3D Data Processing

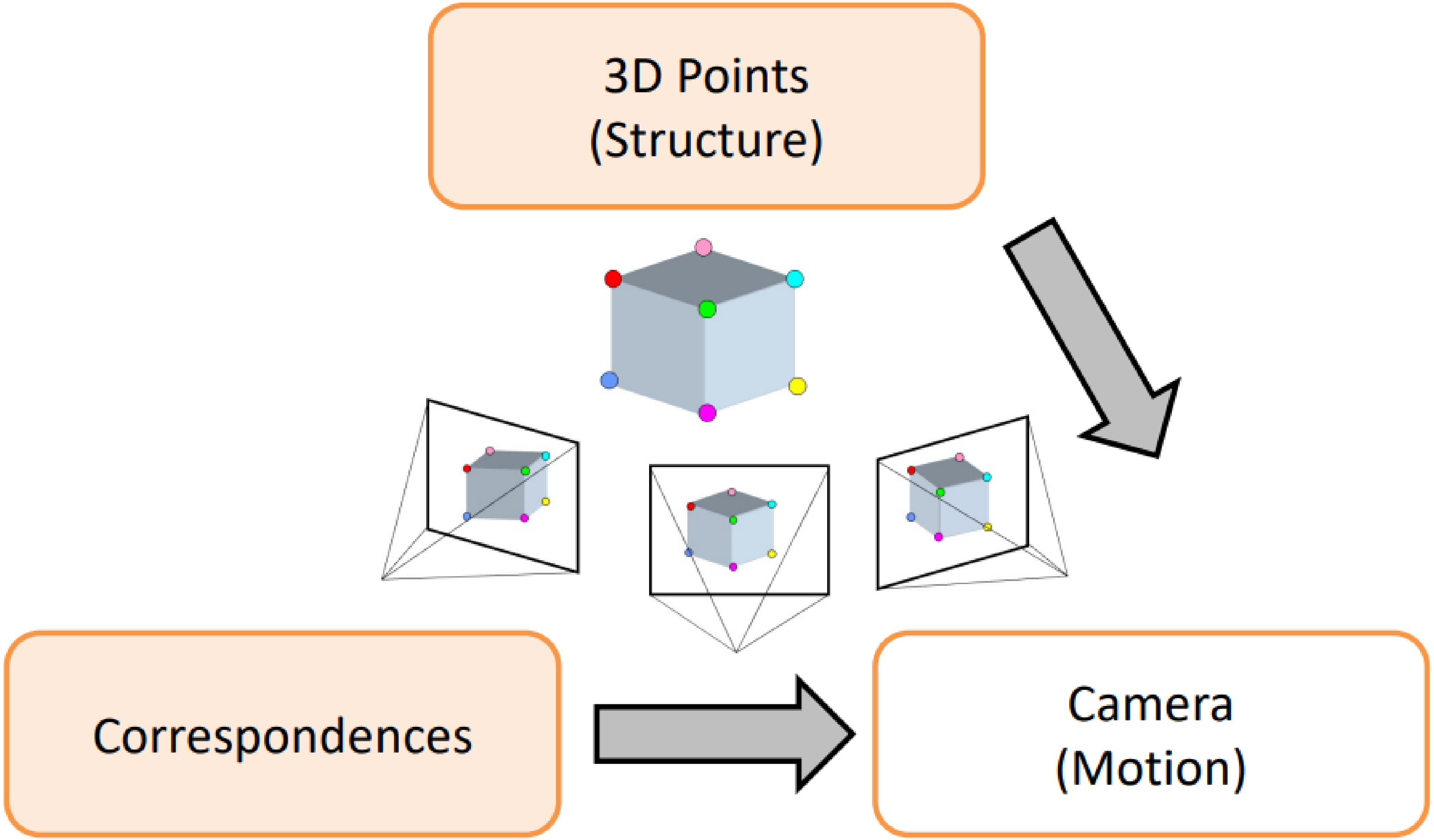
3D Reconstruction From RGB(2D) images

Hyoseok Hwang

Today

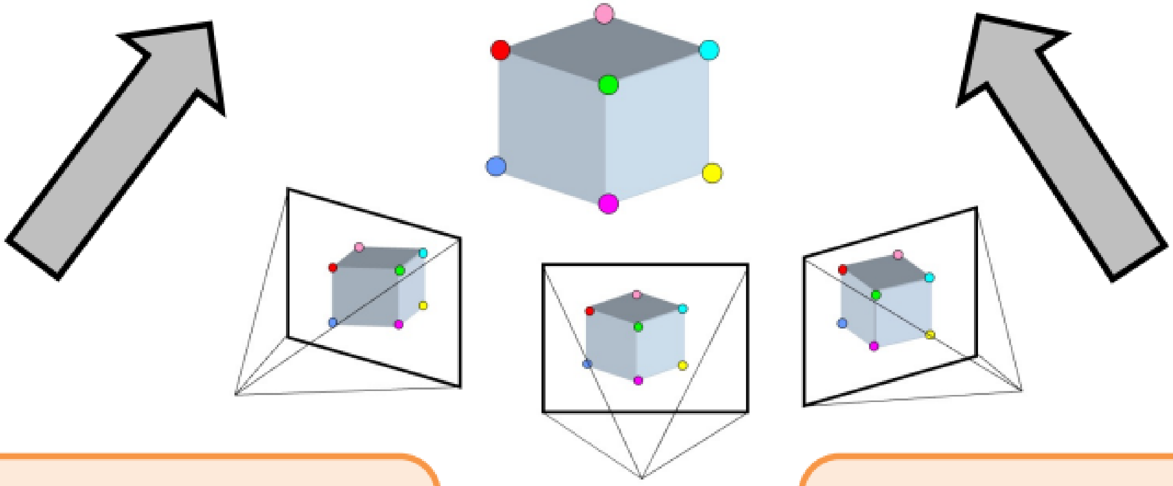
- Epipolar Geometry Review
- Pose estimation
  - Get Fundamental matrix
  - Get Essential matrix
  - Compute R, T

3 key components in 3D



3 key components in 3D

3D Points
(Structure)

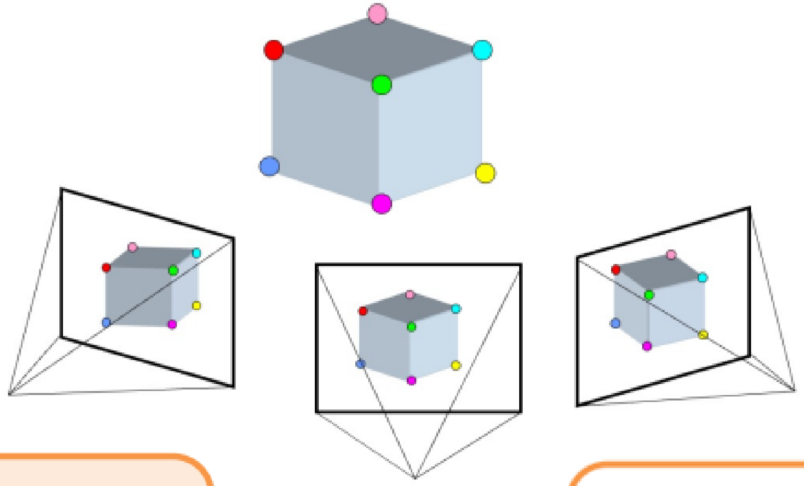


Correspondences

Camera
(Motion)

3 key components in 3D

3D Points
(Structure)

