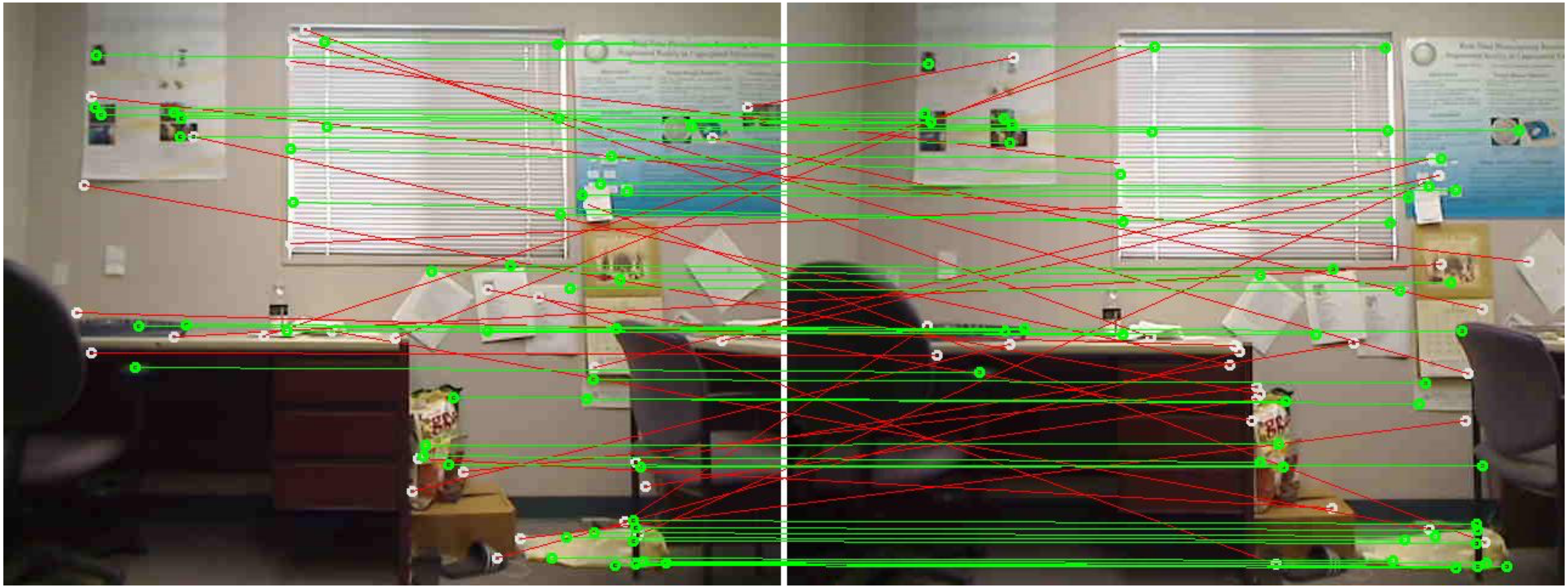
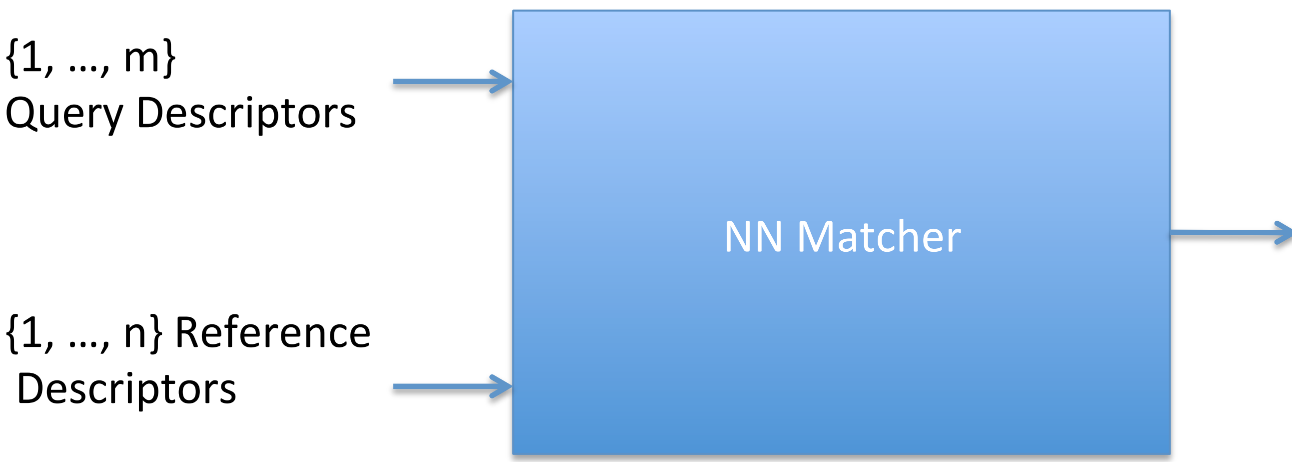
MATCHING FEATURES

Feature Matching Example



Nearest-Neighbor feature matcher



Correspondences

Query Index	Reference Index
1	5
…	…
2	7

Nearest-Neighbor feature matcher (cont’d)

Reference features

Query features

	1	2	…	n
1	d(1, 1)	d(1, 2)		d(1, n)
2	d(2, 1)	d(2, 2)		d(2, n)
…	…			
m	d(m, 1)	d(m, 2)		d(m, n)

The index of the minimum distance for every row is the best match for a given query.

