system is centered at 1st camera pinhole with Z along viewing direction

$$\vec{\mathbf{x}}_{img}^{(1)} \equiv \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_w \\ 1 \end{bmatrix} = \mathbf{x}_w$$

$$\vec{\mathbf{x}}_{img}^{(2)} \equiv \begin{bmatrix} R & \mathbf{t} \end{bmatrix} \begin{bmatrix} \mathbf{x}_w \\ 1 \end{bmatrix} = R\mathbf{x}_w + \mathbf{t}$$

Epipolar geometry - the math

- Assume intrinsic parameters K are identity
- Assume world coordinate system is centered at 1st camera pinhole with Z along viewing direction

$$\vec{\mathbf{x}}_{img}^{(1)} \equiv \mathbf{x}_w$$

$$\vec{\mathbf{x}}_{img}^{(2)} \equiv R\mathbf{x}_w + \mathbf{t}$$

Epipolar geometry - the math

- Assume intrinsic parameters K are identity
- Assume world coordinate system is centered at 1st camera pinhole with Z along viewing direction

$$\lambda_1 \vec{\mathbf{x}}_{img}^{(1)} = \mathbf{x}_w$$

$$\lambda_2 \vec{\mathbf{x}}_{img}^{(2)} = R\mathbf{x}_w + \mathbf{t}$$

Epipolar geometry - the math

$$\lambda_2 \vec{\mathbf{x}}_{img}^{(2)} = \lambda_1 R \vec{\mathbf{x}}_{img}^{(1)} + \mathbf{t}$$

$$\lambda_2 \mathbf{t} \times \vec{\mathbf{x}}_{img}^{(2)} = \lambda_1 \mathbf{t} \times R \vec{\mathbf{x}}_{img}^{(1)} + \underbrace{\mathbf{t} \times \mathbf{t}}_{0}$$

$$\lambda_2 \mathbf{t} \times \vec{\mathbf{x}}_{img}^{(2)} = \lambda_1 \mathbf{t} \times R \vec{\mathbf{x}}_{img}^{(1)}$$

$$\underbrace{\cancel{\lambda_2 \vec{\mathbf{x}}_{img}^{(2)} \cdot \mathbf{t} \times \vec{\mathbf{x}}_{img}^{(2)}}}_{0} = \lambda_1 \vec{\mathbf{x}}_{img}^{(2)} \cdot \mathbf{t} \times R \vec{\mathbf{x}}_{img}^{(1)}$$

$$0 = \lambda_1 \vec{\mathbf{x}}_{img}^{(2)} \cdot \mathbf{t} \times R \vec{\mathbf{x}}_{img}^{(1)}$$

Epipolar geometry - the math

$$\vec{\mathbf{x}}_{img}^{(2)} \cdot \mathbf{t} \times R\vec{\mathbf{x}}_{img}^{(1)} = 0$$

- Can we write this as matrix vector operations?
- Cross product can be written as a matrix

$$[\mathbf{t}]_\times = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

$$[\mathbf{t}]_\times \mathbf{a} = \mathbf{t} \times \mathbf{a}$$

Epipolar geometry - the math

$$\vec{\mathbf{x}}^{(2)}_{img} \cdot [\mathbf{t}]_{\times} R \vec{\mathbf{x}}^{(1)}_{img} = 0$$

- Can we write this as matrix vector operations?
- Dot product can be written as a vector-vector times

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b}$$

Epipolar geometry - the math

$$\vec{\mathbf{x}}_{img}^{(2)} \cdot [\mathbf{t}]_{\times} R \vec{\mathbf{x}}_{img}^{(1)} = 0$$

- Can we write this as matrix vector operations?
- Dot product can be written as a vector-vector times

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b}$$

Epipolar geometry - the math

$$\vec{\mathbf{x}}_{img}^{(2)T}[\mathbf{t}]_\times R\vec{\mathbf{x}}_{img}^{(1)} = 0$$

$$\vec{\mathbf{x}}_{img}^{(2)T} E\vec{\mathbf{x}}_{img}^{(1)} = 0$$

Epipolar geometry - the math

Homogenous coordinates of point in image 2

Homogenous coordinates of point in image 1

$$\vec{\mathbf{x}}_{img}^{(2)T} E \vec{\mathbf{x}}_{img}^{(1)} = 0$$

Essential matrix

Epipolar constraint and epipolar lines

$$\vec{\mathbf{x}}_{img}^{(2)T} E \vec{\mathbf{x}}_{img}^{(1)} = 0$$

- Consider a known, fixed pixel in the first image
- What constraint does this place on the corresponding pixel?
- $\vec{\mathbf{x}}_{img}^{(2)T} \mathbf{l} = 0$ where $\mathbf{l} = E \vec{\mathbf{x}}_{img}^{(1)}$
- What kind of equation is this?