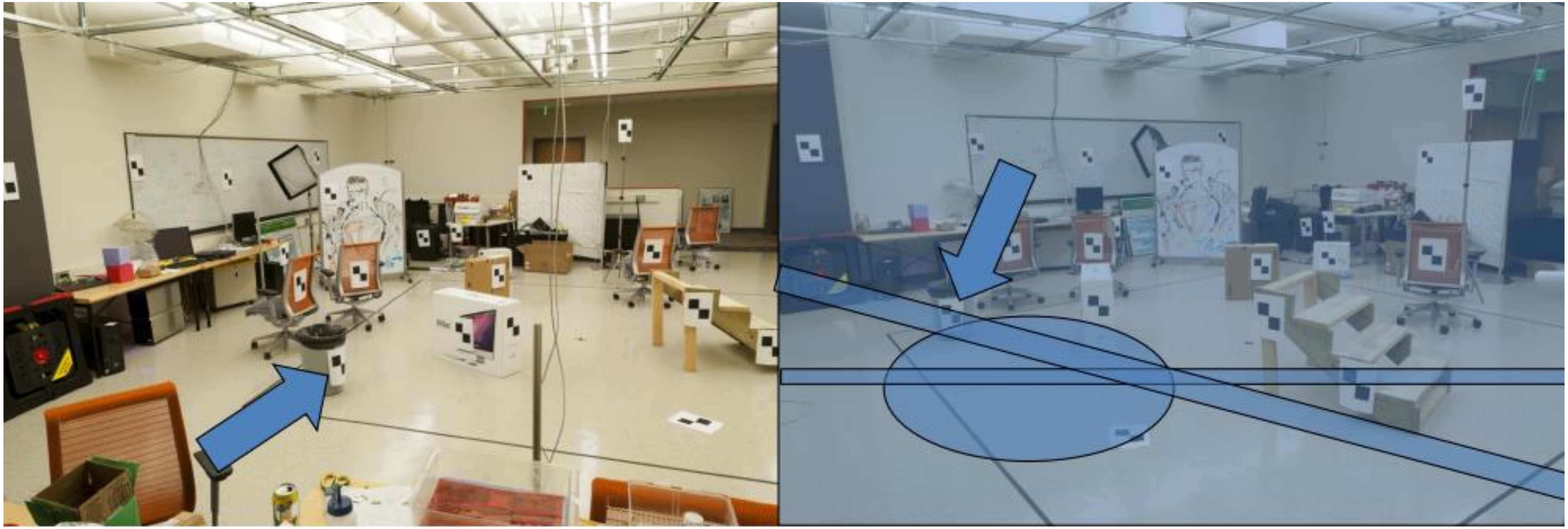


Correspondences

Camera
(Motion)

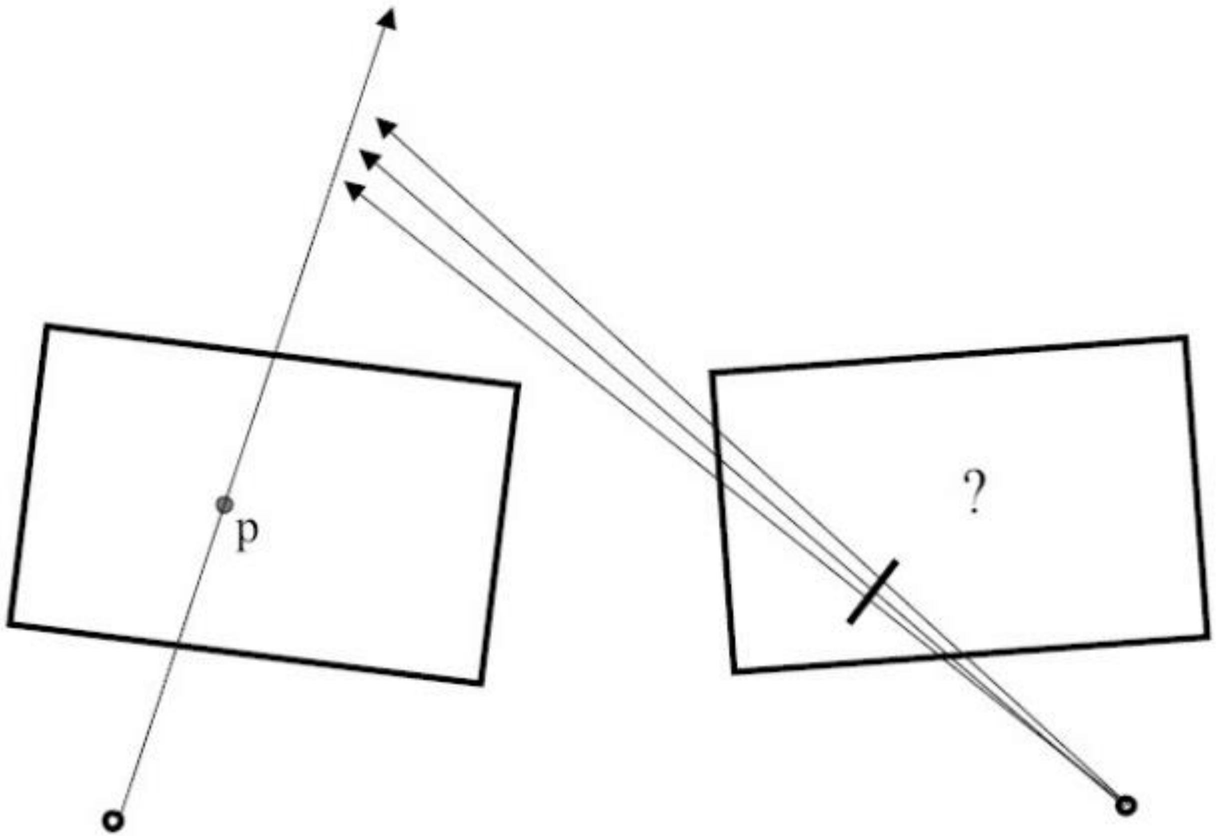
Stereo matching

- Where can I found corresponding point?



