Department of Computer Engineering
University of California at Santa Cruz

Epipolar Geometry and the Eight-Point Algorithm

CMPE 264: Image Analysis and Computer Vision

Hai Tao

Department of Computer Engineering
University of California at Santa Cruz

Camera motion estimation and 3D reconstruction

- Problem statement: Given multiple images of a static scene and the camera parameters, recover the camera motion and the 3D structure of the scene
- The image feature points corresponding to a set of 3D points are first detected in these images. Then the camera motion and the 3D coordinates of these points are estimated based on multiple view geometry.
- We will focus on multiple view geometry first and then discuss how to find feature correspondences across images

Two view geometry

- Notations

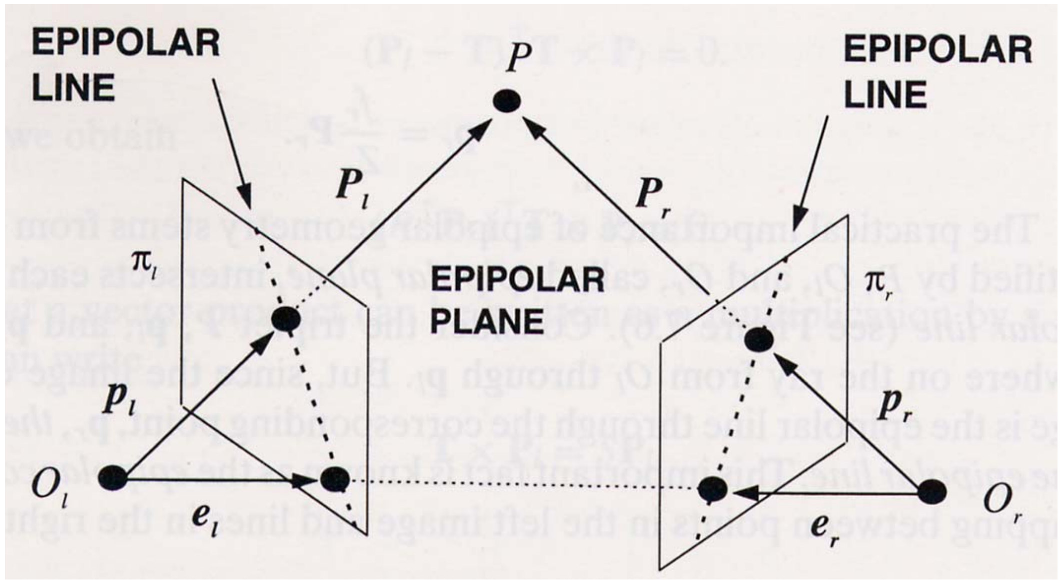
$P_l = [X_l, Y_l, Z_l]^T$ - the 3D coordinates of a point P in the left camera coordinate system

$P_r = [X_r, Y_r, Z_r]^T$ - the 3D coordinates of a point P in the right camera coordinate system

$p_l = [x_l, y_l, 1]^T$ - the coordinates of p in the left image

$p_r = [x_r, y_r, 1]^T$ - the coordinates of p in the right image

$P_r = R(P_l - T)$ - the transformation between the two camera coordinate systems



Two-view geometry

- Fundamental matrix

Since $\overrightarrow{O_l P}$, $\overrightarrow{O_l O_r}$, and $\overrightarrow{O_r P}$ are coplanar, therefore

$$(\overrightarrow{O_r P})_l^T (\overrightarrow{O_l O_r} \times \overrightarrow{O_l P}) = (\overrightarrow{O_r P})_l^T (T \times P_l) = 0$$

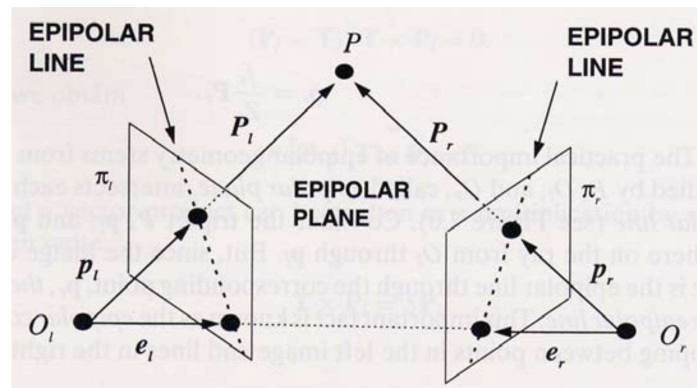
According to the transformation between the two camera frames

$$(\overrightarrow{O_r P})_l = (P_l - T) = (R^{-1} P_r + T - T) = R^T P_r$$

therefore

$$P_r^T R (T \times P_l) = P_r^T (R T_\times) P_l = 0$$

$$T_\times = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix}$$ so that $T \times P = T_\times P$