Pseudo-algorithm for NN matcher

For every feature *i* in Query Set:

    For every feature *j* in Reference Set:

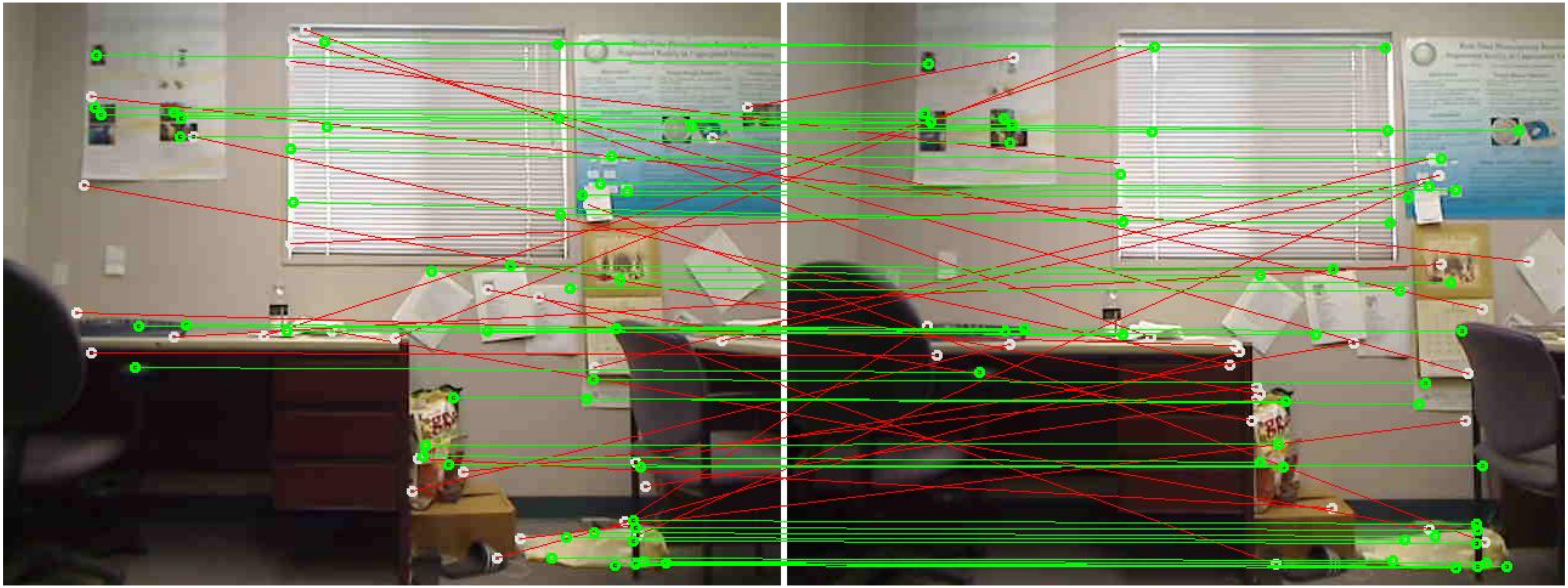
        d(*i, j*) = d(*i, j*) // Compare feature *i* and *j*

    Create correspondence {*i* <-> *j*}, where *j* = min D(*i*, :)

    end for

end for

Feature Matching Example



There are good correspondences and bad ones!

Heuristic to filter bad matches

- David Lowe proposed the following statistic:

$$r = \frac{d_{1_{NN}}}{d_{2_{NN}}}$$

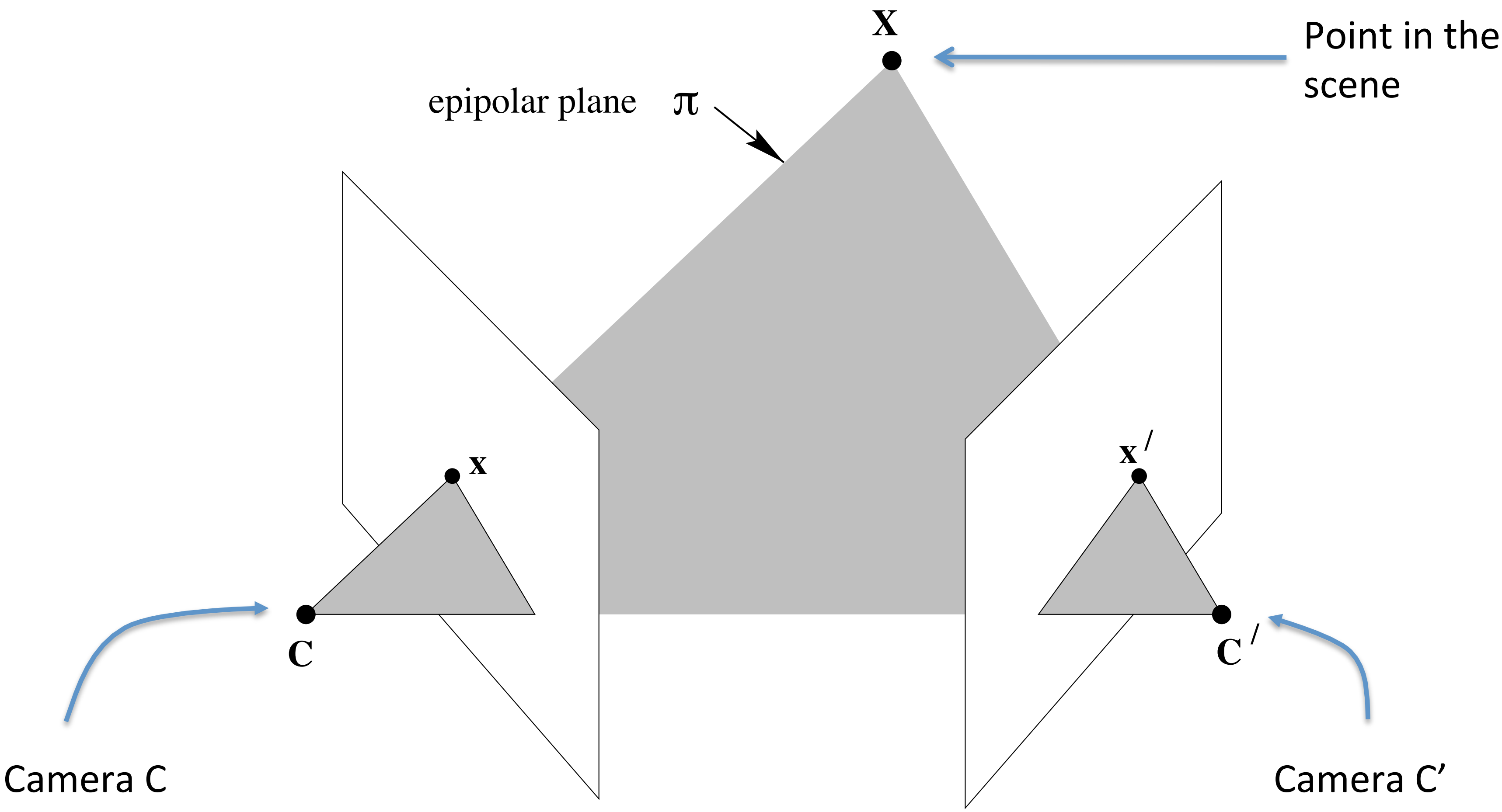
Smallest distance

Second smallest distance

- Accept match if r < 0.75, reject otherwise.