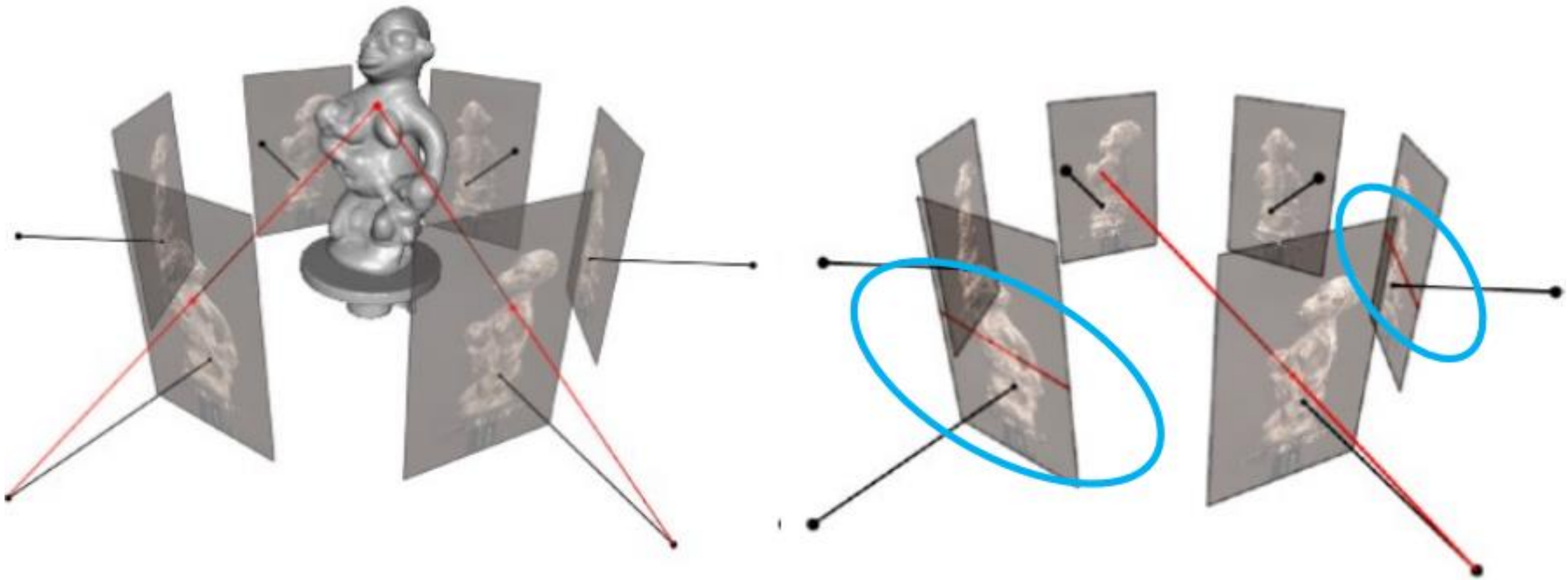
Stereo matching

- Stereo correspondence constraints



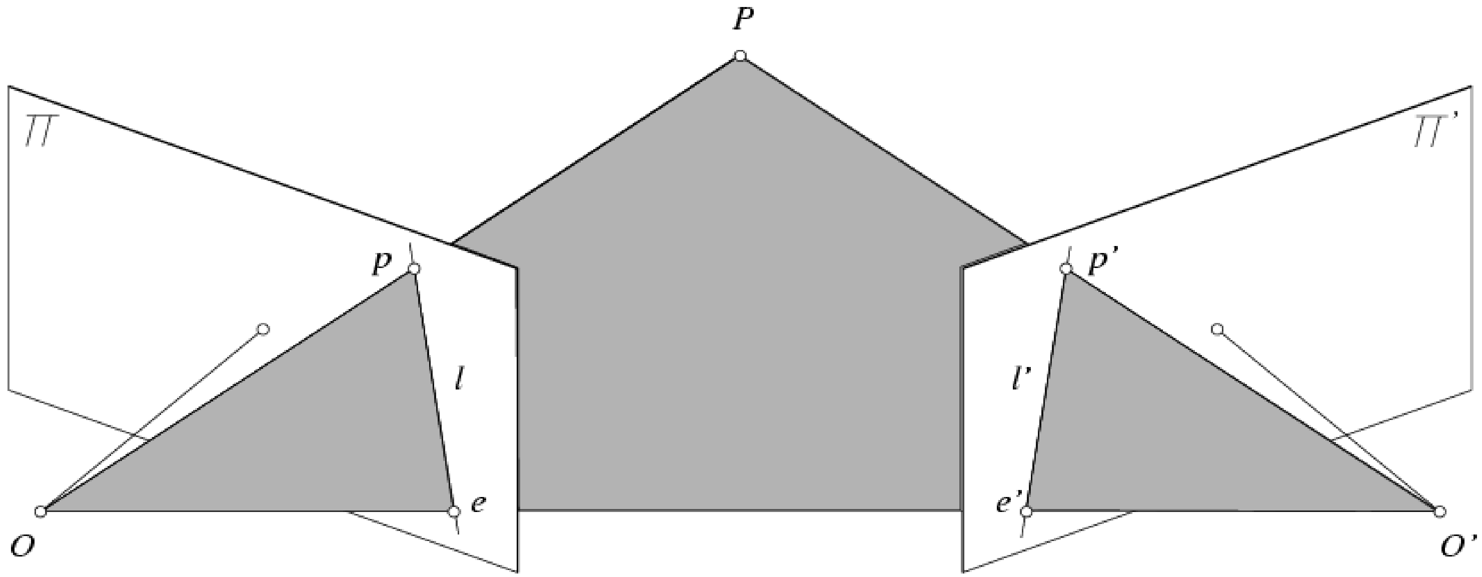
Multi-view case

- Stereo correspondence constraints



Epipolar constraint: Calibrated case

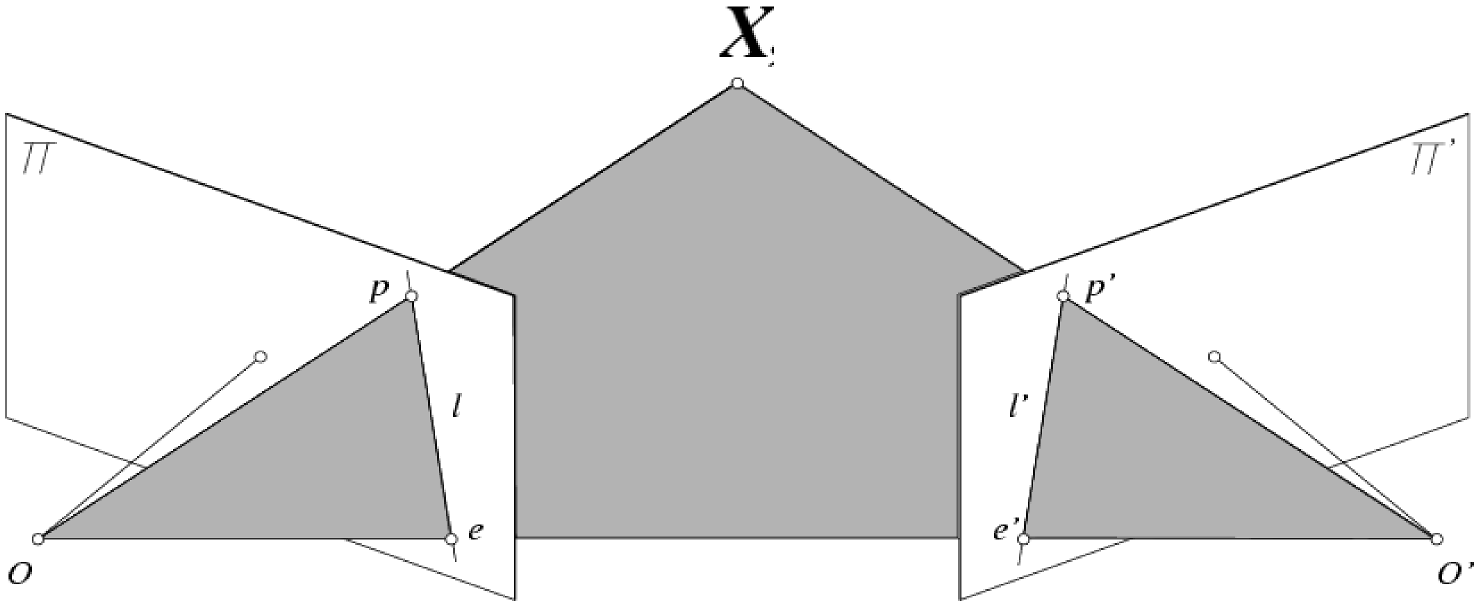
- Intrinsic and extrinsic parameters of the cameras are known, world coordinate system is set to that of the first camera
- Then the projection matrices are given by $\boldsymbol{K}[\boldsymbol{I} \mid \boldsymbol{0}]$ and $\boldsymbol{K}'[\boldsymbol{R} \mid \boldsymbol{t}]$
- We can multiply the projection matrices (and the image points) by the inverse of the calibration matrices to get *normalized* image coordinates:



Epipolar constraint: Calibrated case

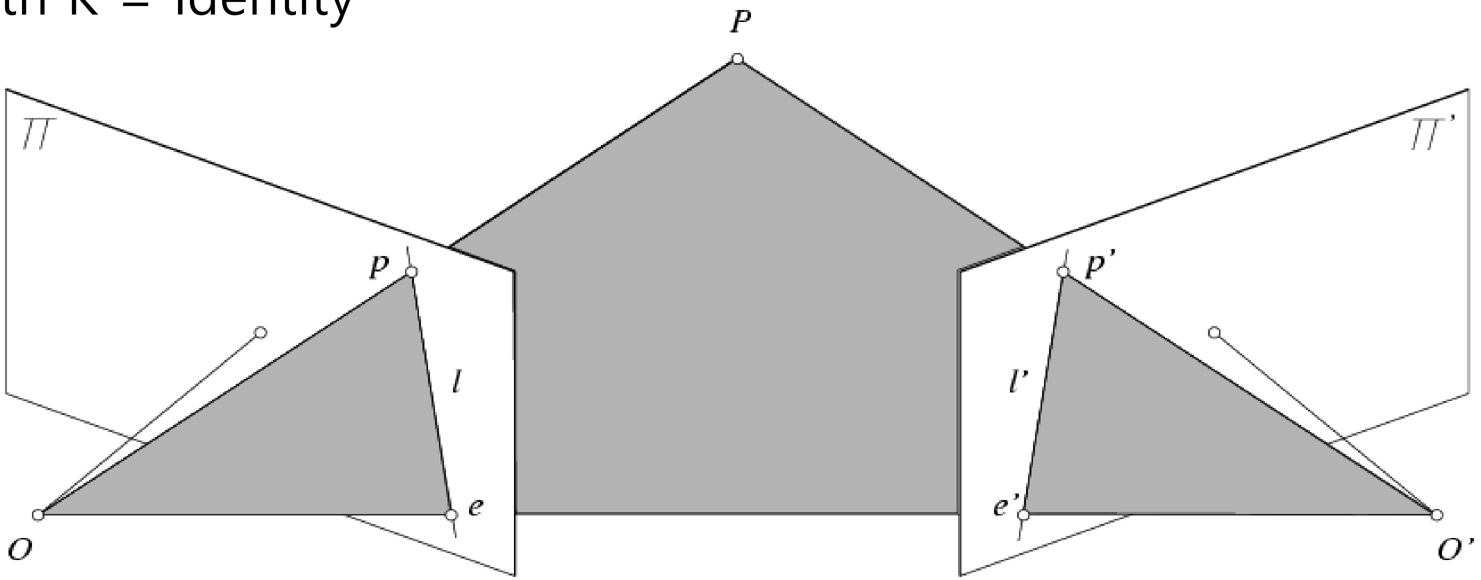
- Intrinsic and extrinsic parameters of the cameras are known, world coordinate system is set to that of the first camera