Epipolar constraint and epipolar lines

$$\vec{\mathbf{x}}_{img}^{(2)T} E \vec{\mathbf{x}}_{img}^{(1)} = 0$$

- Consider a known, fixed pixel in the first image
- $\vec{\mathbf{x}}_{img}^{(2)T} \mathbf{l} = 0$ where $\mathbf{l} = E\vec{\mathbf{x}}_{img}^{(1)}$

$$\vec{\mathbf{x}}_{img}^{(2)T} \mathbf{l} = 0$$

$$\Rightarrow \begin{bmatrix} x_2 & y_2 & 1 \end{bmatrix} \begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix} = 0$$

$$\Rightarrow l_x x_2 + l_y y_2 + l_z = 0$$

Line!

Epipolar constraint: putting it all together

- If **p** is a pixel in first image and **q** is the corresponding pixel in the second image, then:
  $\mathbf{q}^T E \mathbf{p} = 0$
- $E = [\mathbf{t}]_X R$
- For fixed **p**, **q** must satisfy:
  $\mathbf{q}^T \mathbf{l} = 0$, where $\mathbf{l} = E\mathbf{p}$
- For fixed **q, p** must satisfy:
  $\mathbf{l}^T \mathbf{p} = 0$ where $\mathbf{l}^T = \mathbf{q}^T E$, or $\mathbf{l} = E^t \mathbf{q}$
- These are epipolar lines!

Epipolar line in 2nd image

Epipolar line in 1st image

Essential matrix and epipoles

- E = $[\mathbf{t}]_X R$

$$
\begin{aligned}
&\vec{\mathbf{c}_2} = \mathbf{t} \\
&\vec{\mathbf{c_2}}^T E = \mathbf{t}^T E = \mathbf{t}^T [\mathbf{t}]_\times R = 0 \\
&\vec{\mathbf{c}_2}^T E\mathbf{p} = 0 \ \ \forall \mathbf{p}
\end{aligned}
$$

- E**p** is an epipolar line in 2nd image
- All epipolar lines in second image pass through c_2
- c_2 is epipole in 2nd image

Essential matrix and epipoles

- $E = [\mathbf{t}]_X R$

$$\vec{\mathbf{c}_1} = \mathbf{R}^T \mathbf{t}$$

$$E\vec{\mathbf{c}_1} = [\mathbf{t}]_\times R R^T \mathbf{t} = [\mathbf{t}]_\times \mathbf{t} = 0$$

$$\mathbf{q}^T E \vec{\mathbf{c}_1} = 0 \;\; \forall \mathbf{q}$$

- $E^T\mathbf{q}$ is an epipolar line in 1^{st} image
- All epipolar lines in first image pass through c_1
- c_1 is the epipole in 1^{st} image

Epipolar geometry - the math

- We assumed that intrinsic parameters K are identity
- What if they are not?

$$\vec{\mathbf{x}}_{img}^{(1)} \equiv K_1 \begin{bmatrix} R_1 & \mathbf{t}_1 \end{bmatrix} \vec{\mathbf{x}}_w$$

$$\vec{\mathbf{x}}_{img}^{(2)} \equiv K_2 \begin{bmatrix} R_2 & \mathbf{t}_2 \end{bmatrix} \vec{\mathbf{x}}_w$$

Fundamental matrix

$$\vec{\mathbf{x}}_{img}^{(1)} \equiv K_1 \begin{bmatrix} I & 0 \end{bmatrix} \vec{\mathbf{x}}_w$$
$$\vec{\mathbf{x}}_{img}^{(2)} \equiv K_2 \begin{bmatrix} R & \mathbf{t} \end{bmatrix} \vec{\mathbf{x}}_w$$