TWO-VIEW GEOMETRY

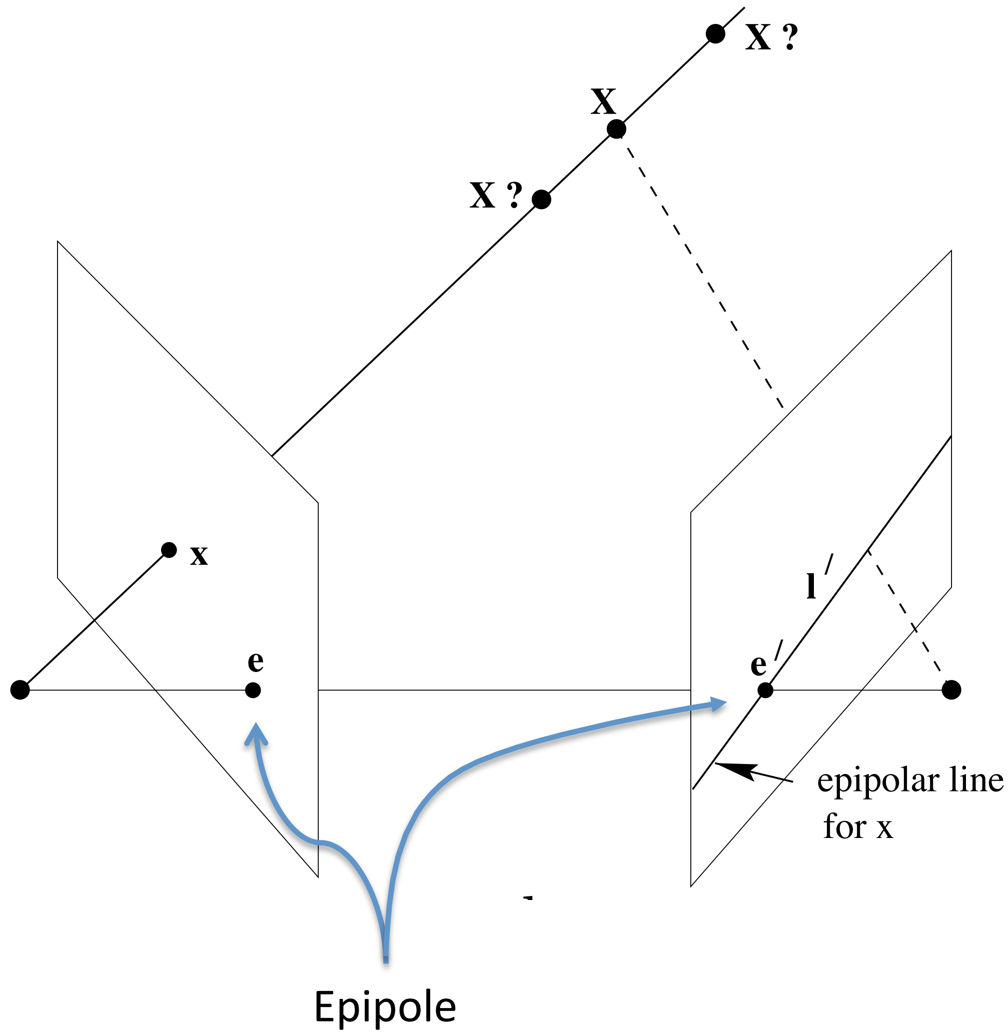
Point Correspondence Geometry



Points x and x’, which are the images of point X, are coplanar.

Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

Geometric constraint of a correspondence



x’ must be somewhere in the epipolar line l’

Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

The baseline intersects the two epipoles

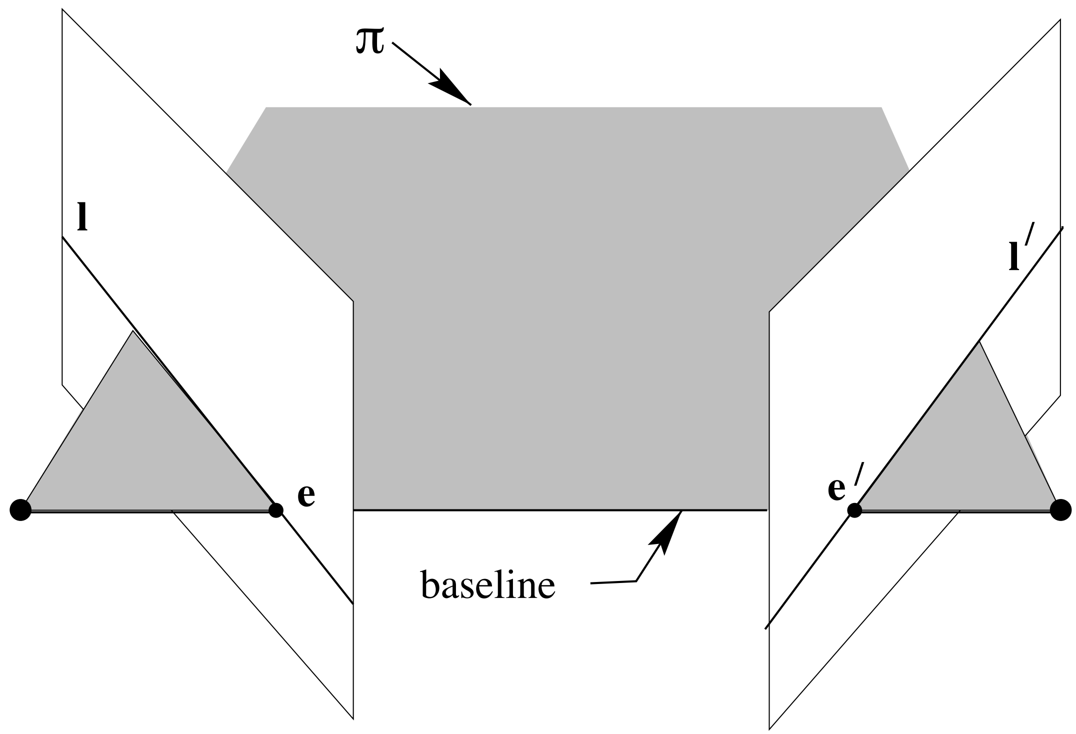
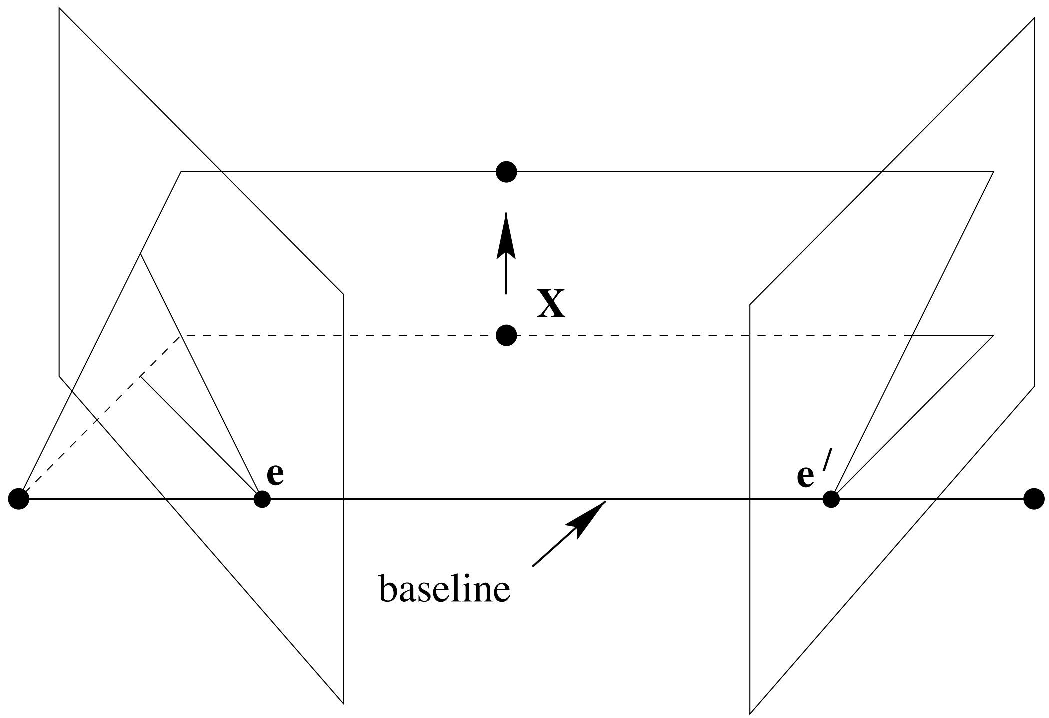


Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

Epipolar pencil



As the position of the 3D point X varies, the epipolar planes rotate about the baseline. This family of planes are called epipolar pencil.

Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

Example of Epipolar geometry

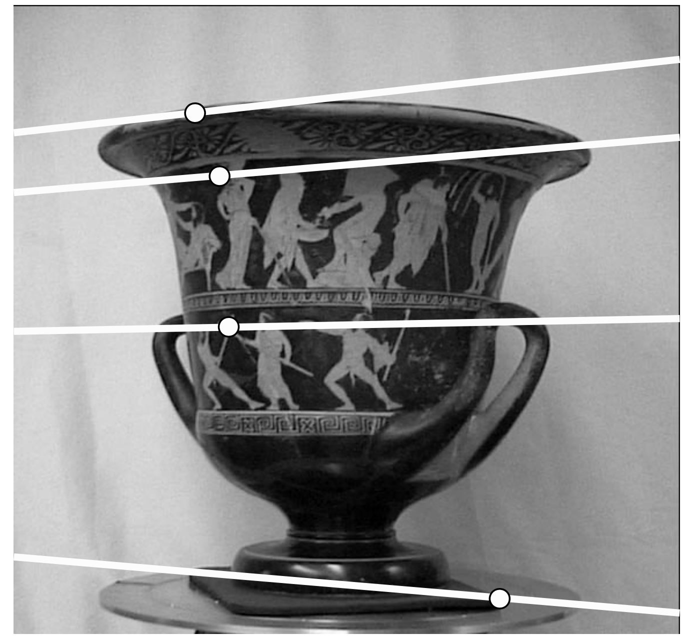
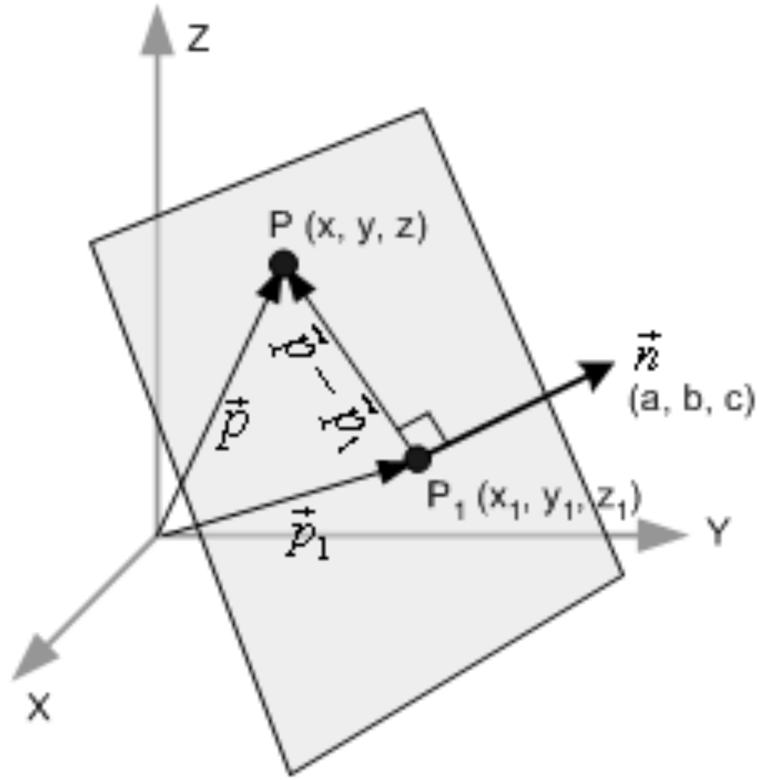


Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

Summary of new concepts

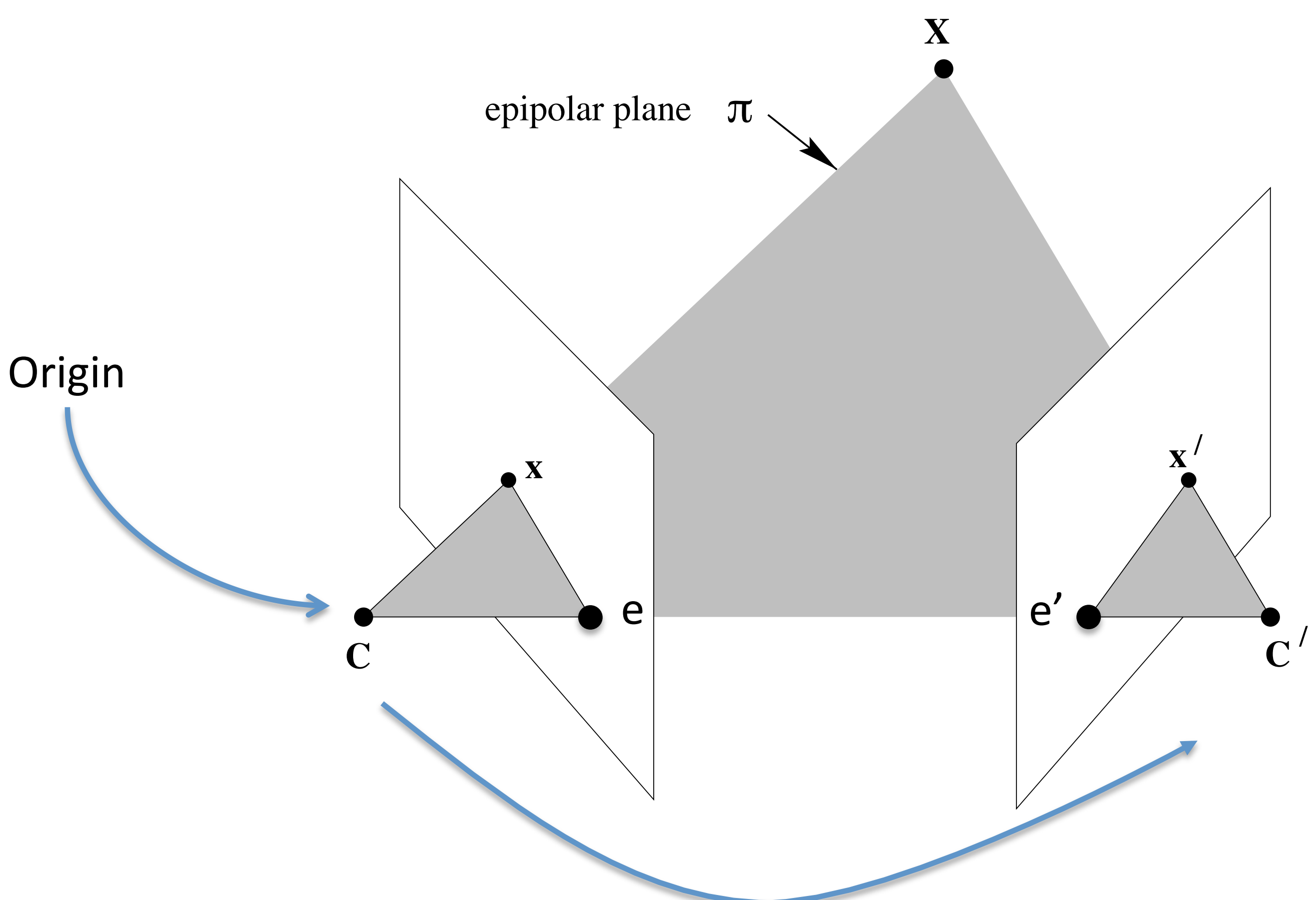
- **Epipole**: The point of intersection of the line joining the camera centers (i.e., the baseline)
- **Epipolar plane**: The plane containing the baseline.
- **Epipolar line**: The intersection of an epipolar plane with the image plane.

Review of planes



$$(\mathbf{P} - \mathbf{P}_1) \cdot \mathbf{n} = 0$$

Essential matrix

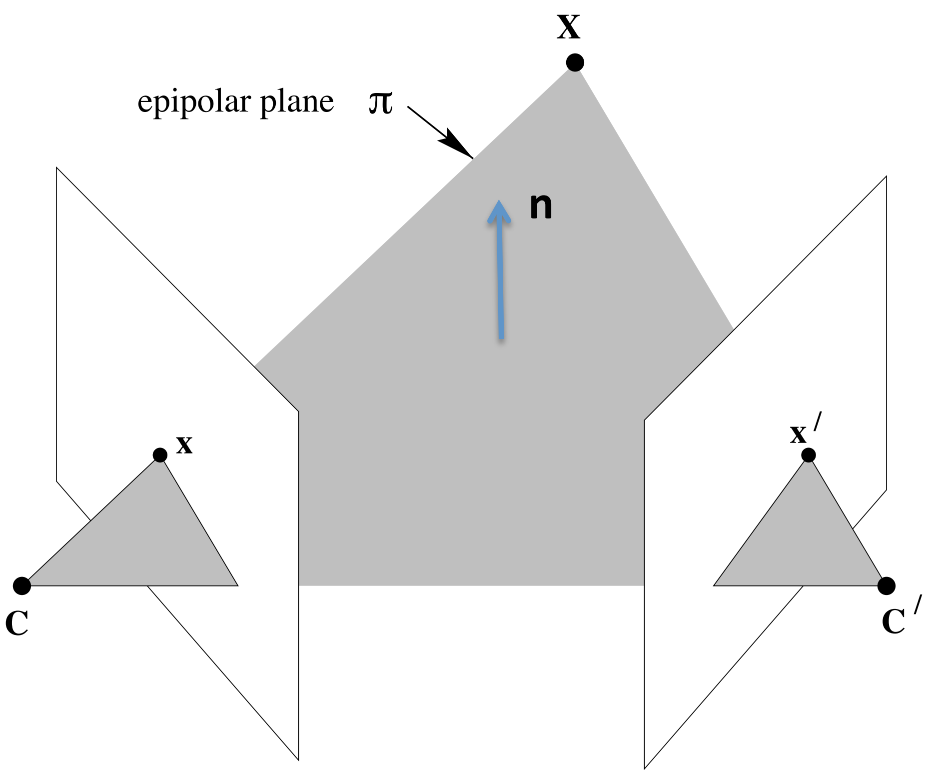


$$T_C^{C'} = \begin{bmatrix} R & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

Assume we know the relative Pose

Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

Essential matrix (cont'd)



$$\mathbf{t} \times (R\mathbf{x}) \propto \mathbf{n}$$

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

What happens if I swap x' with x? Does it still hold?

Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

Cross-product as a matrix product

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

Essential matrix

$$\mathbf{x}' \cdot \left([\mathbf{t}]_\times R \right) \mathbf{x} = 0$$

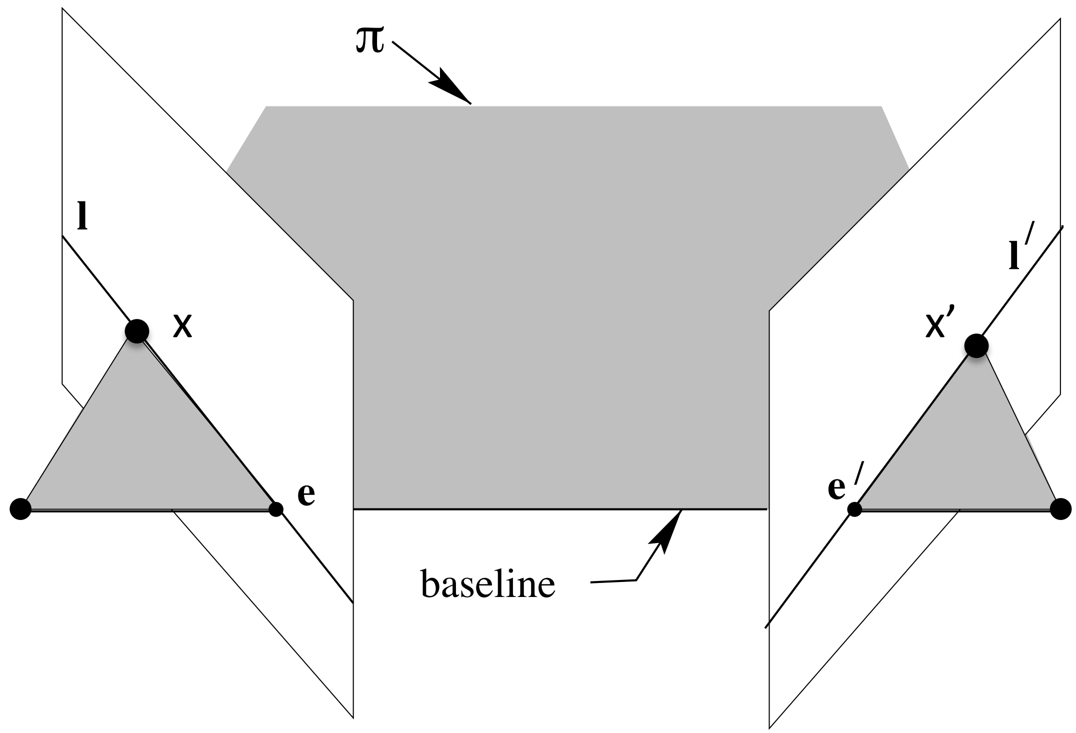
$$E = [\mathbf{t}]_\times R$$

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

How many degrees of freedom does the Essential matrix have?

- We have 3 parameters for rotation
- We have 3 parameters for translation
- Thus, only 5 parameters are required!

Epipolar lines



$$\mathbf{l}' = E\mathbf{x}$$

$$\mathbf{l} = E^T\mathbf{x}'$$

Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

Discussion

- Essential matrix considers that **x** and **x**' are points on the plane (a.k.a. normalized coordinates).

- What can we do to accept pixels?

Review of the pinhole projection model

$$\hat{\mathbf{p}} = \frac{1}{Z} \underbrace{K\,[R\ \mathbf{t}]}_{M} \hat{\mathbf{P}}$$

Pixel coordinates

Point in the world

Normalized coordinates

$$K^{-1}\hat{\mathbf{p}} = \frac{1}{Z}\left[R \;\; \mathbf{t}\right]\hat{\mathbf{P}}$$

Using our notation for the epipolar geometry

$$\mathbf{x} = K^{-1}\hat{\mathbf{p}}_{\mathbf{x}} = \frac{1}{Z}\left[R \;\; \mathbf{t}\right]\hat{\mathbf{X}}$$

Fundamental matrix

$$\underbrace{\left(K_{C'}^{-1}\hat{\mathbf{p}}_{\mathbf{x}'}\right)^{T}}_{\mathbf{x'^T}} E \underbrace{K_{C}^{-1}\hat{\mathbf{p}}_{\mathbf{x}}}_{\mathbf{x}} = 0$$

$$\hat{\mathbf{p}}_{\mathbf{x}'}^{T} \underbrace{K_{C'}^{-T} E K_{C}^{-1}}_{F} \hat{\mathbf{p}}_{\mathbf{x}} = 0$$

$$\boxed{\hat{\mathbf{p}}_{\mathbf{x}'}^{T} F \hat{\mathbf{p}}_{\mathbf{x}} = 0 \qquad F = K_{C'}^{-T} E K_{C}^{-1}}$$

What is the rank of E and F?

- The rank of E and F is 2.
- Why?

Epipolar constraint

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

Pixel coordinate
(homogeneous)

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

Normalized
Coordinates

Checks that the correspondences are coplanar!

