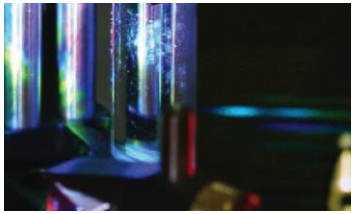
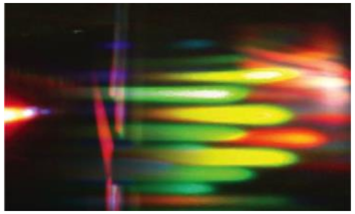
Imperial College London





COMP70058 Computer Vision

Lecture 14 – Computational Stereo (Part 1)

Stamatia (Matina) Giannarou, PhD
The Hamlyn Centre for Robotic Surgery

stamatia.giannarou@imperial.ac.uk

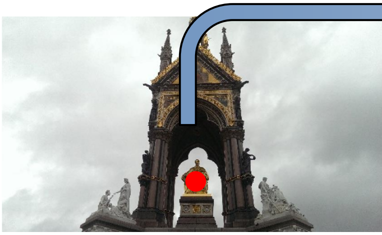
The Hamlyn Centre for Robotic Surgery

Contents

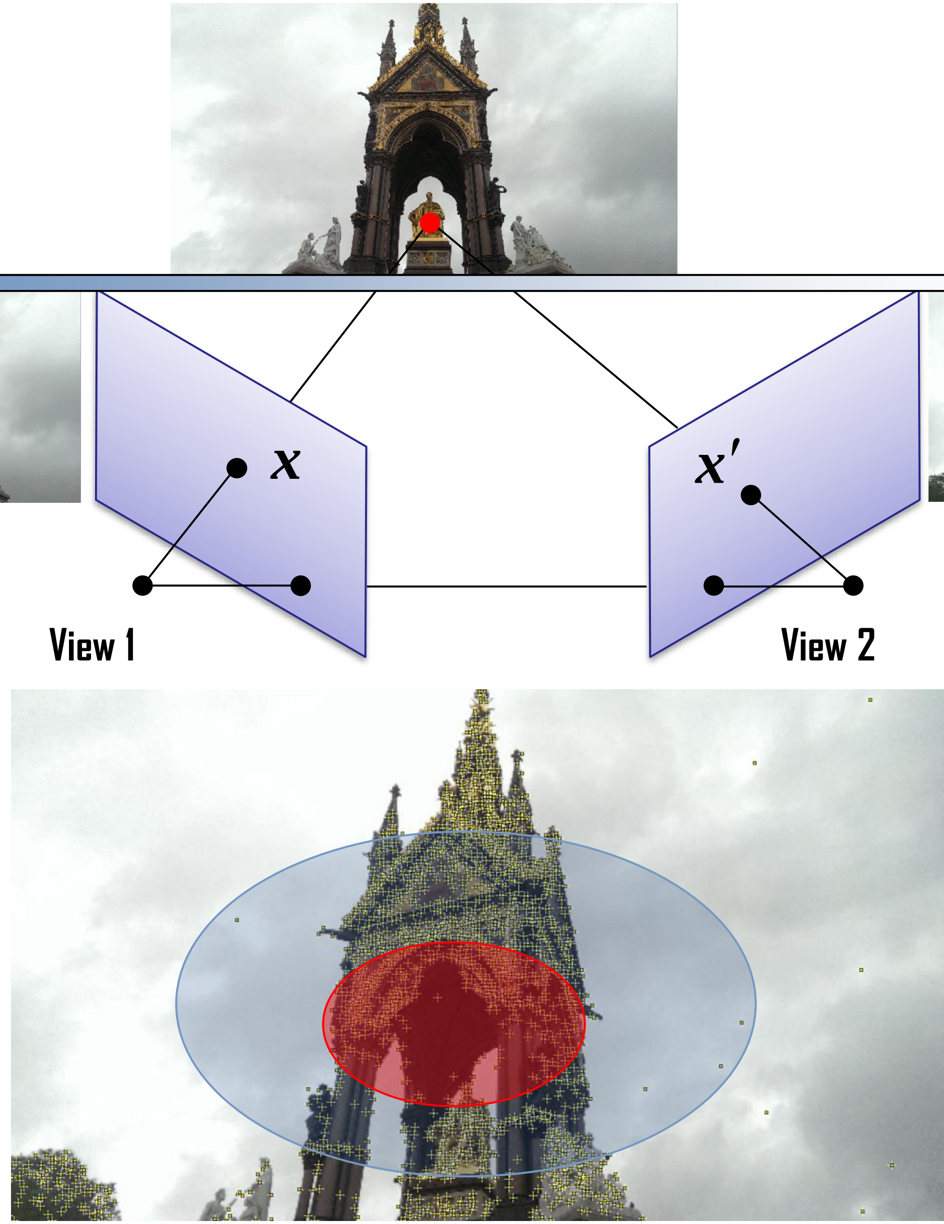
Why Stereo?