$$\boldsymbol{x}_{\text{norm}} = \boldsymbol{K}^{-1}\boldsymbol{x}_{\text{pixel}} = [\boldsymbol{I}\ \ 0]\,\boldsymbol{X}, \qquad \boldsymbol{x}'_{\text{norm}} = \boldsymbol{K}'^{-1}\boldsymbol{x}'_{\text{pixel}} = [\boldsymbol{R}\ \ \boldsymbol{t}]\,\boldsymbol{X}$$



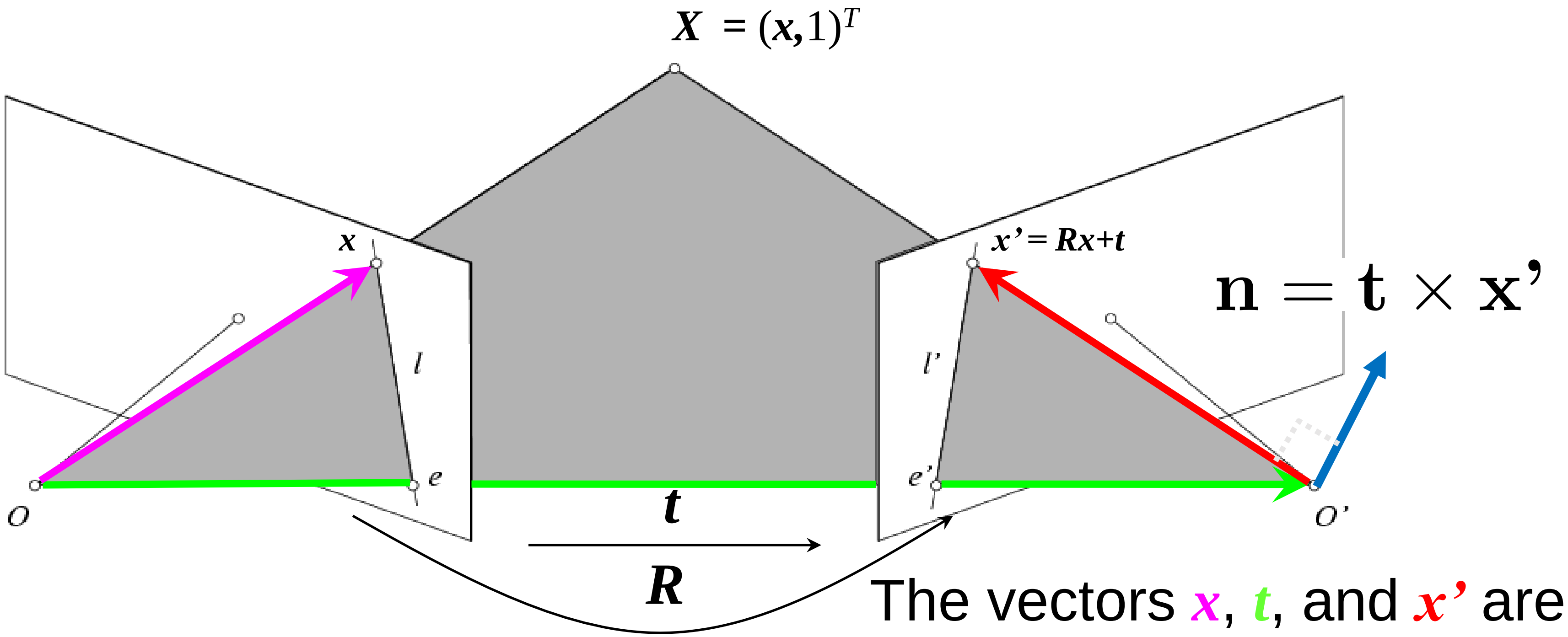
Normalized image coordinates

- We know the intrinsics K
- Recall that with intrinsics, we can unproject pixels to rays
- Make it into a 3D point at the image plane (z = f)
- This is called the *normalized* image coordinates. It may be thought of as a set of points with K = Identity



$$\boldsymbol{x}_{\text{norm}} = \boldsymbol{K}^{-1}\boldsymbol{x}_{\text{pixel}} = [\boldsymbol{I} \;\; 0]\,\boldsymbol{X}, \qquad \boldsymbol{x}'_{\text{norm}} = \boldsymbol{K}'^{-1}\boldsymbol{x}'_{\text{pixel}} = [\boldsymbol{R} \;\; \boldsymbol{t}]\,\boldsymbol{X}$$

Epipolar constraint: Calibrated case



The vectors x, t, and x' are coplanar

What can you say about their relationships, given $\mathbf{n} = \mathbf{t} \times \mathbf{x'}$?

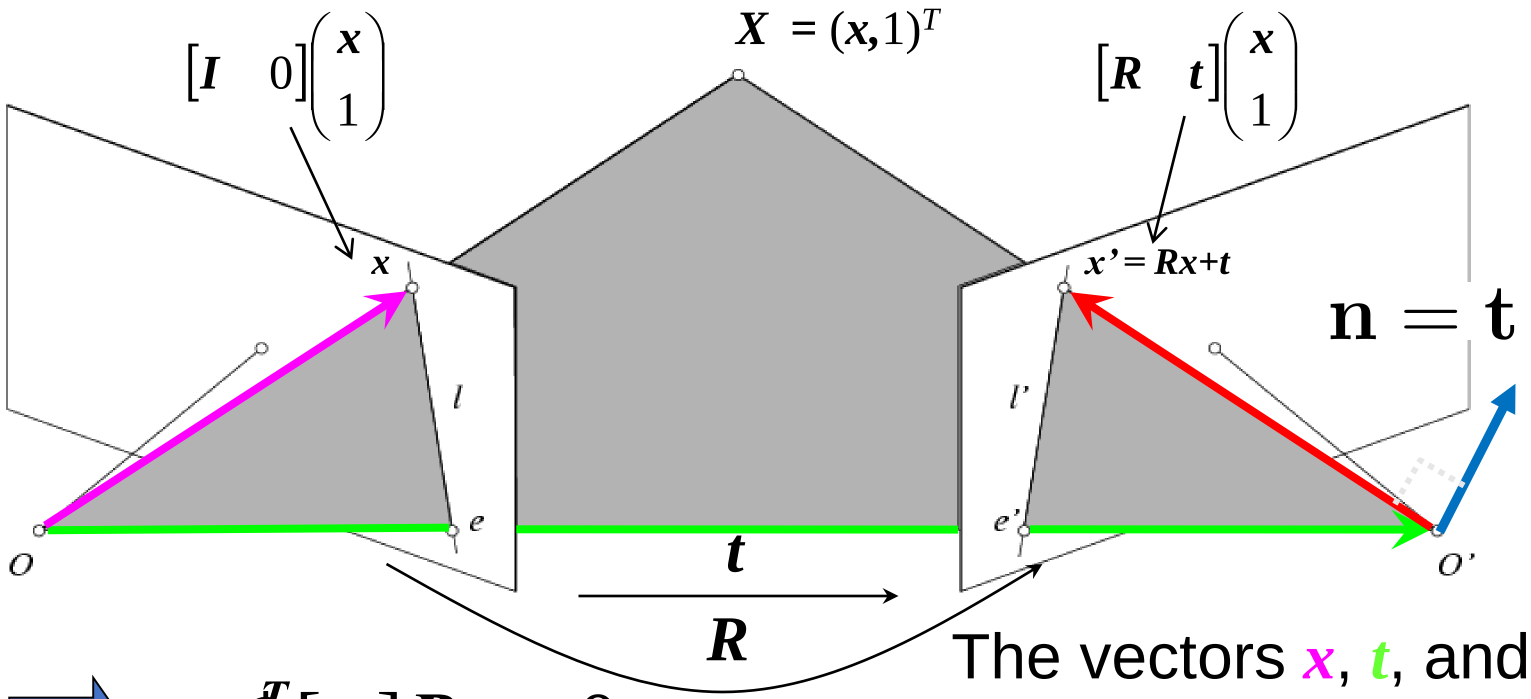
$$\boxed{\mathbf{x'} \cdot (\mathbf{t} \times \mathbf{x'}) = 0}$$

