

# Epipolar Geometry

## Three Questions

### 1. Correspondence Geometry

- Given an image point  $x$  in the first view, how does this constrain the position of the corresponding point  $x'$  in the second image?

### 2. Camera Geometry (motion)

- Given a set of corresponding image points  $x_i \leftrightarrow x'_i, i = 1, \dots, n$
- What are the cameras  $P$  and  $P'$  for the two views?
- What is the geometric transformation between the two views?

### 3. Scene Geometry (structure)

- Given corresponding image points  $x_i \leftrightarrow x'_i$  and cameras  $P, P'$
- What is the position of the point  $X$  in space?

## Geometry

- Epipolar plane  $\pi$  is the plane formed by  $P, P', X$
- Plane passes through camera planes and points  $x, x'$  lie on the edges of the plane
- Epipoles - Intersection of baseline with image plane
  - Projection of projection center in other image
  - Vanishing point of camera motion direction
- Baseline - Line formed by  $\pi$  and the epipoles  $\rightarrow$  Translation between cameras
- Cross Product

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

- Triple Scalar Product  $\rightarrow$  3 points are coplanar

$$(a \times b) \cdot c = 0$$

- Equations

$$P' = RP + t$$

$$p = MP, M = [I|0]$$

$$p' = M'P', M' = [R|t]$$

P is the point in space, p, p' are points in image

$$\vec{O'p'}, \vec{OO'}, \vec{Op} \text{ are coplanar} \implies p' \cdot [t \times (Rp)] = 0, \begin{cases} p = (u, v, 1)^T \\ p' = (u', v', 1)^T \end{cases}$$

$$p' \cdot [t \times (Rp)] = 0 \implies p' \cdot [(t \times R)p] = 0 \implies p'([t_x]R)p = 0$$

$$\underbrace{E = [t_x]R}_{\text{Essential Matrix}} \implies \boxed{p'Ep = 0}$$

- Essential Matrix
  - E can be used to find rotation and translation
  - Computed from point correspondences
- Uncalibrated camera
  - $\hat{p} = M^{-1}p$  and  $\hat{p}' = M'_{int}p'$
  - p and p' are points in pixel coordinates corresponding to  $\hat{p}, \hat{p}'$
  - $p'^T F p = 0$
  - $F = M'^{-T}_{int} E M^{-1}_{int}$
- Properties of fundamental and essential matrix
  - With calibration parameters  $\rightarrow$  essential
  - Without  $\rightarrow$  fundamental
  - Matrix is 3x3
  - **Transpose:** If F is essential matrix of cameras P, P'  $\Rightarrow F^T$  is essential matrix of camera P', P
  - **Epipolar Lines:** p and p' as points in the projective plane  $\Rightarrow Fp$  is projective line in the right image
    - $l' = Fp, l = F^T p'$
  - **Epipole:** For any p the epipolar line  $l' = Fp$  contains the epipole  $e'$   $\Rightarrow (e'^T F)p = 0$  for all p
    - $e'^T F = 0$  and  $Fe = 0$
    - Epipoles are in the Null Space of the Essential Matrix

- Fundamental Matrix
  - Encodes information of the intrinsic and extrinsic parameters
  - F is of rank 2 since S has rank 2 (R and M and M' have full rank)
  - Has 7 DOF
    - 9 elements  $\rightarrow$  scaling is not significant &  $\det(F) = 0$
- Essential Matrix
  - Encodes only extrinsic parameters
  - Rank 2 since S has rank 2
  - Its two nonzero singular values are equal
  - Has only 5 DOF  $\rightarrow$  3 rotation, 2 translation
- Scaling Ambiguity

$$P' = RP + t$$

$$p = \frac{P}{z^T P} \quad p' = \frac{RP + t}{z^T (RP + t)}$$

- Depth Z and Z' and t can only be recovered to a scale factor
- Computing the Fundamental Matrix from Point Correspondences
  - Defined by the equation  $x_i'^T F x_i = 0$
  - For  $n$  point matches  $\rightarrow$  A is an  $n \times 9$  matrix
    - Solve  $Af = 0$
  - Homogeneous set of equations
    - f can be determined only up to a scale, so there are 8 unknowns
    - 8 Point Algorithm
  - Least Squares solution is the singular vector corresponding the smallest singular value of A  $\rightarrow$  last column of V in the SVD
    - This has a lot of error since the matrix is almost singular
  - Non-Linear Least Squares Approach
- Rectification
  - Image Reprojection
    - Reproject image planes onto common plane parallel to line between optical centers
  - Rotate left camera so epipoles go to infinity along the horizontal axis
  - Apply same rotation to the right camera
  - Rotate the right camera by  $R$
  - Adjust the scale
- 3D Reconstruction

