Lecture 10 – Two view geometry

Nov 20nd, 2018

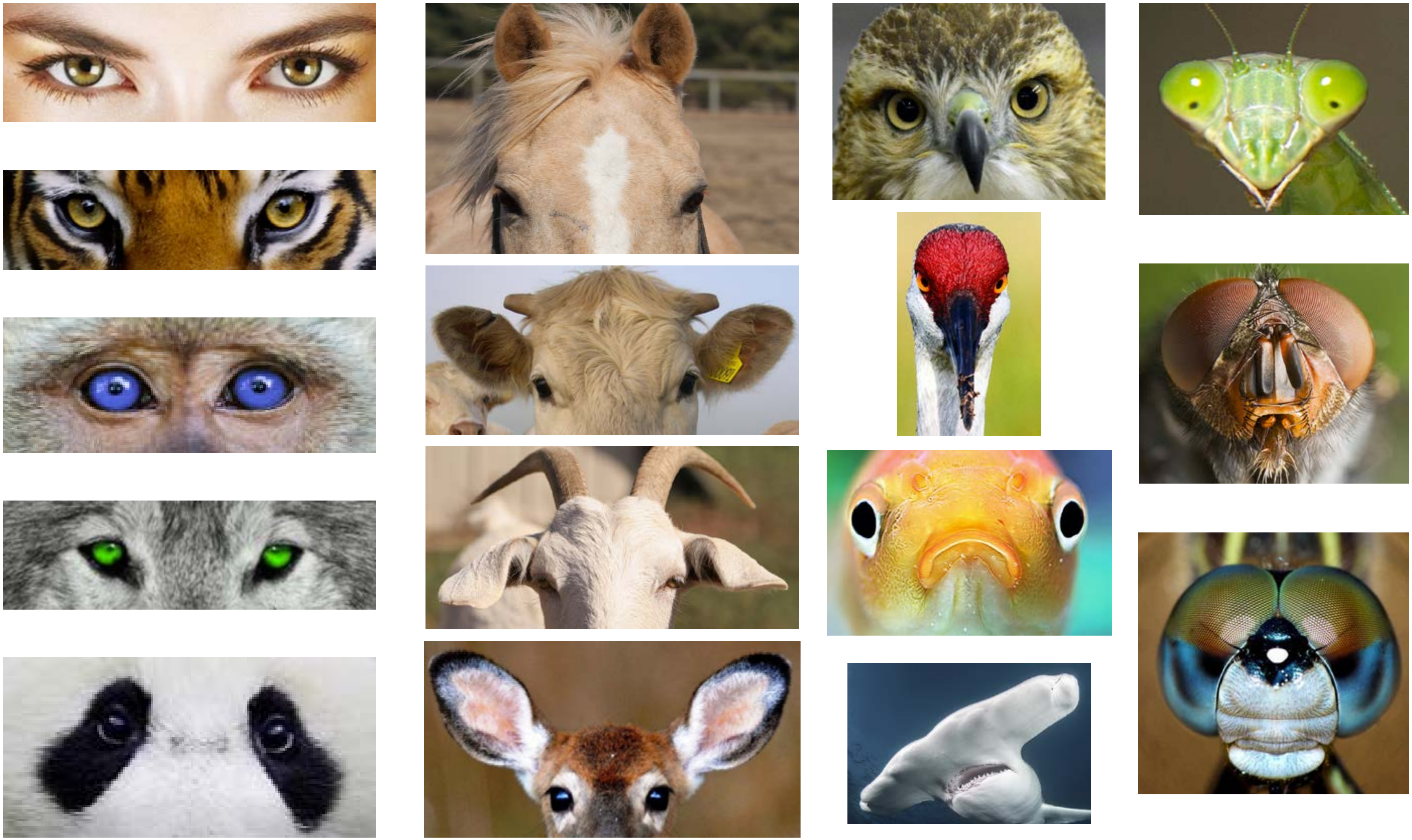
Danping Zou,

Associate Professor

Institute for Sensing and Navigation

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Binocular vision



Binocular vision

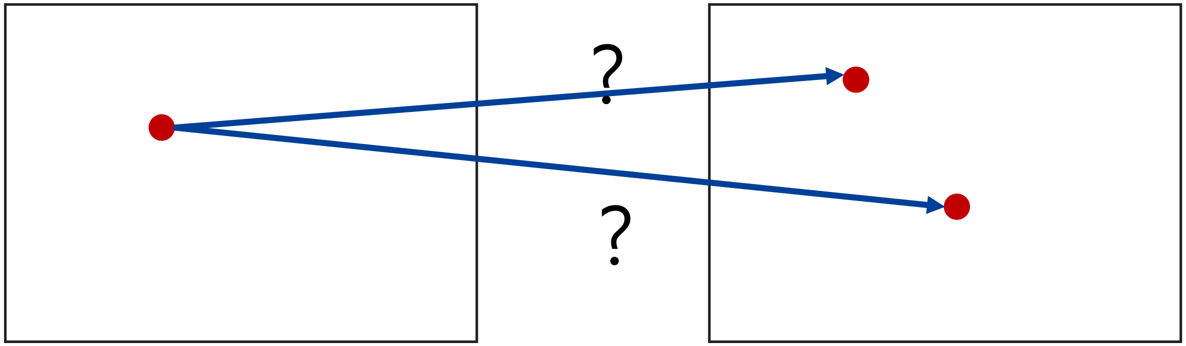
- 3D perception

Two view geometry

3D world

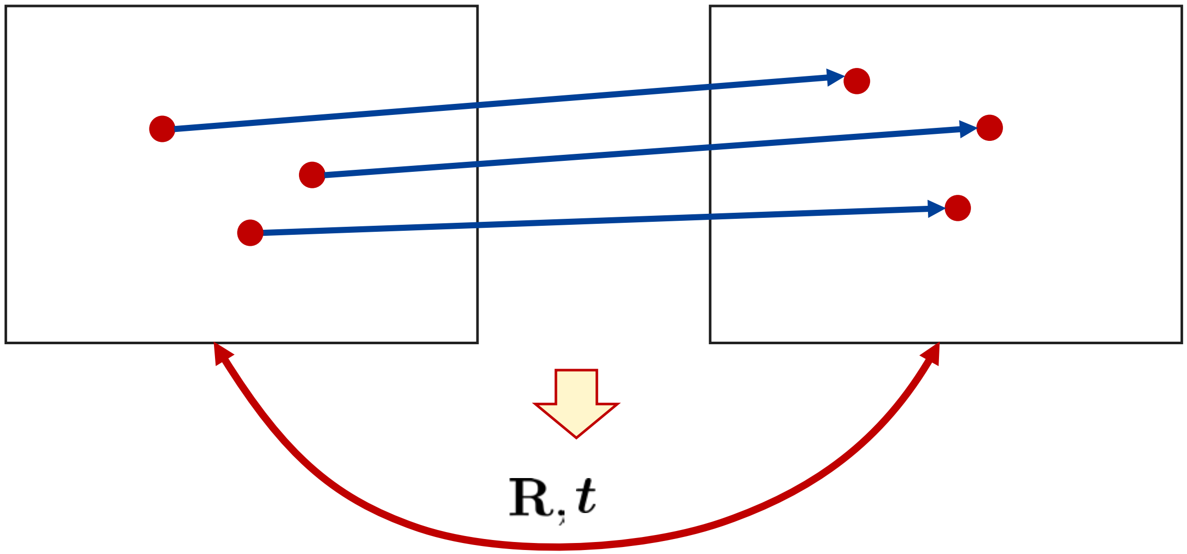
Two view geometry

- Two view geometry tries to answer the following questions.
  - 1. Given a image point in one view, where should its corresponding point be in the other view?
    - **Epipolar constraint (极线约束)**