Summary of fundamental matrix properties

- $\mathtt{F}$ is a rank 2 homogeneous matrix with 7 degrees of freedom.
- **Point correspondence**: If $\mathbf{x}$ and $\mathbf{x}'$ are corresponding image points, then $\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = 0$.
- **Epipolar lines**:
  - ◇ $\mathbf{l}' = \mathtt{F}\mathbf{x}$ is the epipolar line corresponding to $\mathbf{x}$.
  - ◇ $\mathbf{l} = \mathtt{F}^{\top}\mathbf{x}'$ is the epipolar line corresponding to $\mathbf{x}'$.
- **Epipoles**:
  - ◇ $\mathtt{F}\mathbf{e} = \mathbf{0}$.
  - ◇ $\mathtt{F}^{\top}\mathbf{e}' = \mathbf{0}$.
- **Computation from camera matrices** $\mathtt{P}, \mathtt{P}'$:
  - ◇ General cameras, $\mathtt{F} = [\mathbf{e}']_{\times}\mathtt{P}'\mathtt{P}^{+}$, where $\mathtt{P}^{+}$ is the pseudo-inverse of $\mathtt{P}$, and $\mathbf{e}' = \mathtt{P}'\mathbf{C}$, with $\mathtt{P}\mathbf{C} = \mathbf{0}$.
  - ◇ Canonical cameras, $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = [\mathtt{M} \mid \mathbf{m}]$, $\mathtt{F} = [\mathbf{e}']_{\times}\mathtt{M} = \mathtt{M}^{-\top}[\mathbf{e}]_{\times}$, where $\mathbf{e}' = \mathbf{m}$ and $\mathbf{e} = \mathtt{M}^{-1}\mathbf{m}$.
  - ◇ Cameras not at infinity $\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}]$, $\mathtt{F} = \mathtt{K}'^{-\top}[\mathbf{t}]_{\times}\mathtt{R}\mathtt{K}^{-1} = [\mathtt{K}'\mathbf{t}]_{\times}\mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\top}\mathtt{R}\mathtt{K}^{\top}[\mathtt{K}\mathtt{R}^{\top}\mathbf{t}]_{\times}$.

Image from. R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge Press.

Discussion

- Is there any configuration of the two cameras that can cause issues while trying to estimate the fundamental matrix?
- Does this work with cameras equipped with lenses?

Homography

- Assume that the points in the scene are coplanar.

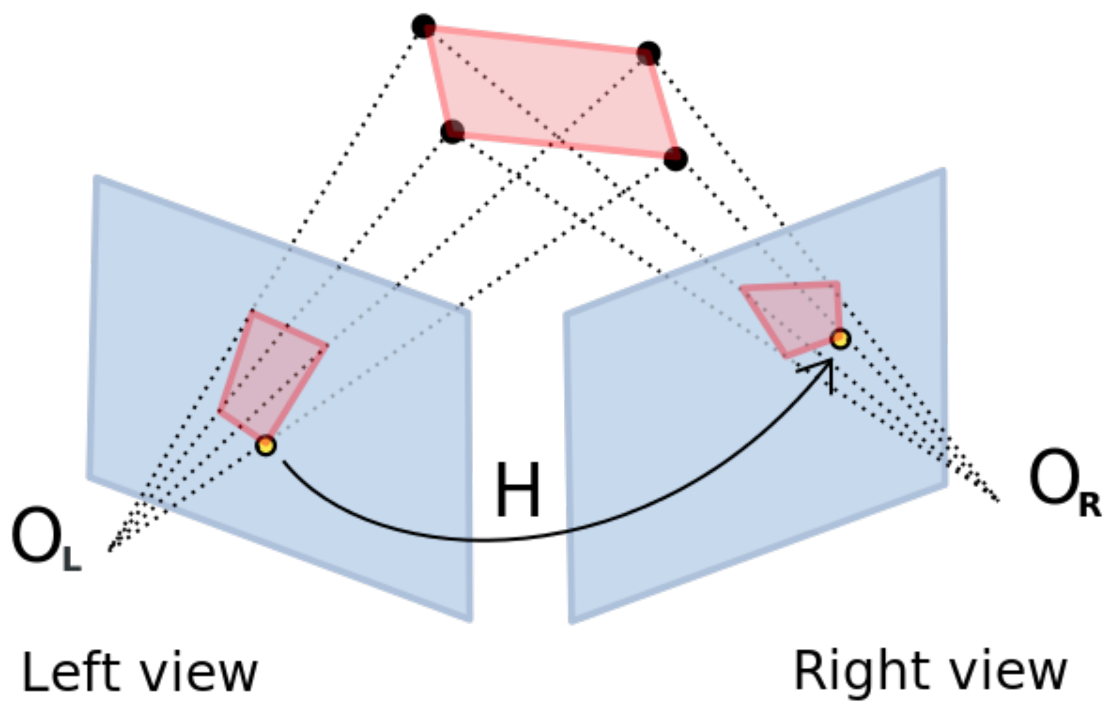


Image from OpenMVG: http://imagine.enpc.fr/~moulonp/openMVG/

Homography (cont'd)

- It is a plane-to-plane transformation.

$$\hat{\mathbf{p}}_{\mathbf{x}'} = H\hat{\mathbf{p}}_{\mathbf{x}}$$

- Maps pixel from one image into another pixel in a second image.
- What rank is it?

ESTIMATIONS FROM CORRESPONDENCES

Estimation

- The goal from the estimation is to calculate a Homography, a Essential matrix, or the Fundamental matrix.
- The input data are the feature matches (or pixel correspondences).

Homography estimation

- If we know that $\mathbf{x}' \leftrightarrow \mathbf{x}$

$$\underbrace{\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix}}_{\mathbf{x}'} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \underbrace{\begin{bmatrix} u \\ v \\ w \end{bmatrix}}_{\mathbf{x}}$$

Our goal is to estimate *H*

- We know that $\mathbf{x}' \times H\mathbf{x} = 0$
- The homography is a 3x3 matrix:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{33} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{h}^{1T} \\ \mathbf{h}^{2T} \\ \mathbf{h}^{3T} \end{bmatrix}$$

- We can write the following expression:

$$\mathbf{x}' \times H\mathbf{x} = \begin{bmatrix} v'\mathbf{h}^{3T}\mathbf{x} - w'\mathbf{h}^{2T}\mathbf{x} \\ w'\mathbf{h}^{1T}\mathbf{x} - u'\mathbf{h}^{3T}\mathbf{x} \\ u'\mathbf{h}^{2T}\mathbf{x} - v'\mathbf{h}^{1T}\mathbf{x} \end{bmatrix}$$

Our goal is to estimate H (cont'd)

- Factoring out the rows of H, we get:

$$\begin{bmatrix} \mathbf{0}^T & -w'\mathbf{x}^T & v'\mathbf{x}^T \\ w'\mathbf{x}^T & \mathbf{0}^T & -u'\mathbf{x}^T \\ -v'\mathbf{x}^T & u'\mathbf{x}^T & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{bmatrix} = \mathbf{0}$$

Eq. 1

- For every correspondence $\mathbf{x}' \leftrightarrow \mathbf{x}$ we have 2 equations that are linearly independent.
- Only the first two equations are used.