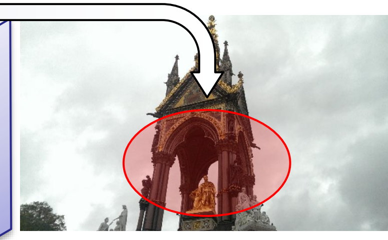
Why Stereo

Stereo Camera Geometry



When we find a feature, say, on the left image, we need to find the corresponding feature point on the right image, after that (hopefully) we can back-project the two rays and find the intersection (where the 3D point is located)

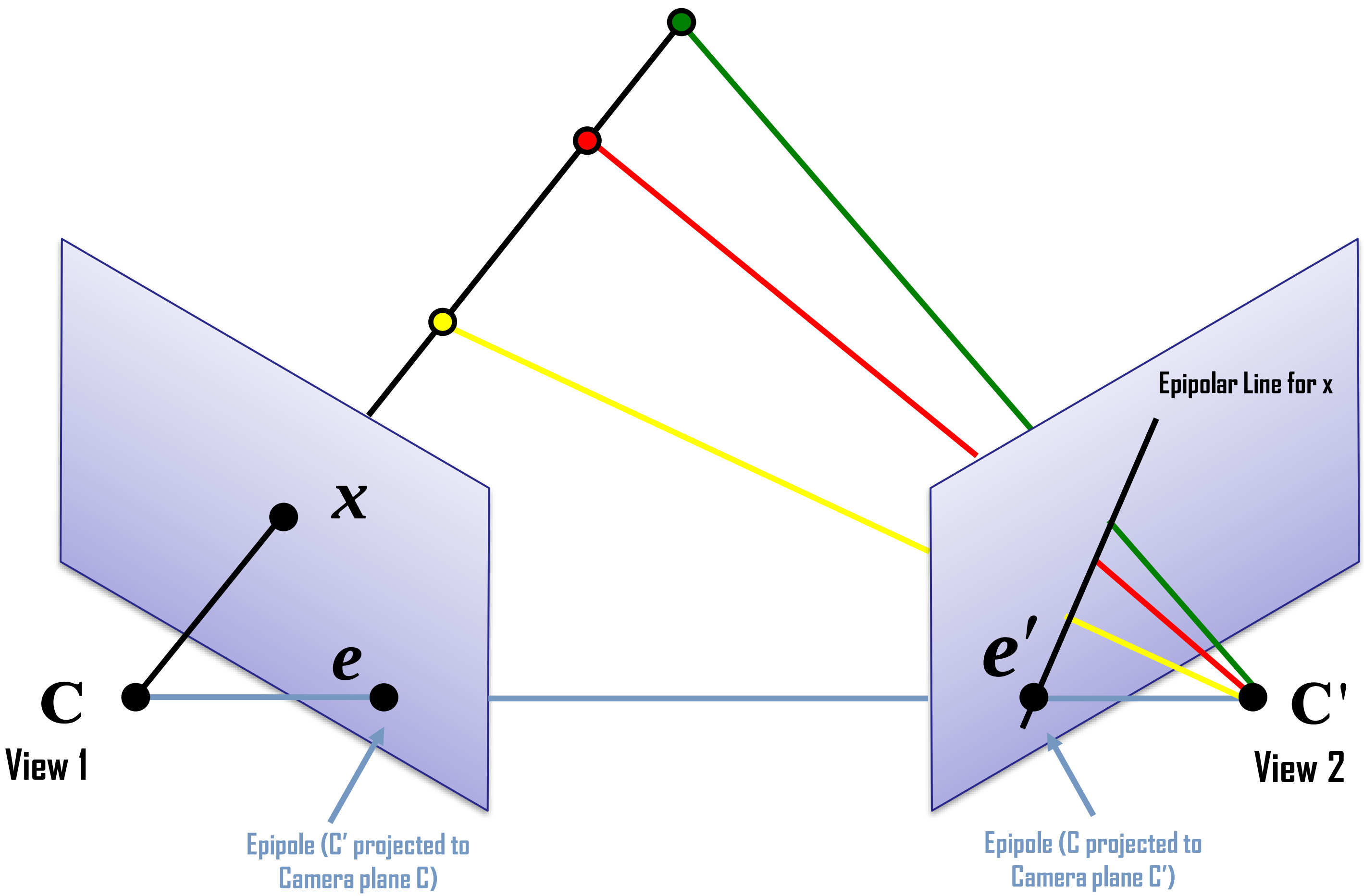




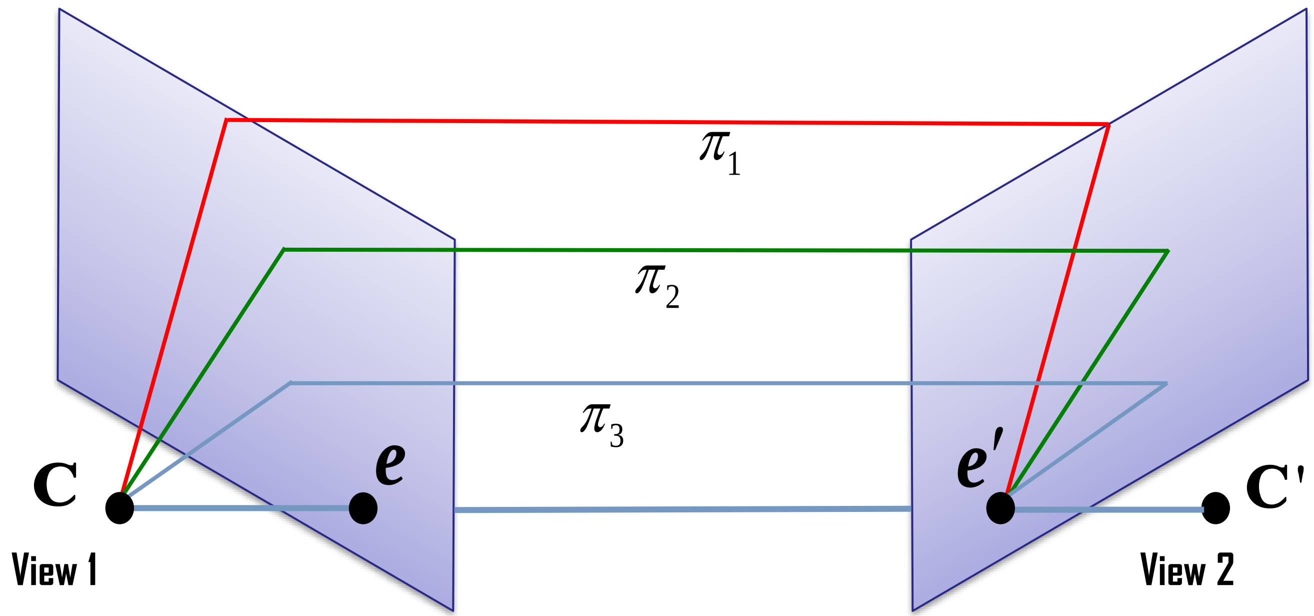