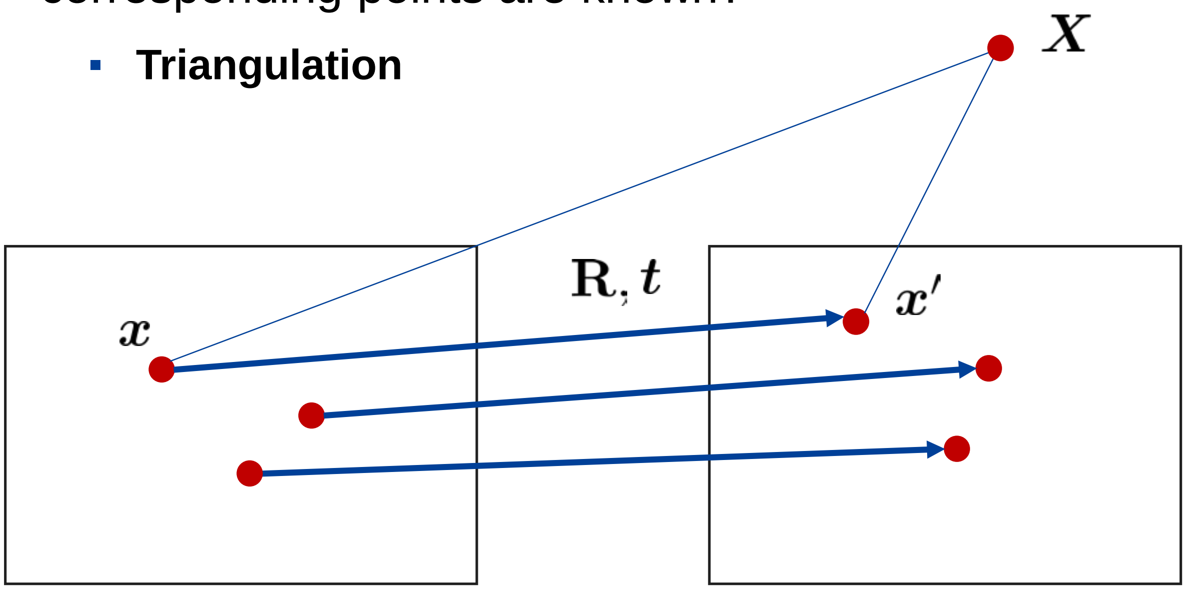
Two view geometry

- Two view geometry tries to answer the following questions.
  - 2. What is the relative pose between two views given a set of correspondences?
    - **Fundamental/Essential matrix estimation**



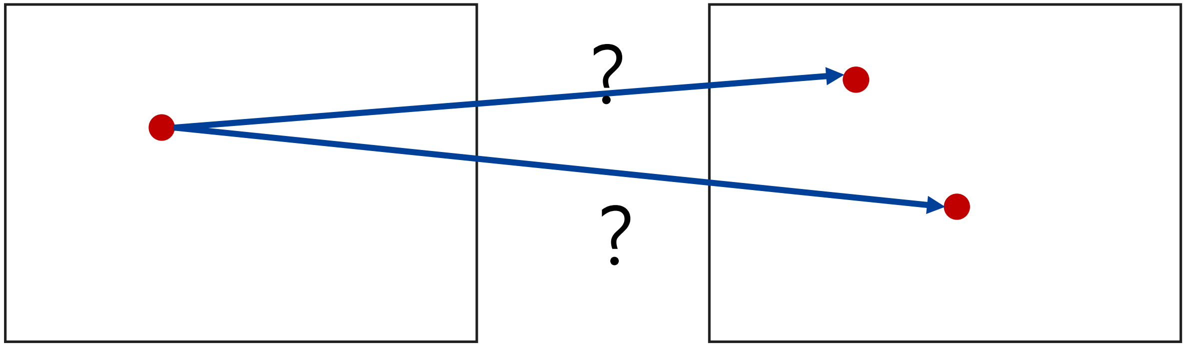
Two view geometry

- Two view geometry tries to answer the following questions.
  - 3. What is the 3D geometry of the scene when the relative pose and corresponding points are known?
    - **Triangulation**



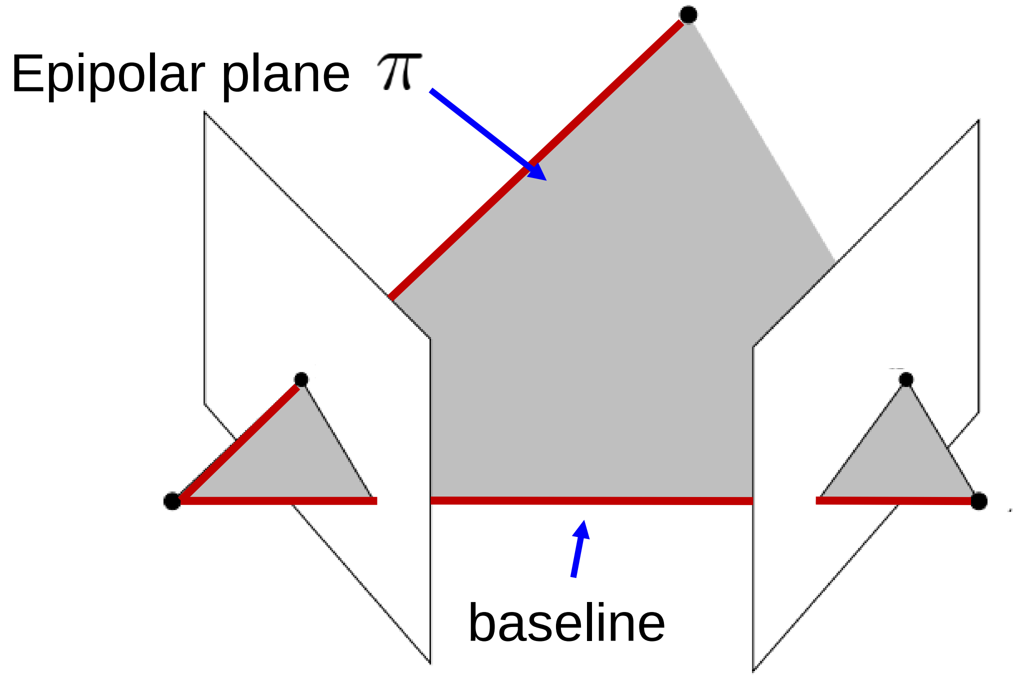
Two view geometry

- **Searching the corresponding point**
  - Given a point in the first view, where can we find its corresponding point in the second view ?