Our goal is to estimate H (cont'd)

- Eq. 1 can be re-written as follows:

$$\underbrace{\begin{bmatrix} \mathbf{0}^T & -w'\mathbf{x}^T & v'\mathbf{x}^T \\ w'\mathbf{x}^T & \mathbf{0}^T & -u'\mathbf{x}^T \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{bmatrix}}_{\mathbf{h}} = \mathbf{0}$$

What can we do to solve for H?

Review of Singular Value Decomposition

- A matrix can be decomposed as:

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

- Where U and V are orthogonal matrices.
- Matrix Σ is a diagonal matrix.
- The singular values are sorted by descending order.

Basic Direct Linear Transform (DLT) for computing Homography

- Input: Given n >= 4 correspondences $\{\mathbf{x}'_i \leftrightarrow \mathbf{x}_i\}$

1. For each correspondence, keep the two equations (rows) from the Eq. 1.
2. Assemble a n * 2 x 9 matrix by stacking up the two equations from each correspondence.
3. Obtain the SVD of matrix A (the assemble). Keep the unit singular vector from V corresponding to the smallest singular value.
4. Matrix H is obtained from the kept vector.

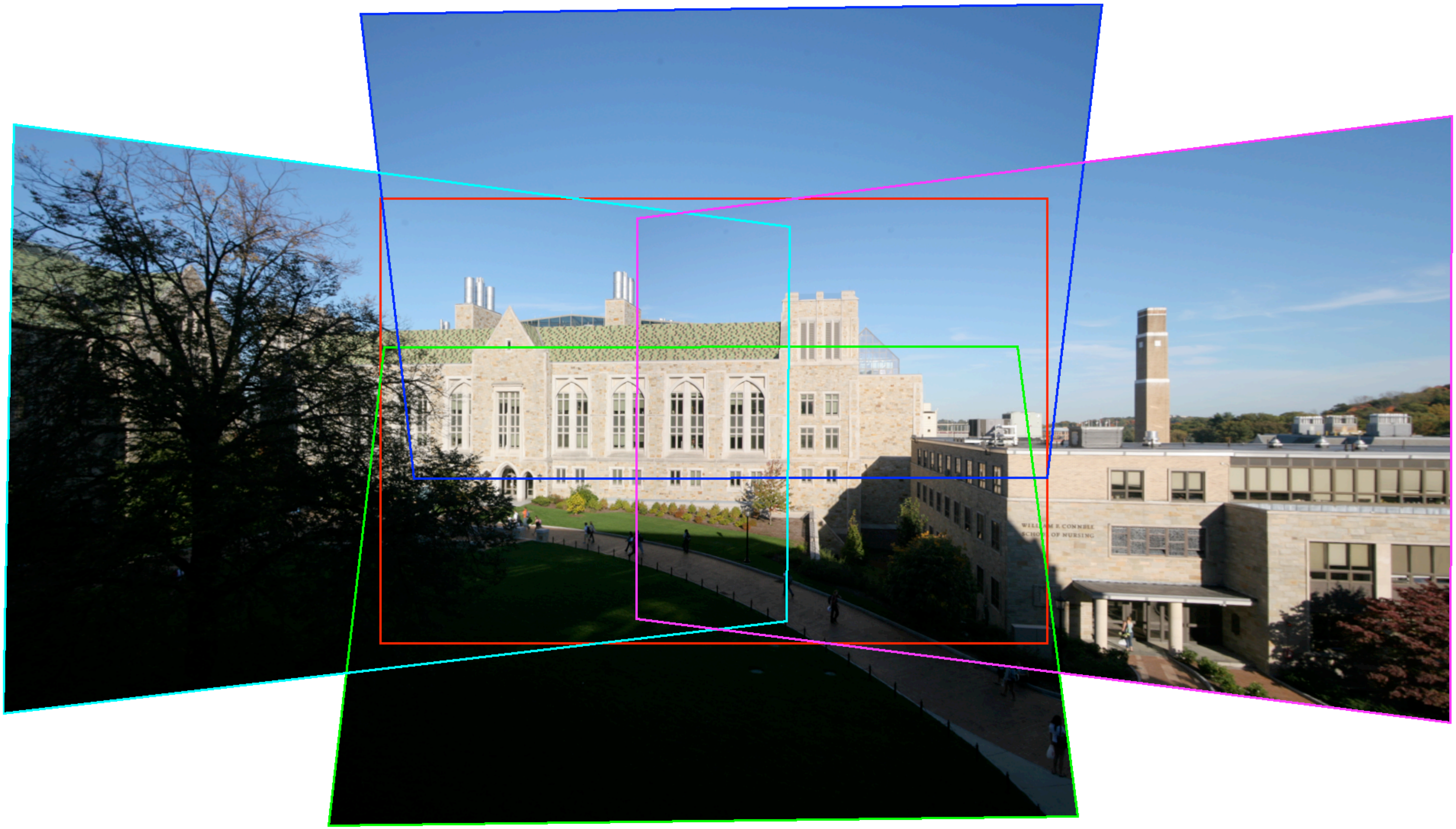
Notes on DLT for Homography

- Usually after solving for H, the element h_{33} is set = 1.
- In other words, after getting H from the SVD, do $H = (1 / h_{33})\ H$.

Discussion

- Is the DLT robust to noise?
- Are there any degenerate cases?
  - If the correspondences are collinear, DLT does not work.

Image Stitching: An applications of Homography



Next class

- We will continue with estimation of Fundamental matrix.
- We will start studying Robust estimation using RANSAC.