But, where should we look for corresponding feature points, there are many of them, how to define the search space, the entire image, or at the vicinity of the corresponding image coordinates, what happens the two views are from very different directions also with large translation?

Epipolar Lines

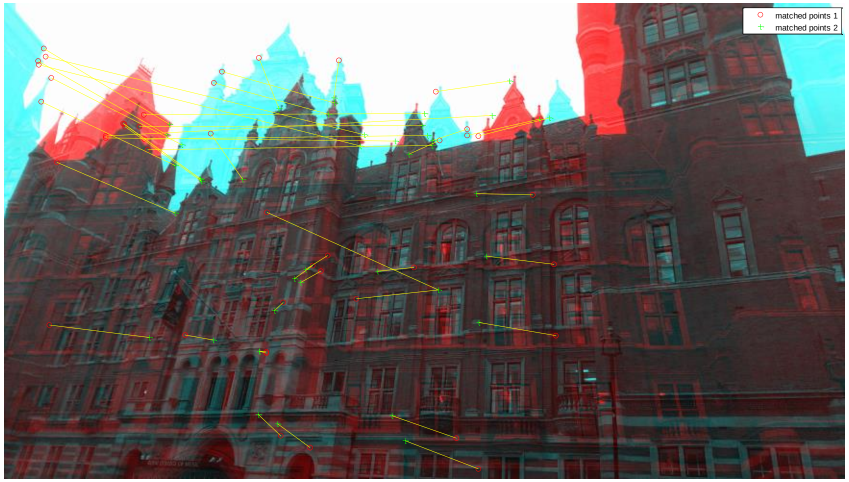
- If we look at the projection geometry, however, things are easier than that, if we look at several points on the same ray, what's the relationship can we find?