Fundamental matrix

$$\lambda_1 \vec{\mathbf{x}}_{img}^{(1)} = K_1 \begin{bmatrix} I & \mathbf{0} \end{bmatrix} \vec{\mathbf{x}}_w$$

$$\lambda_2 \vec{\mathbf{x}}_{img}^{(2)} = K_2 \begin{bmatrix} R & \mathbf{t} \end{bmatrix} \vec{\mathbf{x}}_w$$

Fundamental matrix

$$
\begin{aligned}
\lambda_1 \vec{\mathbf{x}}_{img}^{(1)} =& K_1 \begin{bmatrix} I & \mathbf{0} \end{bmatrix} \vec{\mathbf{x}}_w \\
=& K_1 \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_w \\ 1 \end{bmatrix} \\
=& K_1 \mathbf{x}_w
\end{aligned}
$$

$$
\Rightarrow \lambda_1 K_1^{-1} \vec{\mathbf{x}}_{img}^{(1)} = \mathbf{x}_w
$$

Fundamental matrix

$$
\begin{aligned}
\lambda_2 \vec{\mathbf{x}}_{img}^{(2)} =& K_2 \begin{bmatrix} R & \mathbf{t} \end{bmatrix} \begin{bmatrix} \mathbf{x}_w \\ 1 \end{bmatrix} \\
=& K_2 R \mathbf{x}_w + K_2 \mathbf{t} \\
=& \lambda_1 K_2 R K_1^{-1} \vec{\mathbf{x}}_{img}^{(1)} + K_2 \mathbf{t}
\end{aligned}
$$

$$\Rightarrow \lambda_2 K_2^{-1} \vec{\mathbf{x}}_{img}^{(2)} = \lambda_1 R K_1^{-1} \vec{\mathbf{x}}_{img}^{(1)} + \mathbf{t}$$

$$\Rightarrow \lambda_2 [\mathbf{t}]_\times K_2^{-1} \vec{\mathbf{x}}_{img}^{(2)} = \lambda_1 [\mathbf{t}]_\times R K_1^{-1} \vec{\mathbf{x}}_{img}^{(1)}$$

$$\Rightarrow 0 = \vec{\mathbf{x}}_{img}^{(2)} K_2^{-T} [\mathbf{t}]_\times R K_1^{-1} \vec{\mathbf{x}}_{img}^{(1)}$$

Fundamental matrix

$$\Rightarrow 0 = \vec{\mathbf{x}}_{img}^{(2)} K_2^{-T} [\mathbf{t}]_\times R K_1^{-1} \vec{\mathbf{x}}_{img}^{(1)}$$

$$\Rightarrow 0 = \vec{\mathbf{x}}_{img}^{(2)} F \vec{\mathbf{x}}_{img}^{(1)}$$

Fundamental matrix

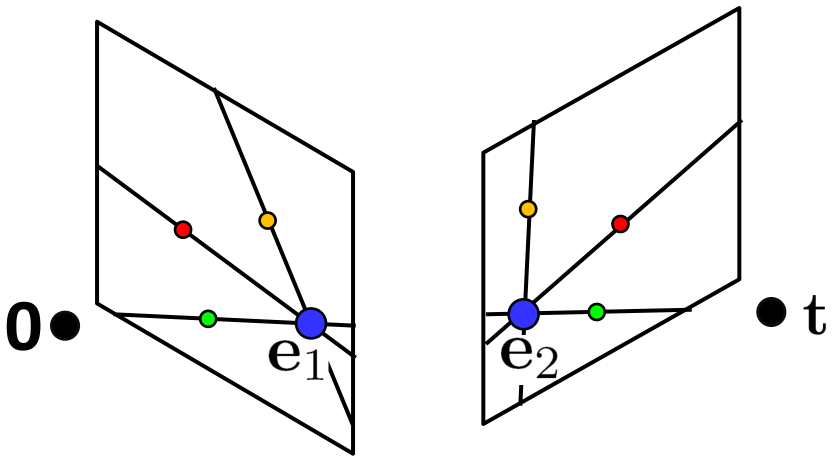
Fundamental matrix result

$$\mathbf{q}^T \mathbf{F} \mathbf{p} = 0$$

(Longuet-Higgins, 1981)