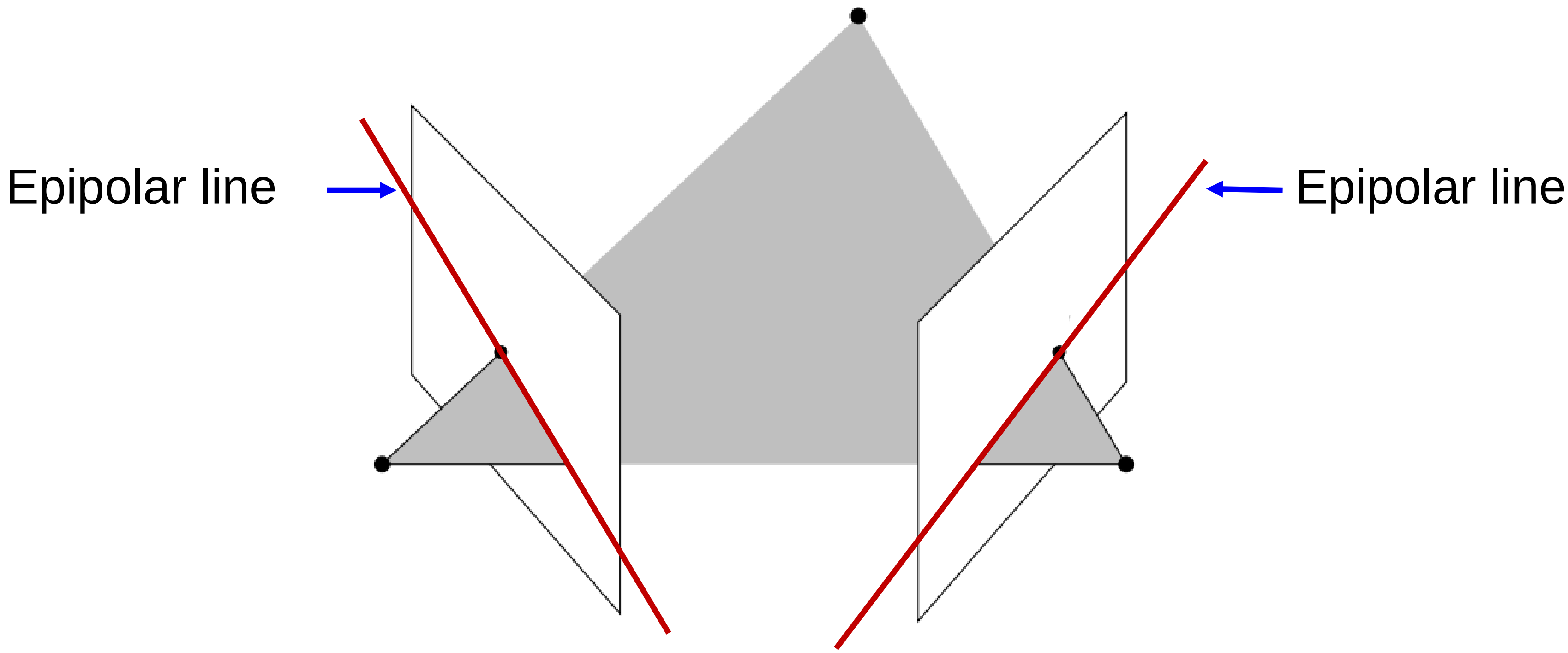
Two view geometry

- **Epipolar plane** – The plane determined by the **baseline** and the ray defined by the image point.



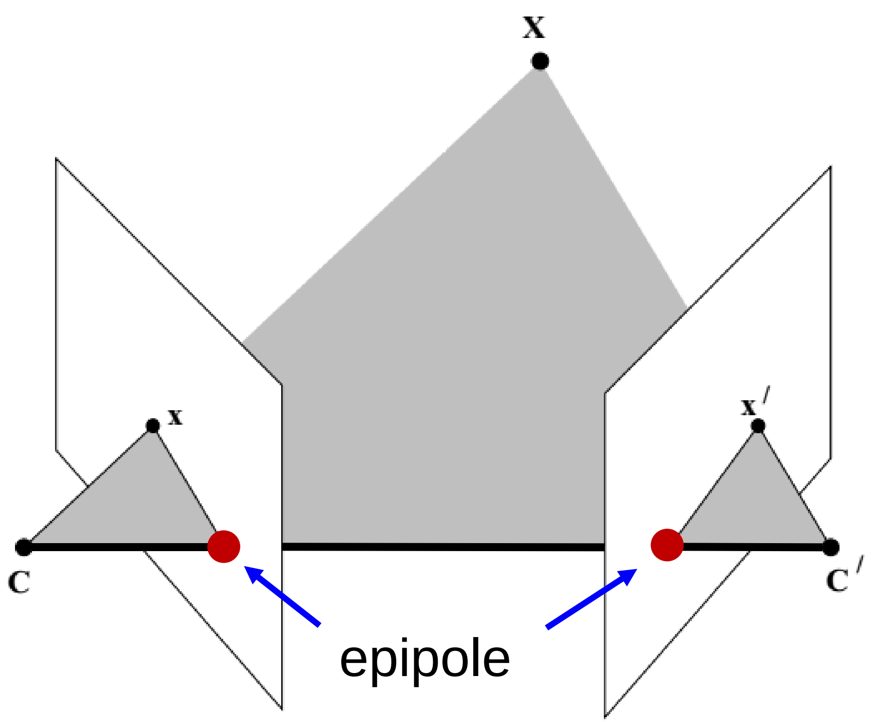
Two view geometry

- **Epipolar line** – Intersection of epipolar plane with the image planes.



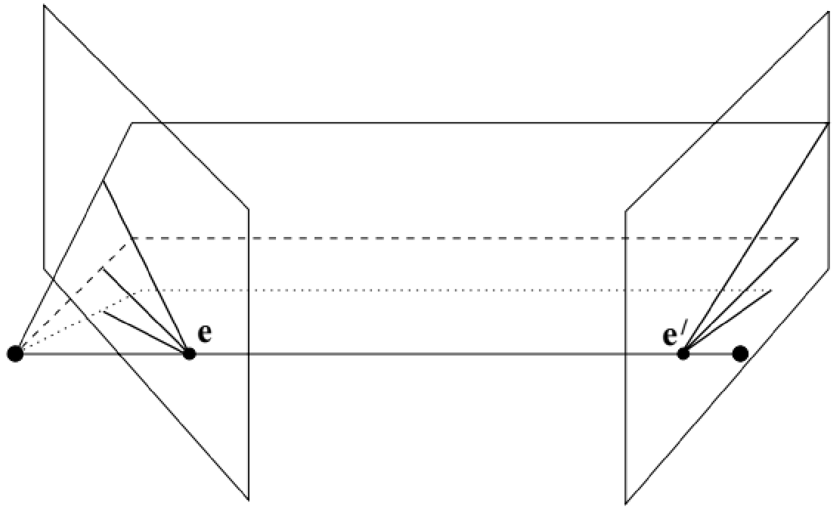
Two view geometry

- **Epipole**
  - intersection of the baseline with the image plane.
  - Projection of the optical center on the other view