Epipolar Lines

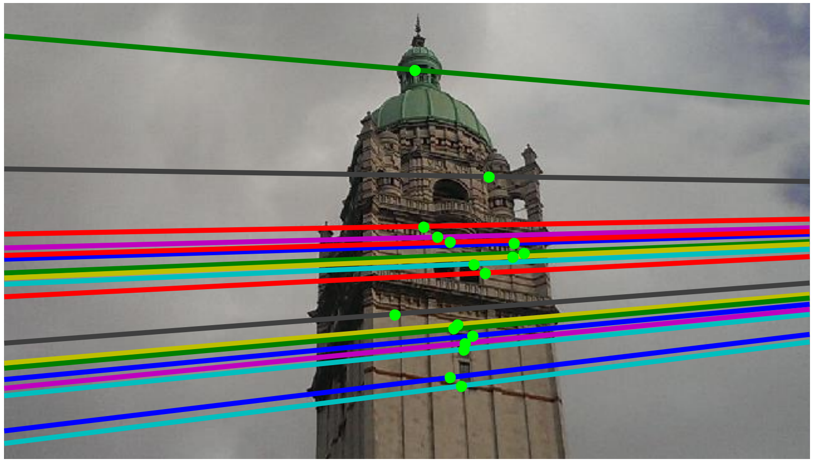


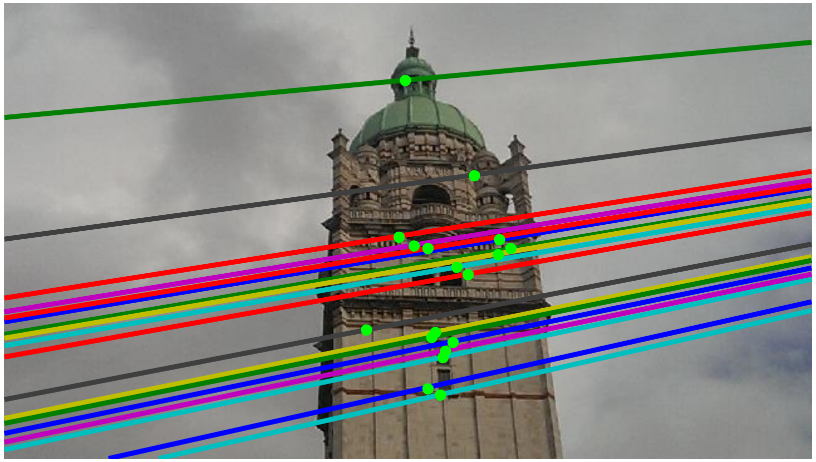
- So for a given feature point on one image plane (say on the left image), how to calculate the corresponding epipolar line on the other image plane (right image)?
- Conceptually, given a feature point **x**, if the relational position and orientation of the cameras are known, the projection plane must pass **C**, **C'** and **x**, and therefore can be easily defined
- The intersection of the projection plane π with the other image defines the epipolar line