$$\mathbf{x'} \cdot (\mathbf{t} \times (R\mathbf{x} + \mathbf{t})) = 0$$

$$\mathbf{x'} \cdot (\mathbf{t} \times R\mathbf{x} + \underbrace{\cancel{\mathbf{t} \times \mathbf{t}}}_{0})) = 0$$

$$\boxed{\mathbf{x'} \cdot (\mathbf{t} \times R\mathbf{x}) = 0}$$

Epipolar constraint: Calibrated case



$$\mathbf{n} = \mathbf{t} \times \mathbf{x'}$$

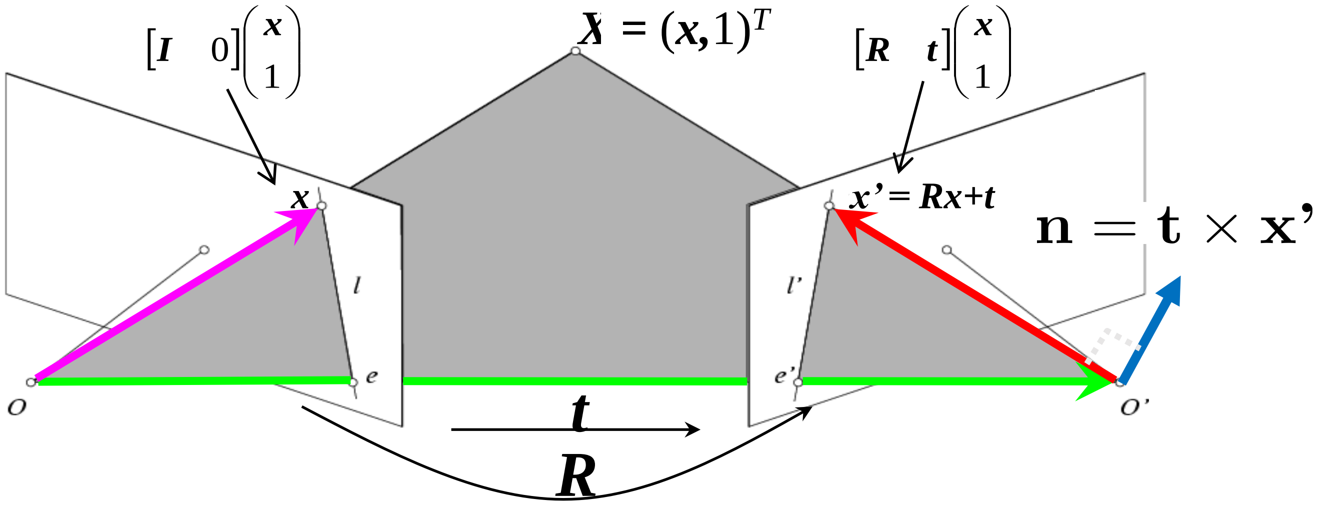
The vectors *x*, *t*, and *x'* are coplanar

$$x' \cdot [t \times (Rx)] = 0 \quad \Rightarrow \quad x'^T [t_\times] R x = 0$$

Recall: $\mathbf{a} \times \mathbf{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = [\mathbf{a}_\times]\mathbf{b}$

The vectors *Rx*, *t*, and *x'* are coplanar

Epipolar constraint: Calibrated case



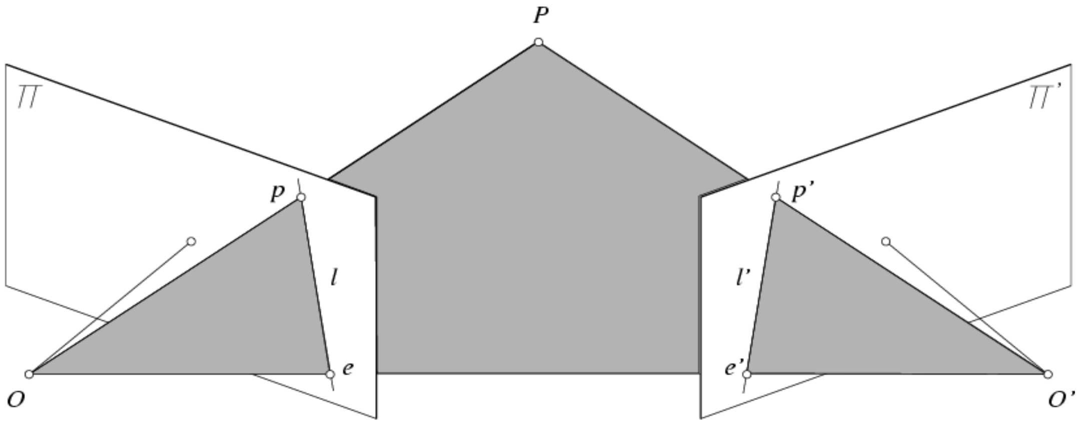
$$x' \cdot [t \times (Rx)] = 0 \;\Rightarrow\; x'^T \underbrace{[t_\times] R}_{E} x = 0 \;\Rightarrow\; x'^T E x = 0$$

Essential Matrix
(Longuet-Higgins, 1981)

Recall: $\mathbf{a} \times \mathbf{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = [\mathbf{a}_\times]\mathbf{b}$

The vectors x, t, and x' are coplanar

Epipolar constraint: Calibrated case

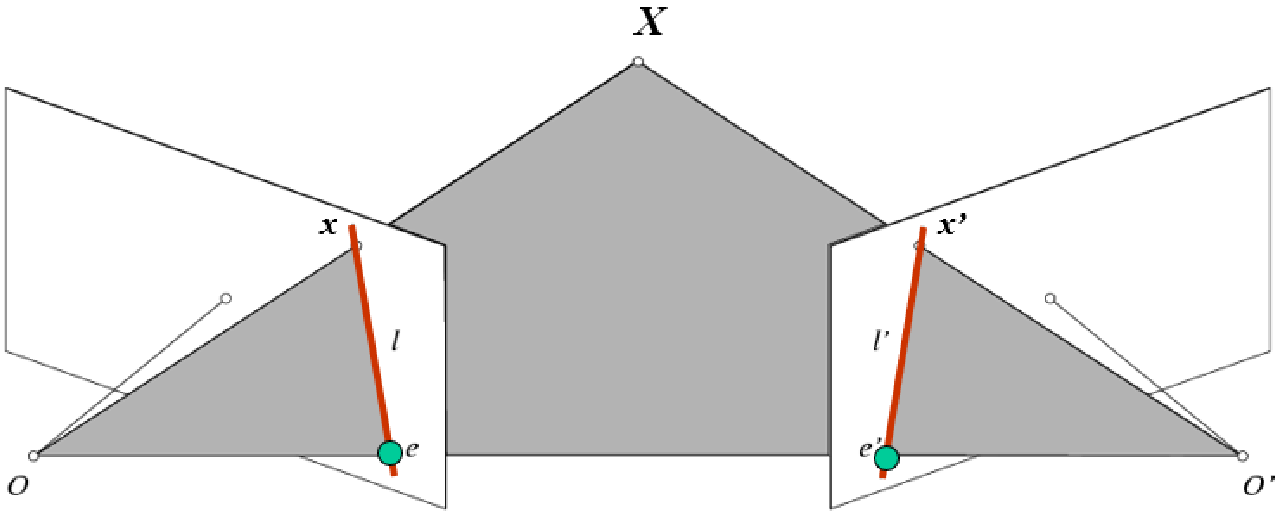


$$\boxed{x'^{T} E x = 0}$$

- ***E x*** is the epipolar line associated with ***x*** (***l'*** = ***E x***)
  - Recall: a line is given by $ax + by + c = 0$ or

$$\mathbf{l}^T\mathbf{x} = 0 \quad \text{where} \quad \mathbf{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Epipolar constraint: Calibrated case