- **Stereo:** We know the viewing geometry (extrinsic parameters) and the intrinsic parameters  $\Rightarrow$  Find correspondences exploiting epipolar geometry, then reconstruct
- **Structure from motion:** Find correspondences  $\rightarrow$  estimate extrinsic parameters (rotation and translation)  $\rightarrow$  reconstruct
- **Uncalibrated Cameras:** Find correspondences  $\rightarrow$  compute projection matrices  $\rightarrow$  reconstruct up to a projective transformation
  - Don't know the position of camera respective to origin
  - Can't find the orientation of camera plane
- **Triangulation**
  - If cameras are intrinsically and extrinsically calibrated  $\Rightarrow$  find P as the midpoint of the common perpendicular to the two rays in space
- Intrinsically calibrated cameras
  - Compute the essential matrix E using normalized points
  - Select  $M = [I|0], M' = [R|T] \rightarrow E = [T_x]R$
  - Find  $T$  and  $R$  using SVD of E

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

$$[T_x] = UZU^T \quad R = UWV^T \text{ or } R = UW^TV^T$$

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Multiple solutions  $\rightarrow$  One with positive values is correct
- Linear Triangulation
  - Homogenous System

$$x = MX \quad x' = M'X$$

$$x \times MX = 0 \quad x' \times M'X' = 0$$

$$A = \begin{bmatrix} xm_3^T & -m_1^T \\ ym_3^T & -m_2^T \\ x'm_3^T & -m_3'^T \\ y'm_3^T & -m_2'^T \end{bmatrix}$$

$$AX = 0$$

- $X$  is the last column of  $V$  in the SVD of  $A$
- Projective Reconstruction Theorem
  - Assume we determine matching points  $x_i$  and  $x'_i$ . Then we can compute a unique Fundamental matrix  $F$
  - The camera matrices  $M, M'$  cannot be recovered uniquely
  - Thus the reconstruction  $X_i$  is not unique
  - There exists a projective transformation  $H$
  - Reconstructive Ambiguity

$$x_i = MX_i = (MH_P^{-1})(H_P X_i)$$

- Two different objects produce the same image since the cameras are uncalibrated
  - Find the original distorted by a projective transformation  $H \rightarrow$  *Projective Reconstruction*
  - If reconstruction is derived from real images  $\rightarrow$  there is a true reconstruction that can produce the actual points  $X_i$  of the scene
- Consequences
  - We can compute a projective reconstruction of a scene from 2 views based on image correspondences alone
  - Don't need anything about poses or calibration
- Projective  $\rightarrow$  Metric
  - Compute homography  $H$  such that  $X_{Ei} = HX_i$  for 5 or more control points  $X_{Ei}$
  - Rectification using 5 points
    - Compute projective reconstruction
    - Find 5+ control points
    - Use them to solve for projective transformation (9 DOF) and rectify the projective reconstruction
- Stratified Reconstruction
  - Begin with projective reconstruction
  - Refine it to affine reconstruction
    - Fix plane to infinity
    - Parallel lines are parallel, ratios along parallel lines are correct
    - Reconstructed scene is then an affine transformation of the actual scene
  - Refine to a metric reconstruction
    - Use orthogonality

- Angles and ratios are correct
- Reconstructed scene is then a scaled version of the actual scene
- Reconstruction from N views
  - Projective or affine reconstruction from large set of images
  - Given a set of camera  $M^i$
  - For each camera  $M^i \rightarrow$  set of image point  $x_j^i$
  - Find 3D points  $X_j$  and cameras  $M^i$ , such that  $M^i X_j = x_j^i$
- Bundle Adjustment
  - Solve following minimization problem
  - Find  $M^i$  and  $X_j$  that minimize

$$\sum_{i,j} d(M^i X_j, x_j^i)^2$$

- Levenberg Marquardt algorithm
- Problems
  - Many parameters  $\rightarrow$  11 per camera, 3 per 3D point
- Useful as final adjustment step for bundles of arrays