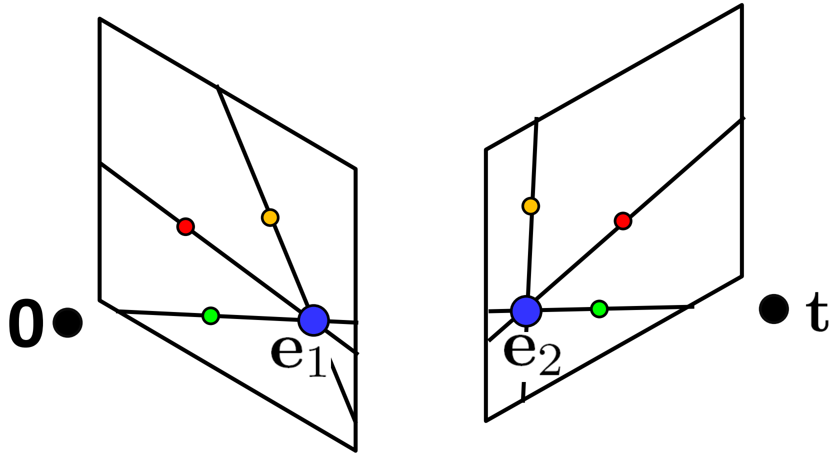
Properties of the Fundamental Matrix

- $\mathbf{F}\mathbf{p}$ is the epipolar line associated with $\mathbf{p}$
- $\mathbf{F}^T\mathbf{q}$ is the epipolar line associated with $\mathbf{q}$



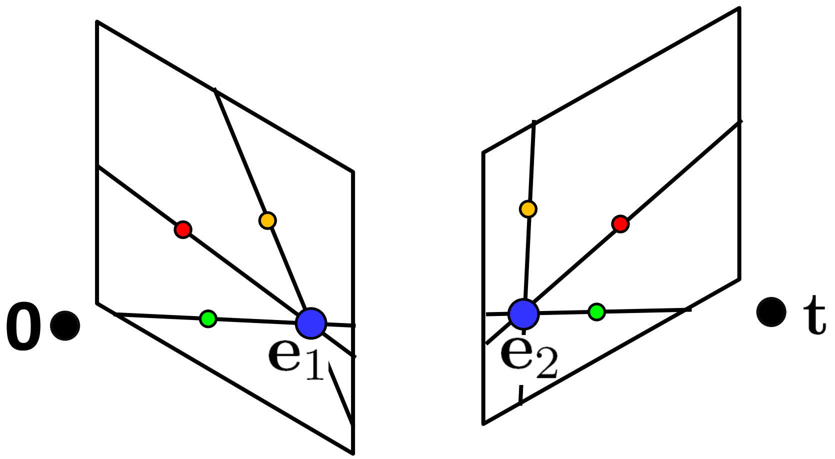
Properties of the Fundamental Matrix

- $\mathbf{F}\mathbf{p}$ is the epipolar line associated with $\mathbf{p}$
- $\mathbf{F}^T\mathbf{q}$ is the epipolar line associated with $\mathbf{q}$
- $\mathbf{F}\mathbf{e}_1 = \mathbf{0}$ and $\mathbf{F}^T\mathbf{e}_2 = \mathbf{0}$
- All epipolar lines contain epipole



Properties of the Fundamental Matrix

- $\mathbf{F}\mathbf{p}$ is the epipolar line associated with $\mathbf{p}$
- $\mathbf{F}^T\mathbf{q}$ is the epipolar line associated with $\mathbf{q}$
- $\mathbf{F}\mathbf{e}_1 = \mathbf{0}$ and $\mathbf{F}^T\mathbf{e}_2 = \mathbf{0}$
- $\mathbf{F}$ is rank 2



Why is F rank 2?

- F is a 3 x 3 matrix
- But there is a vector c_1 and c_2 such that $Fc_1 = 0$ and $F^Tc_2 = 0$

Fundamental matrix song

Estimating **F**

- If we don't know $\mathbf{K}_1$, $\mathbf{K}_2$, $\mathbf{R}$, or $\mathbf{t}$, can we estimate $\mathbf{F}$ for two images?

- Yes, given enough correspondences

Estimating F – 8-point algorithm

- The fundamental matrix F is defined by

$$\mathbf{x}'^{\mathrm{T}}\mathbf{F}\mathbf{x} = 0$$

for any pair of matches x and x' in two images.