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

- $\boldsymbol{E\,x}$ is the epipolar line associated with $\boldsymbol{x}$ ($\boldsymbol{l}' = \boldsymbol{E\,x}$)
- $\boldsymbol{E}^T\boldsymbol{x}'$ is the epipolar line associated with $\boldsymbol{x}'$ ($\boldsymbol{l} = \boldsymbol{E}^T\boldsymbol{x}'$)
- $\boldsymbol{E\,e} = 0$ and $\boldsymbol{E}^T\boldsymbol{e}' = 0$
- $\boldsymbol{E}$ is singular (rank two)
- $\boldsymbol{E}$ has five degrees of freedom

Epipolar constraint: Calibrated case

- Recall that we normalized the coordinates

$$x = K^{-1}\hat{x} \quad x' = K'^{-1}\hat{x}'$$

$$\hat{x} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

- where $\hat{x}$ is the image coordinates
- But now calibration matrices ***K*** and ***K′*** of the two cameras are unknown!
- We can write the epipolar constraint in terms of *unknown* normalized coordinates:

$$x'^T E x = 0$$

$$(K'^{-1}\hat{x}')'^T E (K^{-1}\hat{x}) = 0$$

$$\hat{x}'^T \underbrace{K'^{-T} E (K^{-1}}_{} \hat{x}) = 0$$

$$\hat{x}'^T F \hat{x} = 0$$

$$F = K'^{-T} E K^{-1}$$

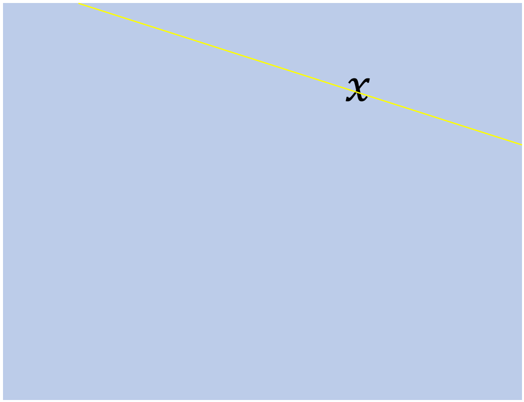
Fundamental Matrix
(Faugeras and Luong, 1992)

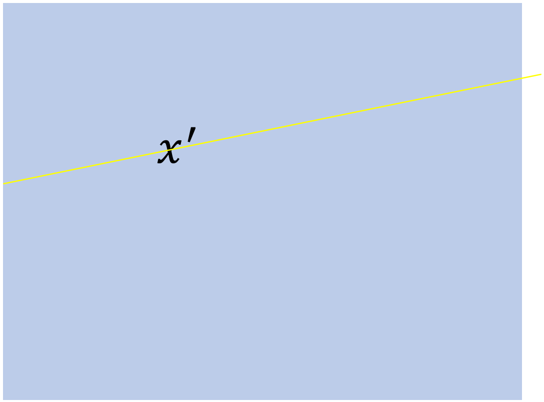
Fundamental Matrix

- Computing the Fundamental Matrix Given Corresponding Points

$$x'^T F x = 0$$

- x′, x : pixel coordinate, Ex) $x' = [300, 100, 1]^T$, $x = [150, 150, 1]^T$





Fundamental Matrix

- Computing the Fundamental Matrix Given Corresponding Points
- Matching points 1~N

$$x_n'^T F x_n = 0$$

- or

$$\underbrace{[x_n', y_n', 1]}_{\text{known}} \underbrace{\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}}_{\text{unknown}} \underbrace{\begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix}}_{\text{known}} = 0$$

Fundamental Matrix

- Linear Dependency

$$[x_n', y_n', 1]\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}\begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} = 0$$

$$x_n' x_n f_{11} + x_n' y_n f_{12} + x_n' f_{13} + y_n' x_n f_{21} \dots . f_{33} = 0$$

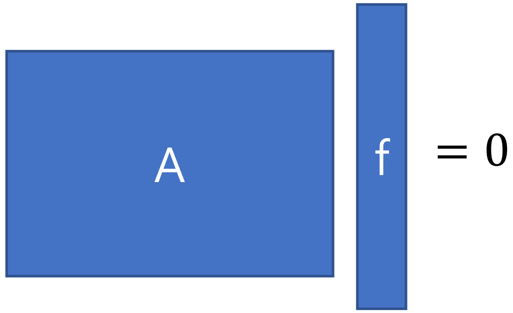
- In the matrix form

$$[x_n' x_n \quad x_n' y_n \quad x_n' \quad y_n' x_n \quad y_n' y_n \quad y_n' \quad x_n \quad y_n \quad 1]\begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$

Fundamental Matrix

- In the matrix form

$$\begin{bmatrix} x_1'x_1 & x_1'y_1 & x_1' & y_1'x_1 & y_1'y_1 & y_1' & x_1 & y_1 & 1 \\ x_2'x_2 & x_2'y_2 & x_2' & y_2'x_2 & y_2'y_2 & y_2' & x_2 & y_2 & 1 \\ & & & & \cdots & & & & \\ x_n'x_n & x_n'y_n & x_n' & y_n'x_n & y_n'y_n & y_n' & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$



f is null space of A → solve using Singular value decomposition

Fundamental Matrix

- Solution of Af=0, reminder

$$[1 \times 9]^T = \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix}$$

- The last row of V^T = the last column of V

reshape

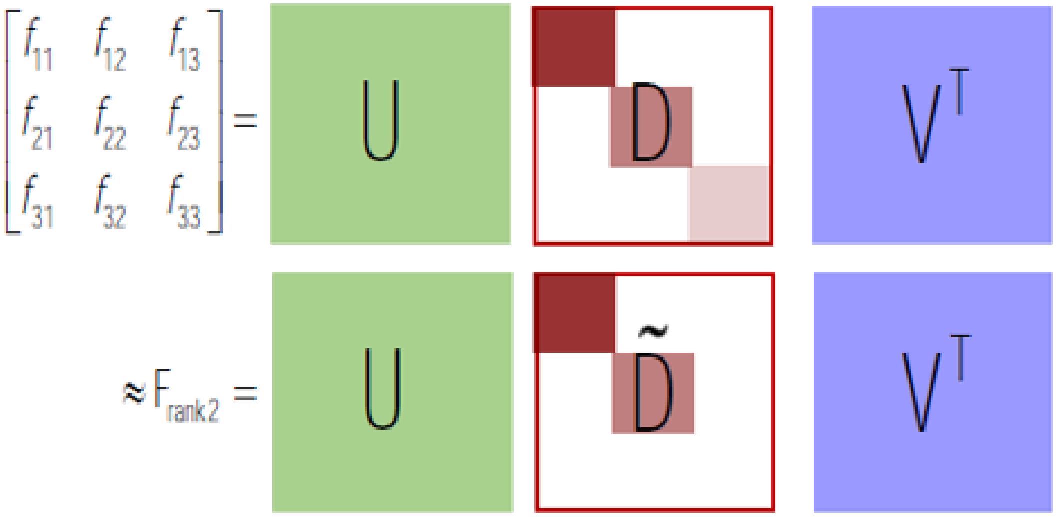
$$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