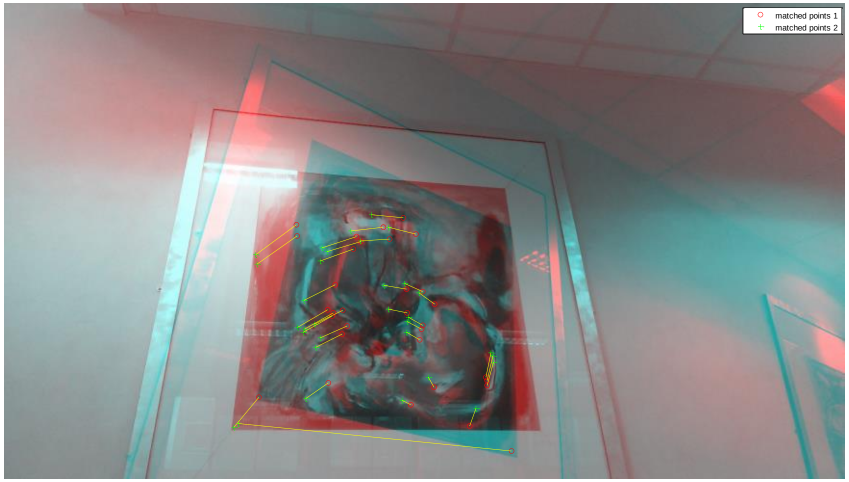
matched points 1
matched points 2

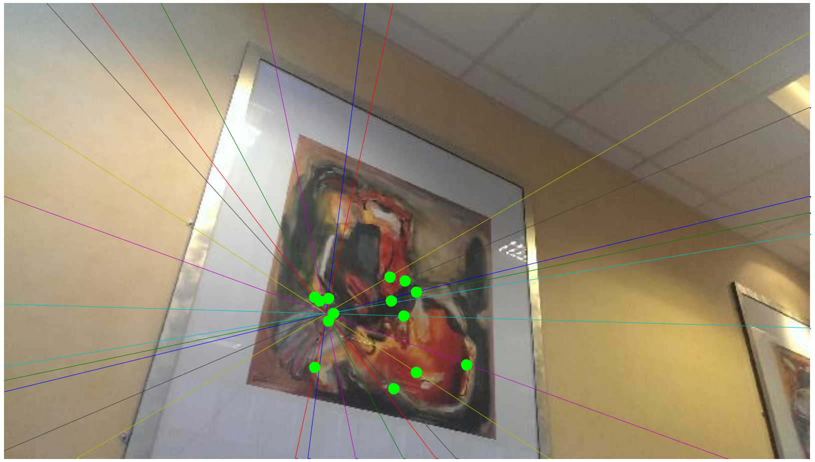

matched points 1
matched points 2

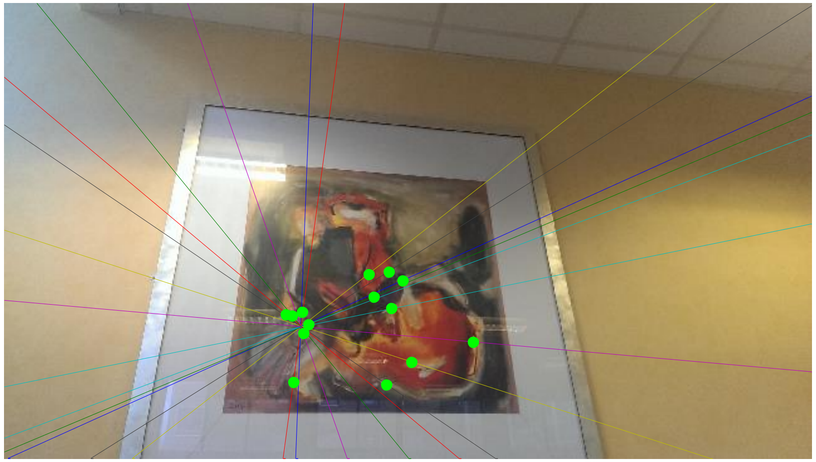




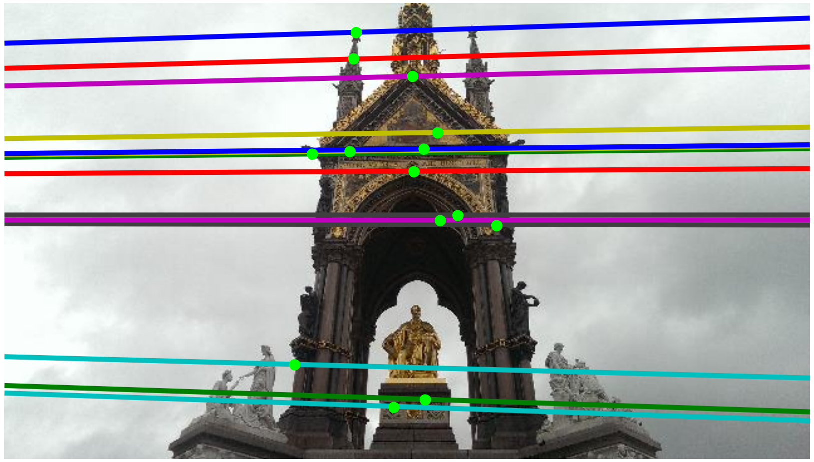
Epipolar Lines - Examples



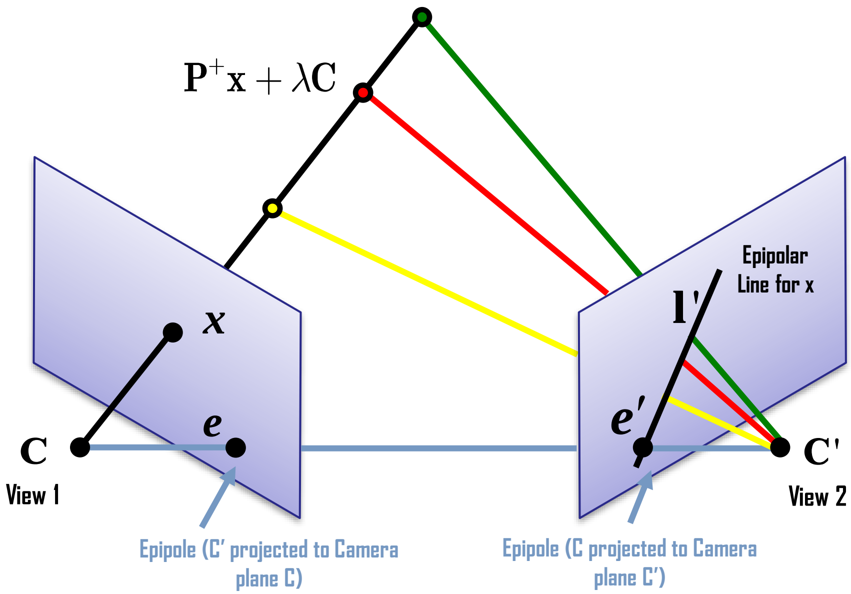





matched points 1
matched points 2



Fundamental Matrix



Note all vectors here are represented in homogeneous form and for cross product, we can use matrix manipulation (skew symmetric matrix)