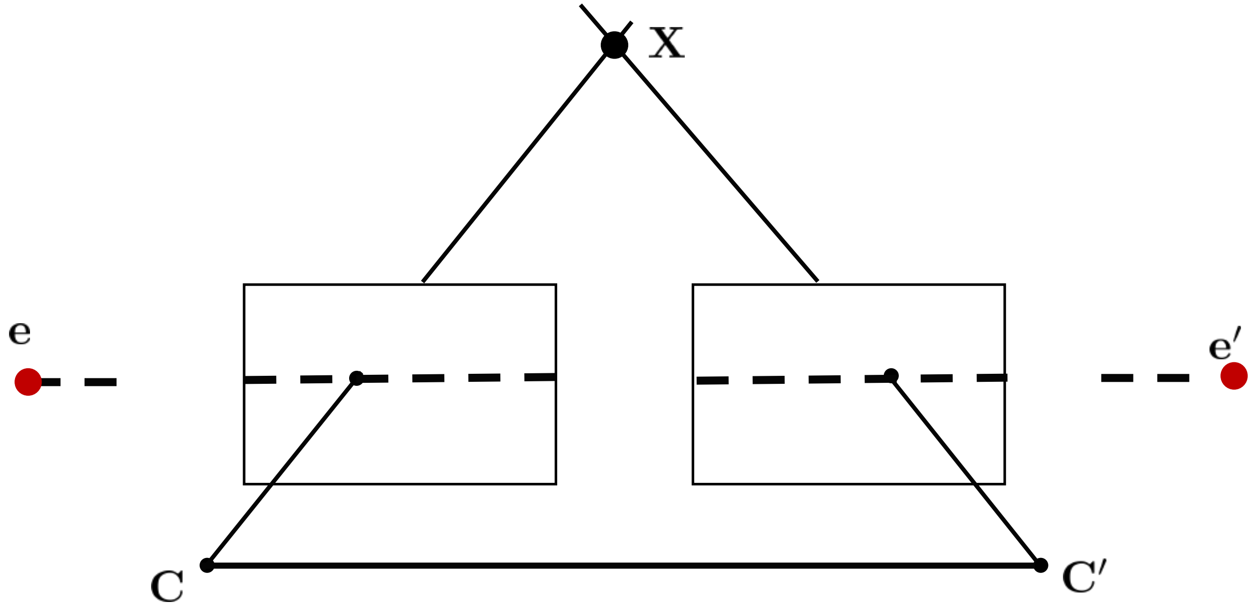
Two view geometry

- Example :



Two view geometry

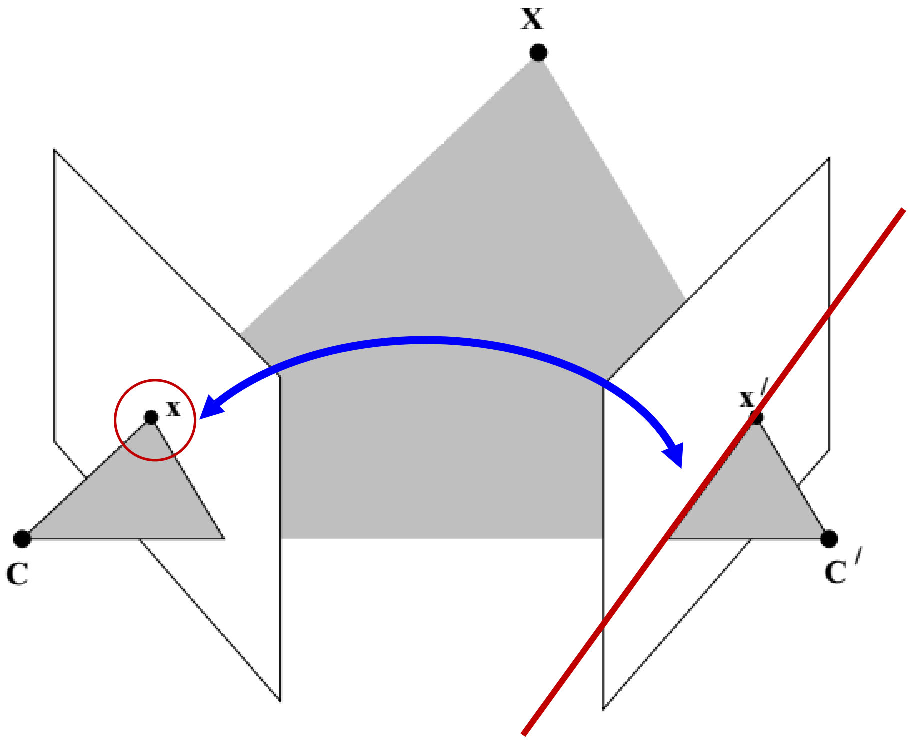
- Example – Parallel image planes



- The baseline intersects with the image plane at infinity
  - Epipoles are at infinity
  - Epipolar lines are collinear

Two view geometry

- **Epipolar constraint:**
  - The corresponding point $\mathbf{x}'$ should lie on the **epipolar line** $\mathbf{l}'$.



Epiploar constraint

- The epipolar constraint is described mathematically as

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

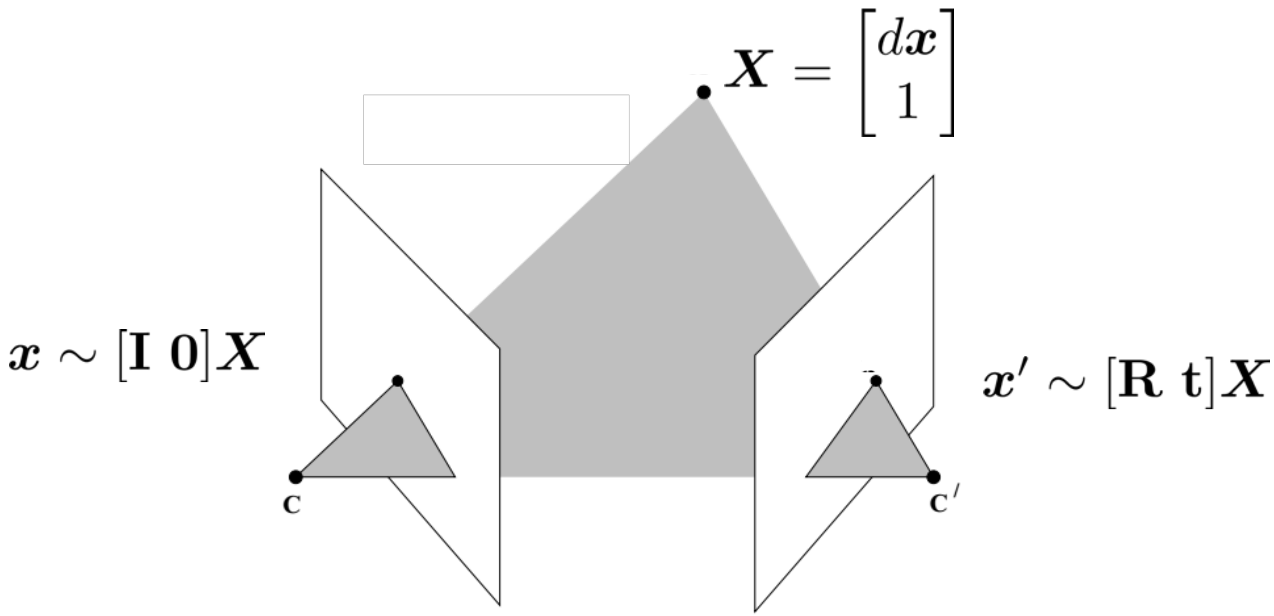
  - Here $\mathbf{F}$ is the **fundamental matrix**
  - $\boldsymbol{l} = \mathbf{F}\boldsymbol{x}$ is the **epipolar line** of $\boldsymbol{x}$

Epiploar constraint

- How do we derive the Epipolar constraint?

Epipolar Constraint

- How do we get the Epipolar constraint?