Fundamental Matrix

- Enforcing Rank 2 matrix
  - SVD decomposition (again) of F

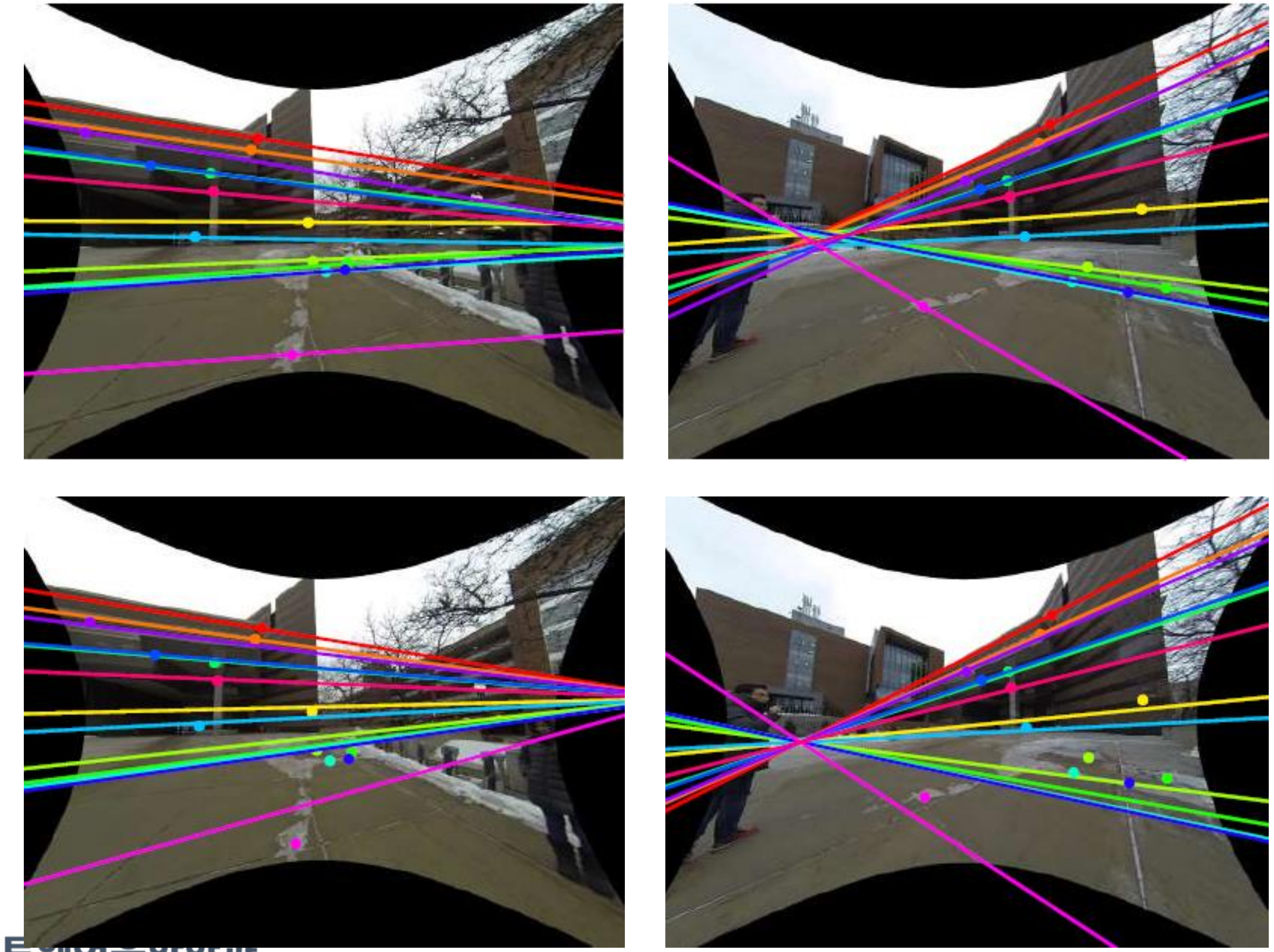
$$F = SVD(F) = UDV^T$$

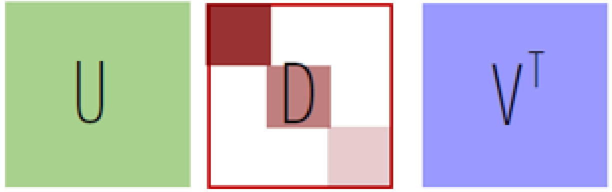
  - Note, these U, D, Vt is irrelevant to the previous ones.
  - Remove the last singular value (to 0) and compose to F again

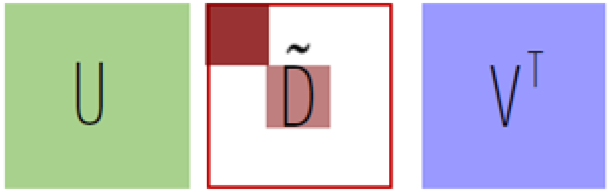


Fundamental Matrix

- F(rank 2) vs. F(full rank)







Fundamental Matrix

- Compute F processing
  - Feature point extraction
  - RANSAC
    - Get sample matching pairs
    - F estimation
    - Scoring (determine inliers)

- The same method to compute homography

Fundamental Matrix

- Scoring (error estimation)
  - Value dependent metric (not useful)

$$\frac{1}{N}\sum_{n=1}^{N} x_n'^T F x_n \leq \theta$$

  - Sampson error

$$\sum \frac{x'^T Fx}{\left(x'^T F\right)_1^2 + \left(x'^T F\right)_2^2 + \left(Fx\right)_1^2 + \left(Fx\right)_2^2}$$

  - Symmetric epipolar error

$$= \sum x'^T Fx \left(\frac{1}{\left(x'^T F\right)_1^2 + \left(x'^T F\right)_2^2} + \frac{1}{\left(Fx\right)_1^2 + \left(Fx\right)_2^2} \right)$$

Essential Matrix

- Pose estimation from Essential Matrix
- This means we know the intrinsic parameters of cameras
- Why can we compute R, t from essential matrix?

$$E = [t_\times]R$$

Note. $[t_\times] = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$

- How can we get Essential matrix?
  - From Fundamental matrix

$$E = K'^T F K$$

  - From least-square method

$$\hat{x}'^T E \hat{x} = 0$$

$\hat{x}', \hat{x}$: normalized points on the camera coordinate

Pose estimation from Essential Matrix

- Properties of the Essential matrix
  - Homogeneous
  - Singular (determinant is 0)
  - Two identical non-zero singular values
- Decomposition

$$E = U\Sigma V^T$$

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

- How to get R, t from U, Σ, V^T
  - There are various solution, we will use Hartley & Zisserman method

Pose estimation from Essential Matrix

- Hartley & Zisserman method

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

- Set two matrix Z(skew-symmetric), W (orthonormal), s.t. ZW = $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Pose estimation from Essential Matrix

- There are for solution of Z, W

$$
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = ZW = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

$$
= -Z^{\mathsf{T}}W = -\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

$$
= -ZW^{\mathsf{T}} = -\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{\mathsf{T}}
$$

$$
= Z^{\mathsf{T}}W^{\mathsf{T}} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{\mathsf{T}}
$$

