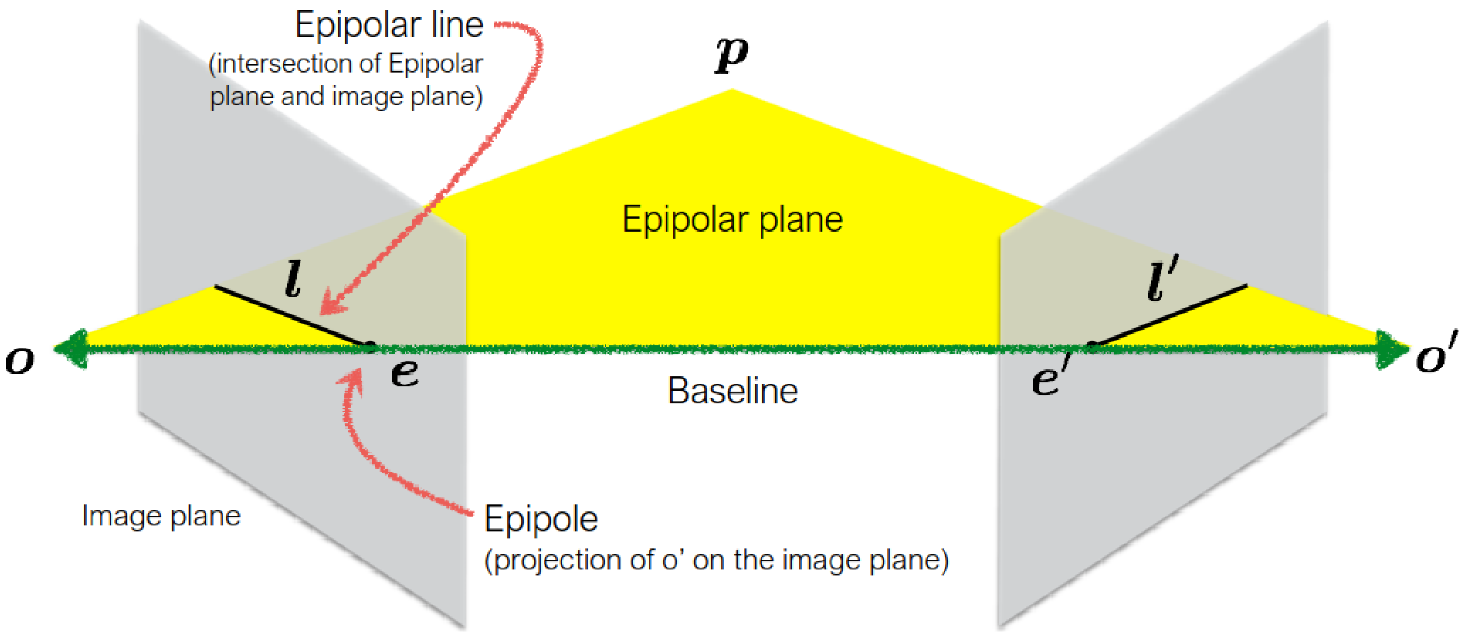
Epipolar Geometry



- For parallel cameras, the epipole is at infinity.

Essential Matrix

- The essential matrix $\mathbf{E}$ is a 3x3 matrix that describes the epipolar geometry between two views.
- It relates corresponding points in two images of a scene taken from two different camera positions.
- Given a point in one image, multiplying by the essential matrix will tell us the epipolar line in the second image
- It works on points on which extrinsic parameters are applied

Properties

- $\mathbf{E}$ is singular, i.e., $\det(\mathbf{E}) = 0$, and has 5 degrees of freedom, and rank 2.
- $(x')^\top l' = 0$, $x^\top l = 0$
- $l' = \mathbf{E}x$, $l = \mathbf{E}x'$
- $(e')^\top \mathbf{E} = 0$, $\mathbf{E}e = 0$
- $(x')^\top \mathbf{E}x = 0$ for corresponding points x and x'. (the epipolar constraint)
- $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$, where $\mathbf{R}$ is the rotation matrix and $[\mathbf{t}]_\times$ is the skew-symmetric matrix of the translation vector $\mathbf{t}$.

$$[\mathbf{t}]_\times = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \text{ where } \mathbf{t} = \begin{bmatrix} t_1 & t_2 & t_3 \end{bmatrix}^\top$$

Fundamental Matrix

- The fundamental matrix is a generalization of the essential matrix, where the assumption of calibrated cameras is removed
- It has 7 degrees of freedom, and is also singular in rank 2.
- All properties of the essential matrix hold for the fundamental matrix as well
- $\mathbf{F} = (\mathbf{K}')^{-\top}\mathbf{E}\mathbf{K}^{-1}$
- It works on points on which intrinsic parameters are applied (image pixels)

Direct Linear Transform (DLT) or Normalized 8-point Algorithm

- Given 8 or more corresponding points, we can estimate the fundamental matrix using the DLT algorithm.
- The DLT algorithm involves solving a linear system of equations to estimate the fundamental matrix.

$$\begin{bmatrix} x'_m & y'_m & 1 \end{bmatrix} \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ 1 \end{bmatrix} = 0$$

$$
\begin{bmatrix}
x_1x'_1 & x_1y'_1 & x_1 & y_1x'_1 & y_1y'_1 & y_1 & x'_1 & y'_1 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_mx'_m & x_my'_m & x_m & y_mx'_m & y_my'_m & y_m & x'_m & y'_m & 1
\end{bmatrix}
\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = 0
$$

$$
\mathbf{AX} = \mathbf{0}
$$

Algorithm Steps

- Normalize the points (since we are working in image coordinates)
  - Let the matrix used for normalizing the points in both cameras be $\mathbf{T}$ and $\mathbf{T}'$
- Construct the matrix $\mathbf{A}$
- Find SVD of $\mathbf{A}$, and using that, find approximate value $\hat{\mathbf{F}}$
- To find $\mathbf{F}$, enforce the rank-2 constraint on $\hat{\mathbf{F}}$ using SVD

$$
\min_F ||F - \hat{F}||_F, \text{subject to } \det F = 0
$$

$$
F = U \begin{bmatrix} \Sigma_1 & 0 & 0 \\ 0 & \Sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^\top
$$

- Denormalize the fundamental matrix to get the final fundamental matrix

$$
F_{\text{final}} = (\mathbf{T}')^\top F \mathbf{T}
$$