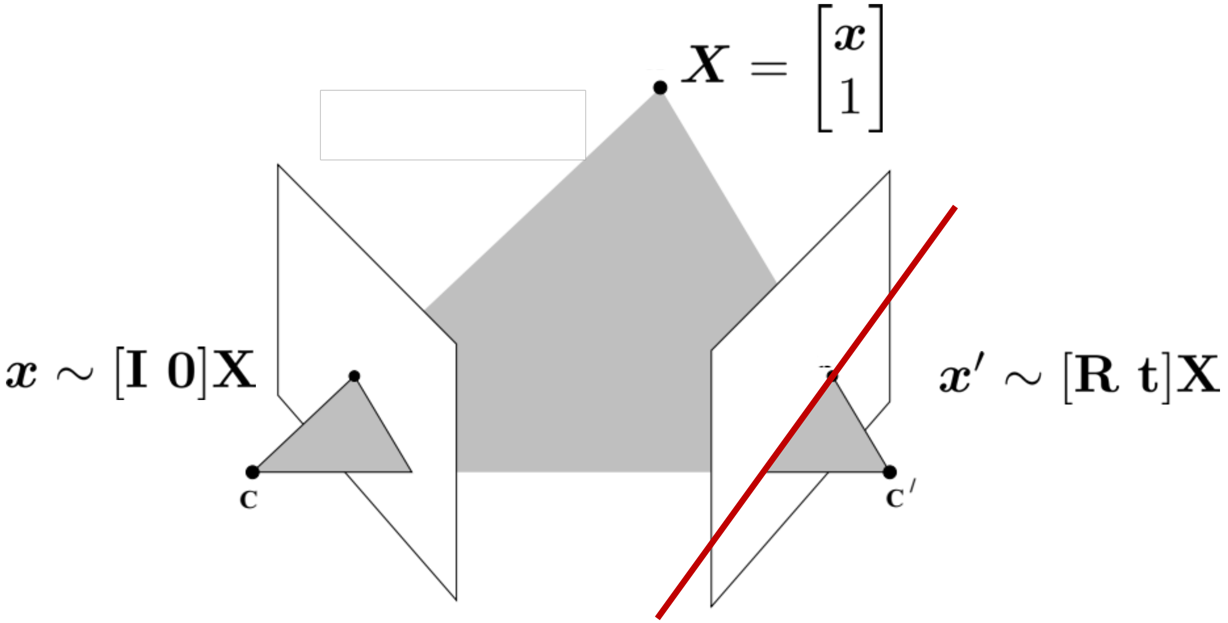


Suppose the camera intrinsic $\mathbf{K}$ is known and the image coordinates have ben normalized.

$$\boldsymbol{x} \sim \mathbf{K}^{-1}\boldsymbol{m}$$

Epipolar Constraint

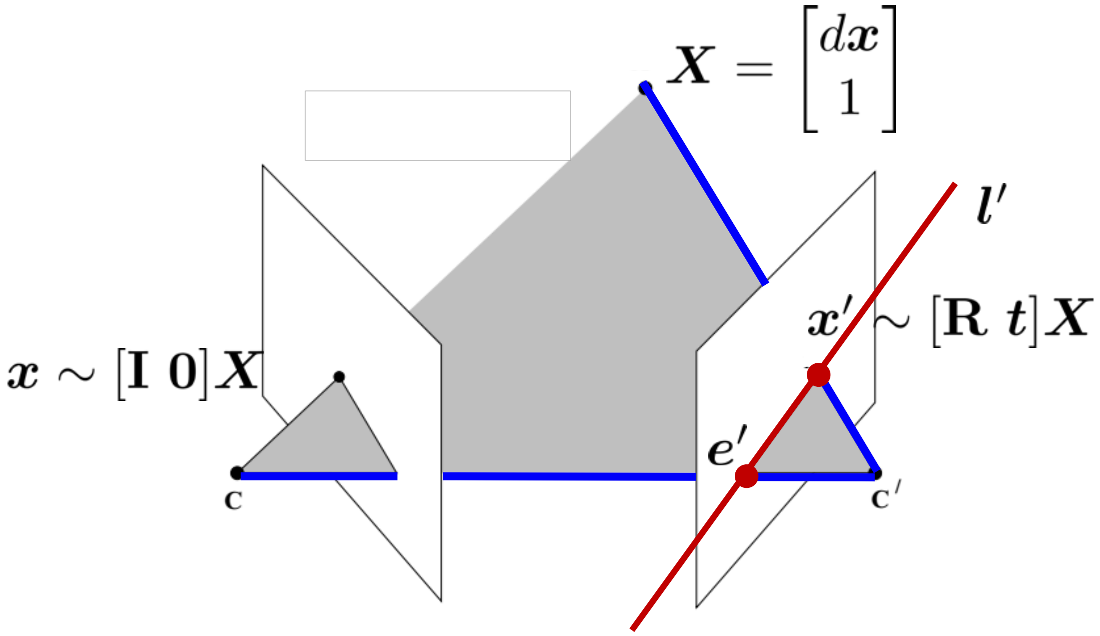
- How do we get the Epipolar constraint?



The key is to get the epipolar line.

Epipolar Constraint

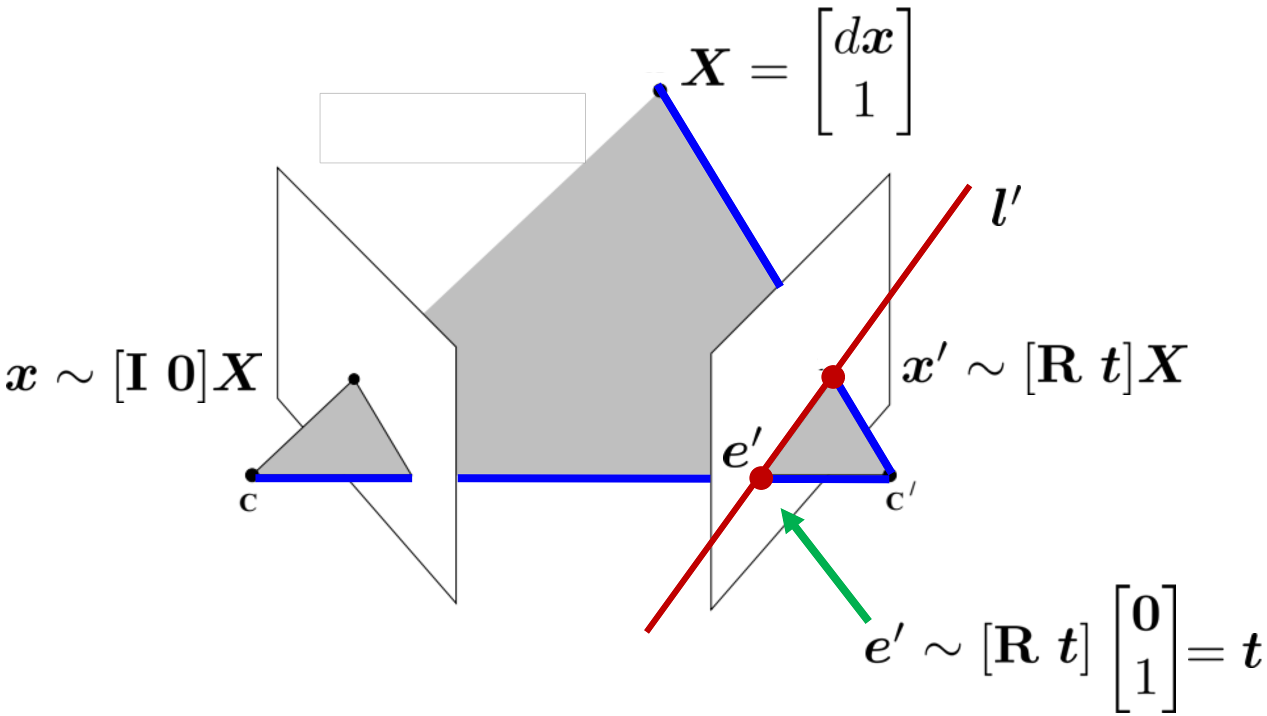
- How do we get the Epipolar constraint?