$$\mathbf{a} \times \mathbf{b} = \left[\mathbf{a}\right]_{\times} \mathbf{b} = \left[\mathbf{b}\right]_{\times}^{T} \mathbf{a}$$

$$\left[\mathbf{a}\right]_{\times} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}$$

$$\mathbf{F} = \mathbf{e}' \times (\mathbf{P}'\mathbf{P}^{+}) = \left[\mathbf{e}'\right]_{\times} \mathbf{P}'\mathbf{P}^{+}$$

- Any point lies on the back-projected ray from C to feature point **x**

$$\mathbf{X}\left(\lambda\right) = \mathbf{P}^{+}\mathbf{x} + \lambda\mathbf{C}$$

- Two known points projected to C' coordinate system **e'** and **x'** (you can choose any point on the ray (other than C of course), e.g., choose λ=0)

$$\begin{aligned} \mathbf{x}' &= \mathbf{P}'(\mathbf{X}(\lambda)) \\ &= \mathbf{P}'(\mathbf{P}^{+}\mathbf{x} + \lambda\mathbf{C}) = \mathbf{P}'\mathbf{P}^{+}\mathbf{x} \\ \mathbf{e}' &= \mathbf{P}'\mathbf{C} \end{aligned}$$

- Two points in C' space define line **l'**

$$\mathbf{l}' = \mathbf{e}' \times (\mathbf{P}'\mathbf{P}^{+})\mathbf{x}$$

- Which yields the fundamental matrix

$$\mathbf{F} = \mathbf{e}' \times (\mathbf{P}'\mathbf{P}^{+})$$

Fundamental Matrix

- The fundamental matrix satisfies the condition that for any pair of corresponding points **x**↔**x**' in the two images

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

- This is because if **x** and **x'** correspond, then **x'** lies on the epipolar line **l'=Fx** or, in other words

$$\mathbf{x}' \bullet (\mathbf{F}\mathbf{x}) = 0, \quad \text{or} \quad \mathbf{x}'^{T}\mathbf{F}\mathbf{x} = 0$$

- Conversely, if the image points satisfy the above condition, then the rays defined by these points are co-planar, which is a necessary condition for points to be in correspondence.

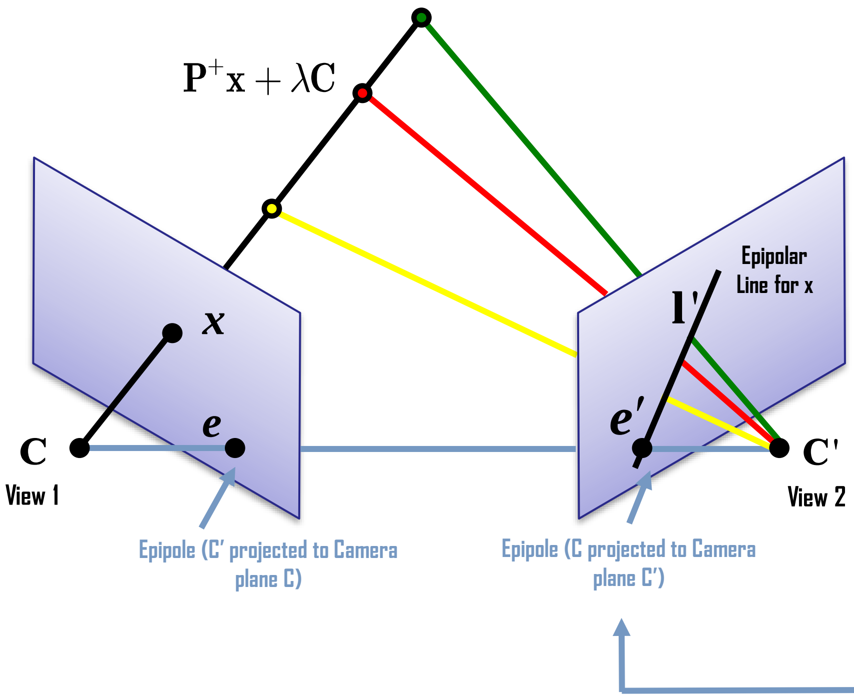
Fundamental Matrix

$$F = \begin{bmatrix} a & b & \alpha a + \beta b \\ c & d & \alpha c + \beta d \\ e & f & \alpha e + \beta f \end{bmatrix}$$