Since $K^{-1} p_l \cong P_l$ and $K^{-1} p_r \cong P_r$

$$P_r^T (R T_\times) P_l \cong p_r^T (K^{-T} R T_\times K^{-1}) p_l = \boxed{p_r^T F p_l = 0}$$

Fundamental matrix

- $F = K^{-T} R T_\times K^{-1}$ is called the *fundamental matrix*. It links correspondences in two images through a linear relation $p_r^T F p_l = 0$. $E = RT_\times$ is called the essential matrix
- Properties of the fundamental matrix
  - F is defined up to a scale factor and is rank 2 – because $T_\times$ is rank 2
  - If e_l is the image of O_r in the left image, then e_l is called the left epipole and vice versa
  - The plane formed by O_l, O_r, and P is called epipolar plane
  - The intersection of the epipolar plane and an images plane is called epipolar line.
  - All epipolar lines passing through the epipoles.
  - Except epipoles, only one epipolar line passing through each image point
  - Corresponding points must lie on conjugated epopolar lines
  - $Fe_l = 0$ and $F^T e_r = 0$ (Homework)
  - e_l is the null space of F
  - Fp_l is the epipolar line in the right image; $F^T p_r$ is the epipolar line in the left image (Homework)

Computing *F* and *E* – the eight-point algorithm

- Each pair of correspondences gives one constraint. To recover the fundamental matrix up to a scale factor, we need eight constraints, i.e. eight pairs of correspondences.
- For the *i*th correspondence

  $p_r^{iT} F p_l^i = 0 \Leftrightarrow$

  $[x_r^i x_l^i, x_r^i y_l^i, x_r^i, y_r^i x_l^i, y_r^i y_l^i, y_r^i, x_l^i, y_l^i, 1][f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33}]^T = 0$

- With $N \geq 8$ correspondences, we have a linear system

  $$A\mathbf{f} = 0$$

  where A is formed by stacking the data vectors as indicated above

  The solution is the eigenvector corresponding to the smallest eigenvalue of $A^T A$
- To compute the essential matrix, use the relation $E = K^T F K$
- To make the algorithm numerical stable, normalize image coordinates so that RMS is $\sqrt{2}$

From essential matrix to camera motion *R* and *T*

- Theorem 1. A 3 by 3 real matrix Q can be factored as the product of a rotation matrix and a non-zero skew symmetric matrix iff Q has two equal non-zero singular values and one singular value equal to 0
- Theorem 2. Suppose Q can be factored into a product *RS* where *R* is orthonormal and S is skew symmetric. Let the SVD of *Q* be $Q = UDV^T$. Then up to a scale factor the factorization is one of the following

$$S = VZV^T$$
$$R = UGV^T \ or \ R = UG^TV^T$$
$$Q = RS$$

where

$$Z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

since $S\mathbf{v}_3 = VZV^T\mathbf{v}_3 = [\mathbf{v}_2, -\mathbf{v}_1, 0][\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]^T \mathbf{v}_3 = \mathbf{v}_2\mathbf{v}_1^T\mathbf{v}_3 - \mathbf{v}_1\mathbf{v}_2^T\mathbf{v}_3 = 0$

therefore $T = \lambda\mathbf{v}_3 = \lambda[0,0,1]V$, where λ is an arbitrary scale factor

From essential matrix to camera motion *R* and *T*

- Algorithm From_E_to_RT
  - Find the singular value decomposition $E = UDV^T$
  - The (R,T) pair is one of the following four solutions

$$(UGV^T, \lambda V[0,0,1]^T)$$

$$(UGV^T, -\lambda V[0,0,1]^T)$$

$$(UG^TV^T, \lambda V[0,0,1]^T)$$

$$(UG^TV^T, -\lambda V[0,0,1]^T)$$

  - Based on the constraint that scene points are in front of both cameras, a single pair of correspondences can be used to find the correct solution

Department of Computer Engineering
University of California at Santa Cruz

Linear triangulation

- After estimating the projection matrix $M_l = K[I,0]$ and $M_r = K[R,-RT]$, 3D coordinates of feature points can be computed from their 2D projections
- Since

$$p_l \cong M_l P_l$$

$$p_r \cong M_r P_l$$

if we denote m_l^{iT} and m_r^{iT} as the rows of M_l and M_r, then

$$\begin{bmatrix} x_l m_l^{3T} - m_l^{1T} \\ y_l m_l^{3T} - m_l^{2T} \\ x_r m_r^{3T} - m_r^{1T} \\ y_r m_r^{3T} - m_r^{2T} \end{bmatrix} P = AP = 0$$

The solution is the eigenvector corresponding to the smallest eigenvalue of of $A^T A$

Bundle Adjustment

- Gradient methods such as the Levenberg-Marquardt method is then employed to refined the result by minimizing the "reprojection error"

$$\min_{R,\mathbf{T},\hat{P}_{j=1,\ldots M}} \sum_{i=1}^{2}\sum_{j=1}^{M}\left\| p_{ij} - \hat{p}_{ij} \right\|^2$$