$$\boldsymbol{l}' \sim \boldsymbol{e}' \times \boldsymbol{x}'$$

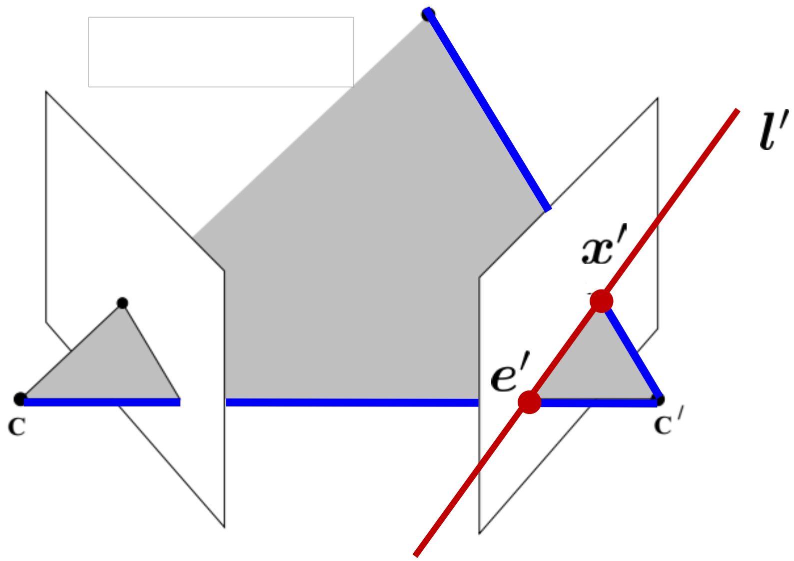
Epipolar Constraint

- How do we get the Epipolar constraint?



Epipolar Constraint

- How do we get the Epipolar constraint?



$$l' \sim e' \times x'$$

$$x' \sim [\mathbf{R}\ t]\mathbf{X} = d\mathbf{R}x + t$$

$$e' \sim [\mathbf{R}\ t]\begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = t$$

$$l' \sim t \times (d\mathbf{R}x + t)$$

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

Epipolar Constraint

- Finally we get the epipolar constraint:

$$\boldsymbol{x}'^{\mathrm{T}}\boldsymbol{l}' = \boldsymbol{x}'^{\mathrm{T}}[\boldsymbol{t}]_{\times}\mathbf{R}\boldsymbol{x} = \mathbf{0}$$

$$\boldsymbol{x}'^{\mathrm{T}}\boxed{\mathbf{E}}\boldsymbol{x} = \mathbf{0}$$

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

is the **essential matrix**

Epipolar constraint

- Take the camera intrinsic parameters into consideration

$$\boldsymbol{x} \sim \mathbf{K}^{-1}\boldsymbol{m}$$

$$\boldsymbol{x}' \sim \mathbf{K}'^{-\mathrm{T}}\boldsymbol{m}'$$

$$\Downarrow$$

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

$$\Downarrow$$

$$\boldsymbol{m}'^{\mathrm{T}}\boxed{\mathbf{K}'^{-\mathrm{T}}\mathbf{E}\mathbf{K}^{-1}}\boldsymbol{m} = \mathbf{0}$$

$$\Downarrow$$

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

$$\mathbf{F} = \mathbf{K}'^{-\mathrm{T}}\mathbf{E}\mathbf{K}^{-1}$$ **Fundamental matrix**

Essential matrix

- Properties of the essential matrix:

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

A 3×3 matrix is an essential matrix if and only if two of its singular values are equal, and the third is zero.

$$\mathbf{E} = \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}^T$$

Essential matrix

- A 3×3 skew-symmetric matrix can be decomposed as

$$[\mathbf{t}]_\times \sim \lambda \mathbf{U} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{U}^{\mathrm{T}}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{W}}$$

$$\mathbf{E} \sim [\mathbf{t}]_\times \mathbf{R} \sim \mathbf{U} \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{W}\,\mathbf{U}^{\mathrm{T}}\mathbf{R} \sim \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}^{\mathrm{T}}$$

Essential matrix

- How do we decompose $\mathbf{t}, \mathbf{R}$ from the essential matrix $\mathbf{E}$?

$$\mathbf{E} \sim [\mathbf{t}]_\times \mathbf{R} \sim \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \boxed{\lambda \mathbf{W} \mathbf{U}^\mathrm{T} \mathbf{R}}$$

$$\mathbf{E} = \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \boxed{\mathbf{V}^\mathrm{T}}$$

1. $[\mathbf{t}]_\times \sim \mathbf{U} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{U}^\mathrm{T}$

2. $\mathbf{R} = \mathbf{U}\mathbf{W}\mathbf{V}^T$ or $\mathbf{U}\mathbf{W}^T\mathbf{V}^T$

Essential matrix

- Extract R and t from essential matrix by SVD

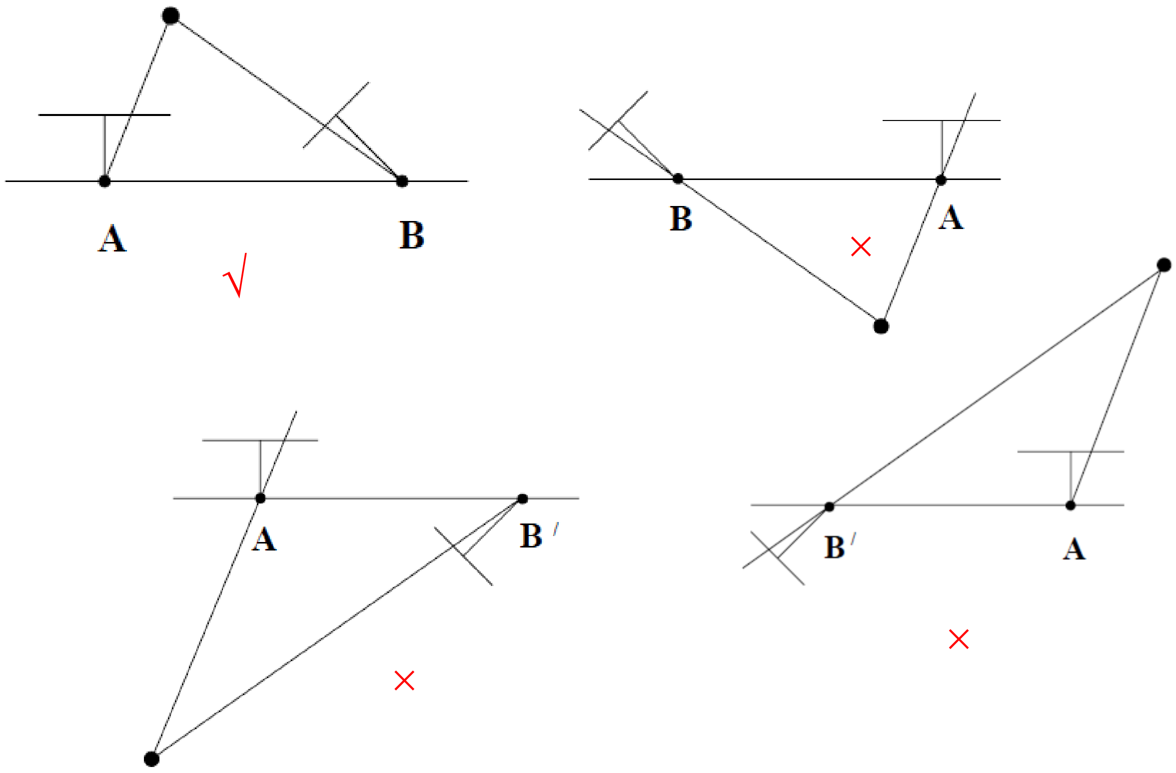
$$\mathbf{E} = \mathbf{U}diag(1,1,0)\mathbf{V}^T$$

$$[\mathbf{t}]_\times = \mathbf{U}\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\mathbf{U}^\mathrm{T} \quad \text{or} \quad [\mathbf{t}]_\times = -\mathbf{U}\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\mathbf{U}^\mathrm{T}$$

$$\mathbf{R} = \mathbf{U}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\mathbf{V}^T \quad \text{or} \quad \mathbf{R} = \mathbf{U}\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\mathbf{V}^T$$