- Let x=$(u,v,1)^{\mathrm{T}}$ and x'=$(u',v',1)^{\mathrm{T}}$,

$$\mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

each match gives a linear equation

$$uu' f_{11} + vu' f_{12} + u' f_{13} + uv' f_{21} + vv' f_{22} + v' f_{23} + uf_{31} + vf_{32} + f_{33} = 0$$

8-point algorithm

$$\begin{bmatrix} u_1u_1' & v_1u_1' & u_1' & u_1v_1' & v_1v_1' & v_1' & u_1 & v_1 & 1 \\ u_2u_2' & v_2u_2' & u_2' & u_2v_2' & v_2v_2' & v_2' & u_2 & v_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_nu_n' & v_nu_n' & u_n' & u_nv_n' & v_nv_n' & v_n' & u_n & v_n & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$

- In reality, instead of solving $\mathbf{Af} = 0$, we seek $\mathbf{f}$ to minimize $\|\mathbf{Af}\|$, least eigenvector of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$.

8-point algorithm – Problem?

- **F** should have rank 2
- To enforce that **F** is of rank 2, F is replaced by F' that minimizes $\|\mathbf{F}-\mathbf{F}'\|$ subject to the rank constraint.

- This is achieved by SVD. Let $\mathbf{F}=\mathbf{U}\Sigma\mathbf{V}^{\mathrm{T}}$, where

$$\Sigma=\begin{bmatrix}\sigma_1 & 0 & 0\\ 0 & \sigma_2 & 0\\ 0 & 0 & \sigma_3\end{bmatrix}, \text{ let } \Sigma'=\begin{bmatrix}\sigma_1 & 0 & 0\\ 0 & \sigma_2 & 0\\ 0 & 0 & 0\end{bmatrix}$$

then $\mathbf{F}'=\mathbf{U}\Sigma'\mathbf{V}^{\mathrm{T}}$ is the solution.

Recovering camera parameters from F / E

- Can we recover R and t between the cameras from F?

$$F = K_2^{-T}[\mathbf{t}]_{\times} R K_1^{-1}$$

- No: K_1 and K_2 are in principle arbitrary matrices
- What if we knew K_1 and K_2 to be identity?

$$E = [\mathbf{t}]_{\times} R$$

Recovering camera parameters from E

$$E = [\mathbf{t}]_\times R$$
$$\mathbf{t}^T E = \mathbf{t}^T [\mathbf{t}]_\times R = 0$$
$$E^T \mathbf{t} = 0$$

- **t** is a solution to $E^T\mathbf{x} = 0$
- Can't distinguish between **t** and c**t** for constant scalar c
- How do we recover R?

Recovering camera parameters from E

$$E = [\mathbf{t}]_\times R$$

- We know E and **t**
- Consider taking SVD of E and $[\mathbf{t}]_X$

$$[\mathbf{t}]_\times = U \Sigma V^T$$

$$E = U' \Sigma' V'^T$$

$$U' \Sigma' V'^T = E = [\mathbf{t}]_\times R = U \Sigma V^T R$$

$$U' \Sigma' V'^T = U \Sigma V^T R$$

$$V'^T = V^T R$$

Recovering camera parameters from E

$$E = [\mathbf{t}]_{\times} R$$

$$\mathbf{t}^T E = \mathbf{t}^T [\mathbf{t}]_{\times} R = 0$$

$$E^T \mathbf{t} = 0$$

- **t** is a solution to $E^T\mathbf{x} = 0$
- Can't distinguish between **t** and c**t** for constant scalar c

8-point algorithm

- Pros: it is linear, easy to implement and fast
- Cons: susceptible to noise
- Degenerate: if points are on same plane

- Normalized 8-point algorithm: Hartley
  - Position origin at centroid of image points
  - Rescale coordinates so that center to farthest point is sqrt (2)

Structure-from-motion

- Given 2 (or more) images from *unknown* cameras and with 3D world *unknown*
  - *Can we recover both the cameras and the 3D world structure?*
- Step 1: Get correspondences
- Step 2: Estimate Essential matrix, get R and t (assuming K known)
- Step 3: Use calibrated cameras + correspondence to get 3D locations of points