- **F** has 7 degrees of freedom (3x3 – 1 (homogeneous) -1(rank 2))
- If **F** is the fundamental matrix for camera pair (**P, P'**), then $\mathbf{F}^T$ is for (**P', P**), and $\mathbf{l}=\mathbf{F}^T\mathbf{x}'$ represents the epipolar line for **x'** in the second image
- For any point **x** except **e**, the epipolar line **l'=Fx** contains the epipole **e'** , therefore

$$\mathbf{e}'^T \mathbf{F}\mathbf{x} = (\mathbf{e}'^T \mathbf{F})\mathbf{x} = 0 \text{ for all } \mathbf{x}, \text{ therefore } \mathbf{e}'^T \mathbf{F} = 0, \text{ or } \mathbf{F}^T\mathbf{e}' = 0, \text{ similarly } \mathbf{F}\mathbf{e} = 0$$

Fundamental Matrix - Example



- Let’s put the left camera to coincide with the world coordinates

$$\mathbf{P} = \mathbf{K}[\mathbf{I} \mid 0], \quad \mathbf{P}' = \mathbf{K}'[\mathbf{R} \mid \mathbf{t}]$$

$$\mathbf{P}^{+} = \begin{bmatrix} \mathbf{K}^{-1} \\ 0 \end{bmatrix} \text{ and } \mathbf{C} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- So the fundamental matrix can be derived as

$$\begin{aligned} \mathbf{F} &= \mathbf{e}' \times (\mathbf{P}'\mathbf{P}^{+}) \\ &= [\mathbf{P}'\mathbf{C}]_{\times} \mathbf{P}'\mathbf{P}^{+} \\ &= [\mathbf{K}'\mathbf{t}]_{\times} \mathbf{K}'\mathbf{R}\mathbf{K}^{-1} \end{aligned}$$

Essential Matrix

- When the calibration matrix K is known, we can apply the inverse

$$\hat{\mathbf{x}} = \mathbf{K}^{-1}\mathbf{x} \text{ then } \hat{\mathbf{x}} = [\mathbf{R} \mid \mathbf{t}]\mathbf{X}$$

- This effectively removes all the intrinsic calibration factors and is called a normalised camera matrix. In this case, we can write

$$\mathbf{P} = [\mathbf{I} \mid 0], \quad \mathbf{P}' = [\mathbf{R} \mid \mathbf{t}]$$

- The fundamental matrix corresponding to the pair of normalised cameras is called **the Essential Matrix**, which can be written as

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

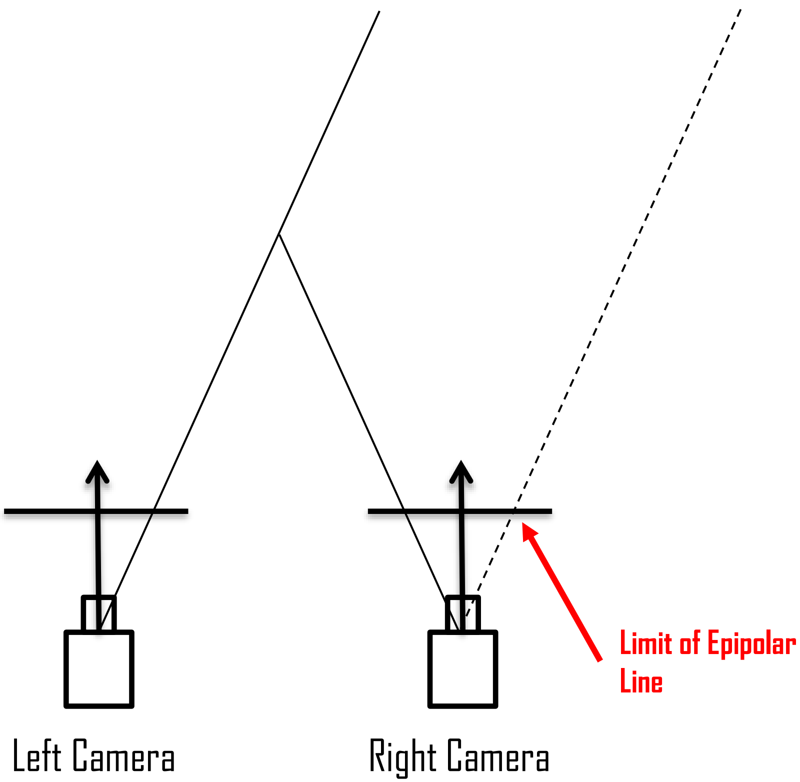
- The defining equation for the essential matrix is