Essential matrix

- Four possible solutions – we can select the right camera poses that triangulated 3D points are in front of the cameras



Epipolar geometry

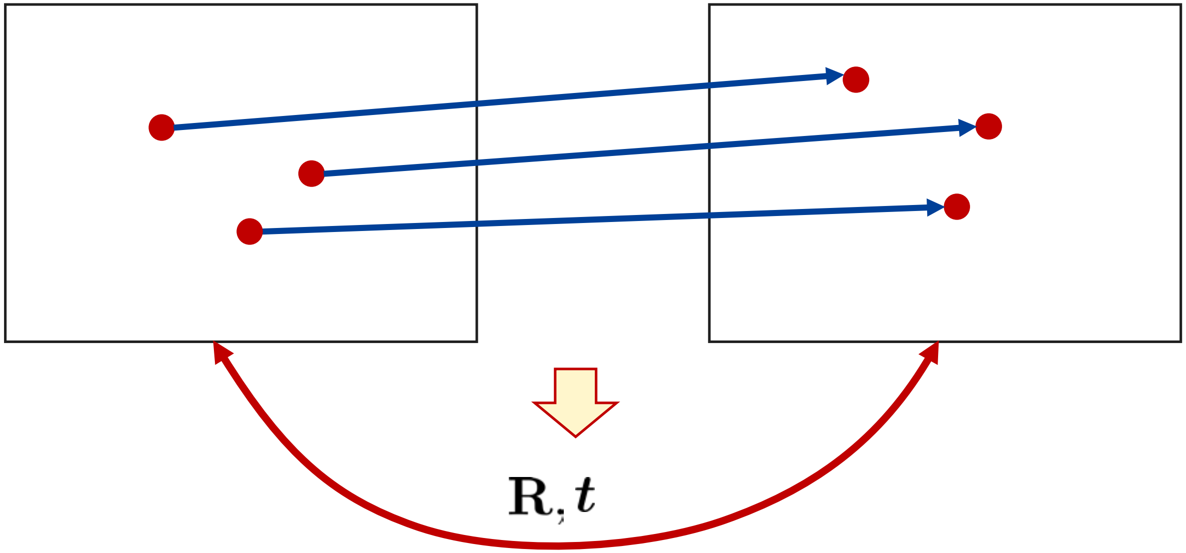
- Why is the essential matrix useful?
  - It captures information about the epipolar geometry of 2 views + camera parameters
  - It encodes the relative poses between the two views (given that the camera intrinsic are known)
  - For 3D reconstruction (Triangulation)
- Fundamental matrix <-> essential matrix

Summary

- Epipolar plane / Epipolar line / Epipole
- Epipolar constraint
- Essential matrix / Fundamental matrix
- Extract the relative pose $\mathbf{R}, \boldsymbol{t}$ from the essential matrix

Two view geometry

- Two view geometry tries to answer the following questions.
  - 2. What is the relative pose between two views given a set of correspondences?
    - **Fundamental/Essential matrix estimation**



Fundamental/Essential matrix estimation

- **Eight point algorithm**
  - For each correspondence $\mathbf{x} \leftrightarrow \mathbf{x}'$, we have the equation :

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

  - Let $\mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$
  - The epipolar constraint can be written as

$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'f_{23} + xf_{31} + yf_{32} + f_{33} = 0$$

Fundamental/Essential matrix estimation

- We can further write it in matrix-vector production:

$$(x'x, x'y, x', y'x, y'y, y', x, y, 1)\begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = 0$$

Fundamental/Essential matrix estimation

- If we have n point correspondences, we have

$$\mathbf{Af} = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix} \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = \mathbf{0}$$

Two view geometry

- Since is $\mathbf{F}$ determined up to scale only, at least eight points are required to solve $\mathbf{F}$.
- Homogenous system again
- The solution can also be obtained by SVD decomposition (The vector corresponds to the minimum singular value)

$$\mathbf{A} = \mathbf{U}\mathrm{diag}(\sigma_1, \sigma_2 \ldots \boxed{\sigma_9}) \left[\mathbf{v}_1, \mathbf{v}_2 \ldots \boxed{\mathbf{v}_9}\right]^{\mathrm{T}} \qquad (\sigma_1 > \sigma_2 \ldots \sigma_9)$$

Constraint enforcement

- Singularity correction
  - The solution by 8 point algorithm does not satisfy the singularity condition.
    - $\mathbf{F}$ is rank-2 matrix or mathematically, $\det(\mathbf{F}) = 0$
  - SVD approximation
    - Decompose $\mathbf{F}$ by SVD : $\mathbf{F} = \mathbf{U}\Sigma\mathbf{V}^T$
    - Here $\Sigma = diag(r, s, t)$
    - The SVD approximation of $\mathbf{F}$ is $\mathbf{F}' = \mathbf{U}diag(r, s, 0)\mathbf{V}^T$

$\mathbf{F}'$ is the 'closest' singular matrix to $\mathbf{F}$ in Frobenius norm.

$\mathbf{E}' = \mathbf{U}diag(1,1,0)\mathbf{V}^T$ (for essential matrix)

Essential matrix estimation

- Compute essential matrix
  - Once the camera has been calibrated.
  - Only five points are required to solve essential matrix since there is only five degree of freedom in essential matrix

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

  - Nister's **five point algorithm**

    Nistér, David. "An efficient solution to the five-point relative pose problem."*Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26.6 (2004): 756-770.

Essential matrix estimation

- Nister's **five point algorithm** :
  - Five points give five equations: $\mathbf{B}\tilde{\mathbf{E}} = 0$
  - B is 5x9 matrix. We can get the null space solution

$$\mathbf{E} = x\mathbf{X} + y\mathbf{Y} + z\mathbf{Z} + w\mathbf{W}$$

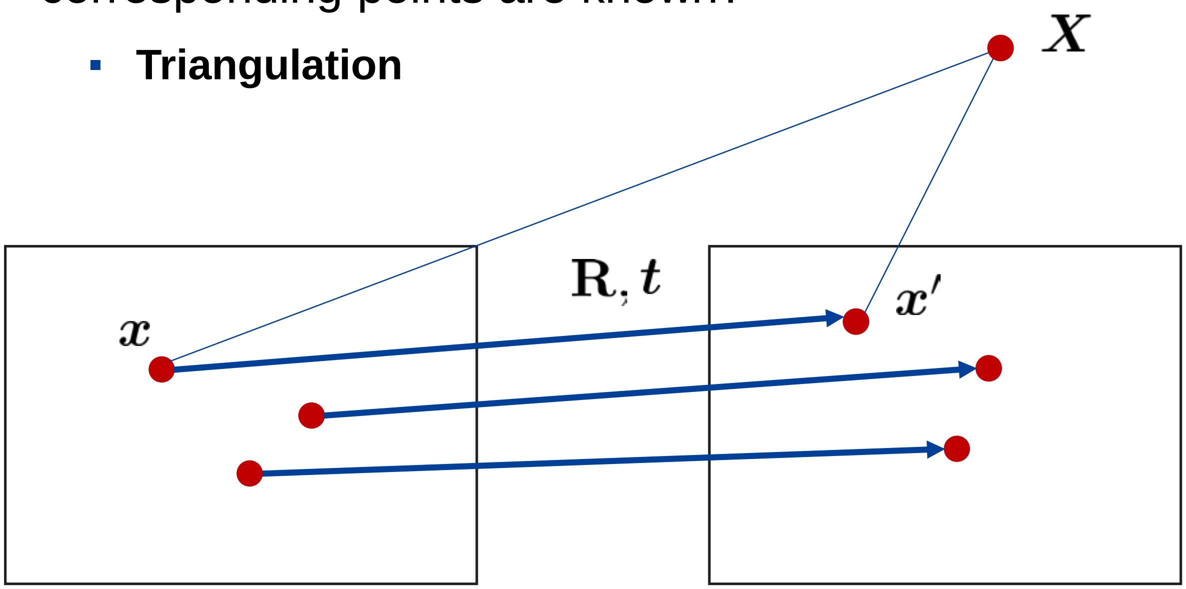
  - Let $w = 1$ to fix the scaling. Solve (x, y, z) using the following constraints.