Pose estimation from Essential Matrix

- Solution by Hartley & Zisserman

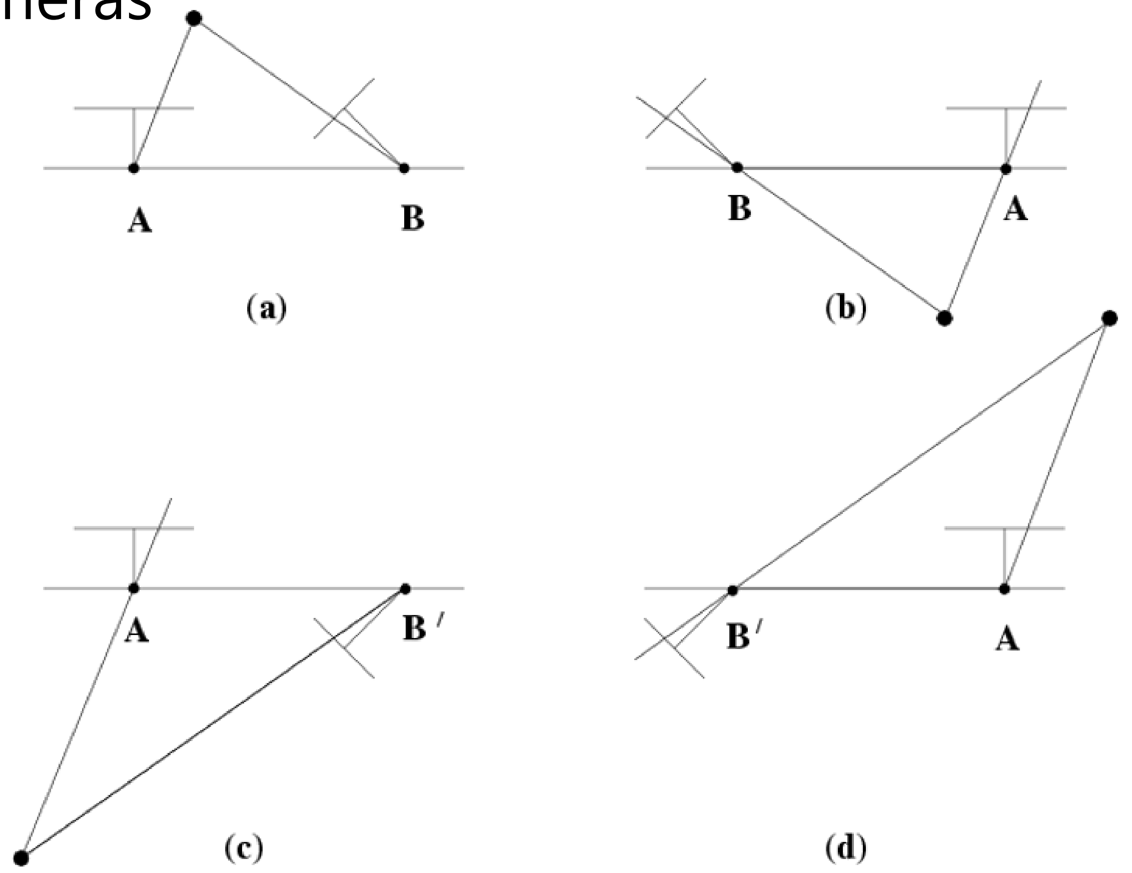
$$
\begin{aligned}
\mathrm{E} &= U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^{\mathsf{T}} \\
&= U \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{Z} \underbrace{\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{W} V^{\mathsf{T}} \\
&= U Z \underbrace{U^{\mathsf{T}} U}_{I} W V^{\mathsf{T}} \\
&= \underbrace{U Z U^{\mathsf{T}}}_{[t_\times]} \underbrace{U W V^{\mathsf{T}}}_{R^{\mathsf{T}}}
\end{aligned}
$$

Note. $[t_\times] = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$

Pose estimation from Essential Matrix

- Solution by Hartley & Zisserman
- There are 4 solutions, but only 1 solution is feasible
  - 3D Points are in front of both cameras

$$E^1 = UZU^T \; UWV^T$$
$$E^2 = UZ^TU^T \; UWV^T$$
$$E^3 = UZU^T \; UW^TV^T$$
$$E^4 = UZ^TU^T \; UW^TV^T$$



3D Reconstruction

- Linear triangulation
- Not rectified case

$$x = K[I \mid 0]X \qquad x = K'[R \mid t]X$$

$$x = \alpha PX \qquad x' = P'X$$

$$x \times PX = 0$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} — & \boldsymbol{p}_1^\top & — \\ — & \boldsymbol{p}_2^\top & — \\ — & \boldsymbol{p}_3^\top & — \end{bmatrix} \begin{bmatrix} | \\ \boldsymbol{X} \\ | \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} \boldsymbol{p}_1^\top \boldsymbol{X} \\ \boldsymbol{p}_2^\top \boldsymbol{X} \\ \boldsymbol{p}_3^\top \boldsymbol{X} \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \times \begin{bmatrix} \boldsymbol{p}_1^\top \boldsymbol{X} \\ \boldsymbol{p}_2^\top \boldsymbol{X} \\ \boldsymbol{p}_3^\top \boldsymbol{X} \end{bmatrix} = \begin{bmatrix} y\boldsymbol{p}_3^\top \boldsymbol{X} - \boldsymbol{p}_2^\top \boldsymbol{X} \\ \boldsymbol{p}_1^\top \boldsymbol{X} - x\boldsymbol{p}_3^\top \boldsymbol{X} \\ \boxed{x\boldsymbol{p}_2^\top \boldsymbol{X} - y\boldsymbol{p}_1^\top \boldsymbol{X}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

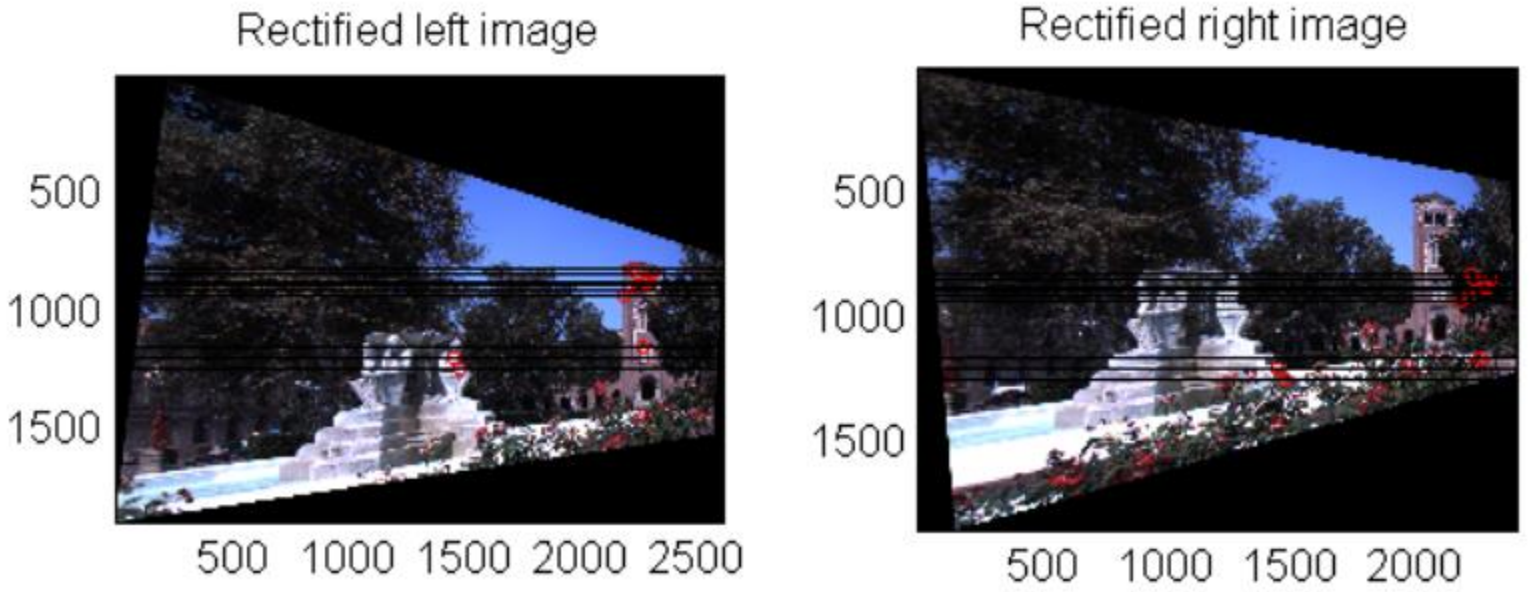
Linearly dependent

Including right image

$$\begin{bmatrix} y\boldsymbol{p}_3^\top - \boldsymbol{p}_2^\top \\ \boldsymbol{p}_1^\top - x\boldsymbol{p}_3^\top \\ y'\boldsymbol{p}_3'^\top - \boldsymbol{p}_2'^\top \\ \boldsymbol{p}_1'^\top - x'\boldsymbol{p}_3'^\top \end{bmatrix} \boldsymbol{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

3D Reconstruction

- Dense matching
  - Stereo image rectification
  - Block matching
  - Triangulation



Thank you