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

Remember the general representation of the fundamental matrix? Basically now you can consider K and K' are identity matrices

$$\begin{aligned}\mathbf{F} &= \mathbf{e}' \times (\mathbf{P}'\mathbf{P}^+) \\ &= [\mathbf{P}'\mathbf{C}]_\times \mathbf{P}'\mathbf{P}^+ \\ &= [\mathbf{K}'\mathbf{t}]_\times \mathbf{K}'\mathbf{R}\mathbf{K}^{-1}\end{aligned}$$

Essential Matrix - Example



- Cameras have the same intrinsic parameters and are in correspondence (pointing at the same direction, same height and translate only horizontally by t_x

$$\mathbf{P} = [\mathbf{I} \mid 0], \quad \mathbf{P}' = [\mathbf{I} \mid \mathbf{t}]$$

$$\mathbf{R} = \mathbf{I} \quad \text{and } \mathbf{t} = \begin{bmatrix} t_x & 0 & 0 \end{bmatrix}^T$$

$$\mathbf{E} = \left[\mathbf{t}\right]_\times \mathbf{R} = \left[\mathbf{t}\right]_\times = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix}$$

$$\mathbf{x}'^T \mathbf{E} \mathbf{x} = \begin{bmatrix} u' & v' & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0$$

$$v' t_x - v t_x = 0 \text{ therefore } v = v'$$

The epipolar lines are all horizontal and of equal height in the image planes (in fact there is a limit of the epipolar lines, i.e., they don't cover the entire width of the image plane).

Epipolar Lines - Examples

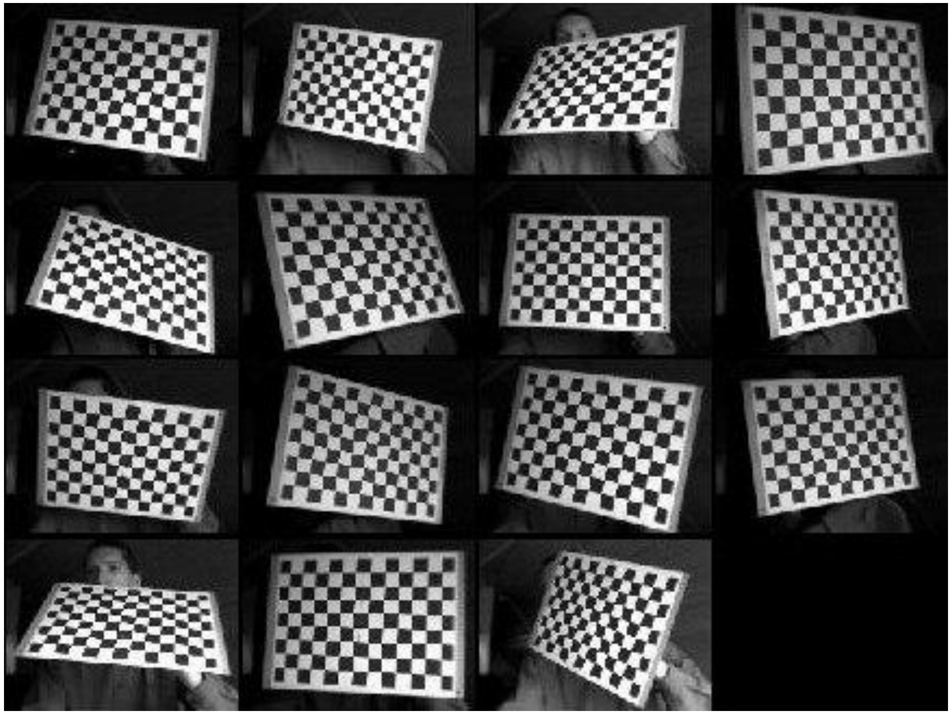
- The epipole is a fixed point and has the same coordinates in both images
- Points appear to move along lines radiating from the epipole
- In the case of pure translation **T**, 3D points move on straight lines parallel to **T**, and the imaged intersection of these parallel lines is the vanishing point **v** in the direction of **T**. It is evident that **v** is the **epipole** for both views, and the imaged parallel lines are the **epipolar lines**.

Camera Calibration

- Intrinsic camera parameters: the principle point, focal length, scaling factors for pixels, skew factor, and distortion factors

$$K=\begin{bmatrix}\alpha_x & s & x_0\\ 0 & \alpha_y & y_0\\ 0 & 0 & 1\end{bmatrix}$$

- 2D checkerboard patterns are commonly used for camera calibration
- The corners of the squares are used as calibration points with known 2D to 3D correspondences



$$\begin{bmatrix}u\\ v\\ w\end{bmatrix} = K[R\ T]\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}$$

http://www.vision.caltech.edu/bouguetj/calib_doc/

Z. Zhang: A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, 2000.

Conclusions

- Camera geometry and transformations
- Camera projection matrix
- Stereo camera geometry
- Epipolar lines and examples
- The fundamental matrix
- The essential matrix
- Camera calibration