$$\mathbf{E}\mathbf{E}^{\mathrm{T}}\mathbf{E} - \frac{1}{2}trace(\mathbf{E}\mathbf{E}^{\mathrm{T}})\mathbf{E} = 0$$

This is a system of polynomial equations. Recall the ***Sylvester Resultant*** introduced in Lecture 09.

Two view geometry

- Two view geometry tries to answer the following questions.
  - 3. What is the 3D geometry of the scene when the relative pose and corresponding points are known?
    - **Triangulation**



3D Reconstruction

- After we extract the relative pose between two views

$$\mathbf{E} \rightarrow (\boldsymbol{t}, \mathbf{R})$$

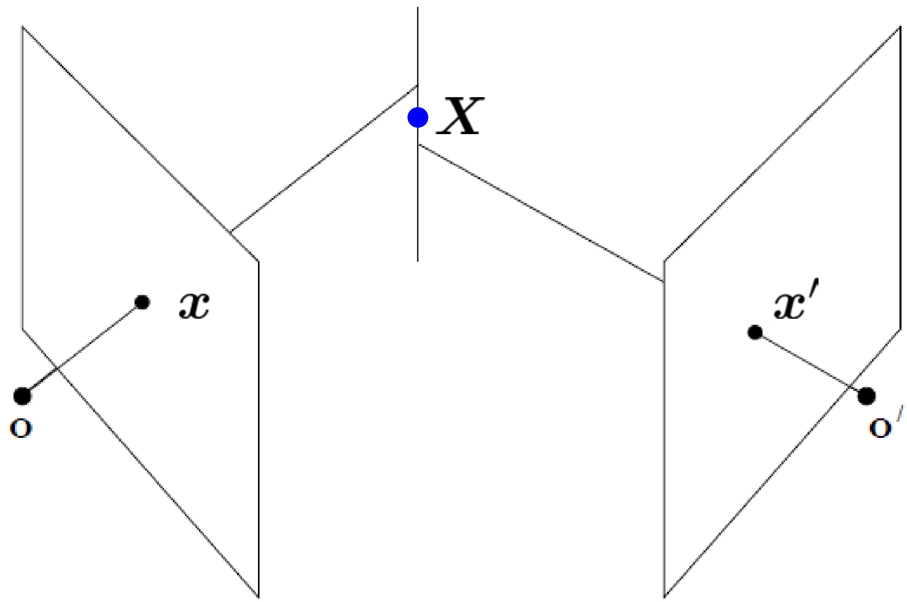
- We can get the camera matrices for each view:

$$\mathbf{P} = \mathbf{K}\,[\mathbf{I}\ \mathbf{0}] \quad \mathbf{P} = \mathbf{K}[\mathbf{R}\ \boldsymbol{t}]$$

- Given a pair of corresponding points , $\boldsymbol{x} \leftrightarrow \boldsymbol{x}'$, how do we compute its 3D coordinates?

Triangulation

- Knowing $\boldsymbol{x}$ and $\boldsymbol{x}'$
- Knowing $\mathbf{P}$ and $\mathbf{P}'$
- Compute $\boldsymbol{X}$



Triangulation

- Using homogenous coordinates

$$\boldsymbol{x} \sim \mathbf{P}\boldsymbol{X} \qquad \boldsymbol{x}' \sim \mathbf{P}\boldsymbol{X}'$$

$$\Downarrow$$

$$\boldsymbol{x} \times \mathbf{P}\boldsymbol{X} = \mathbf{0}$$

$$\boldsymbol{x}' \times \mathbf{P}'\boldsymbol{X} = \mathbf{0}$$

$$\Downarrow$$

$$\begin{bmatrix} [\boldsymbol{x}]_\times \mathbf{P} \\ [\boldsymbol{x}']_\times \mathbf{P}' \end{bmatrix} \boldsymbol{X} = \mathbf{0}$$

$$\Downarrow$$

$$\mathbf{A}_{6\times 4}\boldsymbol{X} = \mathbf{0}_{6\times 1}$$

Triangulation

- Using Cartesian coordinates

$$u = \frac{\mathbf{P}_1^{\mathrm{T}} \boldsymbol{X}}{\mathbf{P}_3^{\mathrm{T}} \boldsymbol{X}} \qquad v = \frac{\mathbf{P}_2^{\mathrm{T}} \boldsymbol{X}}{\mathbf{P}_3^{\mathrm{T}} \boldsymbol{X}}$$

$$\Rightarrow$$

$$p_1 x + p_2 y + p_3 z + p_4 = u(p_9 x + p_{10} y + p_{11} z + p_{12})$$

$$p_5 x + p_6 y + p_7 z + p_8 = v(p_9 x + p_{10} y + p_{11} z + p_{12})$$

$$\Downarrow$$

$$u' = \frac{\mathbf{P}_1'^{\mathrm{T}} \boldsymbol{X}}{\mathbf{P}_3'^{\mathrm{T}} \boldsymbol{X}} \qquad v' = \frac{\mathbf{P}_2'^{\mathrm{T}} \boldsymbol{X}}{\mathbf{P}_3'^{\mathrm{T}} \boldsymbol{X}} \quad \Rightarrow$$

$$\boxed{\begin{aligned}
&(p_1 - u p_9)x + (p_2 - u p_{10})y + (p_3 - u p_{11})z = u p_{12} - p_4 \\
&(p_5 - v p_9)x + (p_6 - v p_{10})y + (p_7 - v p_{11})z = v p_{12} - p_8 \\
&(p_1' - u' p_9')x + (p_2' - u' p_{10}')y + (p_3' - u' p_{11}')z = u' p_{12}' - p_4' \\
&(p_5' - v' p_9')x + (p_6' - v' p_{10}')y + (p_7' - v' p_{11}')z = v' p_{12}' - p_8'
\end{aligned}}$$

$$\Downarrow$$

$$\mathbf{A}_{4\times 3} \boldsymbol{x} = \boldsymbol{b}_{4\times 1}$$

Triangulation

- Refinement
  - Minimizing the re-projection errors

$$\|\mathbf{x} - f(\mathbf{P}, \mathbf{X})\|^2 + \|\mathbf{x}' - f(\mathbf{P}', \mathbf{X})\|^2$$

Here $f(\mathbf{P}, \mathbf{X}) = \begin{pmatrix} \frac{p_1 x + p_2 y + p_3 z + p_4}{p_9 x + p_{10} y + p_{11} z + p_{12}} \\ \frac{p_5 x + p_6 y + p_7 z + p_8}{p_9 x + p_{10} y + p_{11} z + p_{12}} \end{pmatrix}$ It is a nonlinear least square problem and can be solved by Gauss-Newton or Levenberg-Marquardt algorithm efficiently.

Summary