# Problem Set 7 Solutions

*Computer Vision 2020*
*University of Bern*

## 1 Epipolar Geometry

1. In this exercise we will derive the essential and fundamental matrices for a pair of cameras related by a rigid transformation $[\mathbf{R}|\mathbf{t}]$, by looking at the common 3D point $\mathcal{P}$.

   Let $\mathbf{X}_1 = [X_1\ Y_1\ Z_1]^T \in R^3$ be the coordinates of $\mathcal{P}$ defined in the first camera coordinate system with origin in the first camera center $C_1$. Let $\mathbf{R}$ be the rotation matrix and $\mathbf{t}$ be the translation vector that transform the first camera coordinate system into the second camera coordinate system with origin $C_2$. Remember that $\mathbf{R}$ is a $3 \times 3$ orthogonal matrix with $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ and $\mathbf{t} \in R^3$. Finally, let the coordinates of the 3D point $\mathcal{P}$ in the second camera coordinate system be $\mathbf{X}_2 = [X_2\ Y_2\ Z_2]^T \in R^3$.

   (a) How do you compute the projection $\mathbf{m}_1 = [x_1\ y_1\ 1]^T$ of $\mathcal{P}$ in the first coordinate system using the coordinate of $\mathbf{X}_1$ and how do you compute the projection $\mathbf{m}_2 = [x_2\ y_2\ 1]^T$ of $\mathcal{P}$ in the second coordinate system using the coordinate of $\mathbf{X}_2$?

   **Solution** $\mathbf{m}_1 = \mathbf{X}_1/Z_1$ and $\mathbf{m}_2 = \mathbf{X}_2/Z_2$.

   (b) Give the relationship between $\mathbf{X}_1$ and $\mathbf{X}_2$ using $\mathbf{R}$ and $\mathbf{t}$.

   **Solution** $\mathbf{X}_2 = \mathbf{R}\mathbf{X}_1 + \mathbf{t}$.

   (c) How are the projections $\mathbf{m}_1$ and $\mathbf{m}_2$ related (always taking the first image plane as coordinate system)?

   **Solution** $Z_2\mathbf{m}_2 = Z_1\mathbf{R}\mathbf{m}_1 + \mathbf{t}$.

   (d) Suppose that $\mathcal{P}$ lies on a plane $\mathcal{P} \in \pi$, where $\pi$ has a normal vector $\mathbf{n}_1$ (expressed in the first camera coordinate system) and its distance to $C_1$ is $d_1$. Assuming that $\mathbf{n}_1$ verifies $\|\mathbf{n}_1\| = \sqrt{\mathbf{n}_1^T\mathbf{n}_1} = 1$, give the equation of the plane $\pi$.

**Solution** We can use the Hessian normal form of the equation of a plane, that is $\mathbf{n}_1^T\mathbf{X} = d_1$.

(e) Use the results above, including $\mathcal{P} \in \pi$, to find a transformation of the form $\mathbf{m}_2 = T\mathbf{m}_1$.

**Solution**

$$\begin{aligned}
Z_2\mathbf{m}_2 &= Z_1\mathbf{R}\mathbf{m}_1 + \mathbf{t} \\
Z_2\mathbf{m}_2 &= Z_1\mathbf{R}\mathbf{m}_1 + \mathbf{t}\frac{\mathbf{n}_1^T\mathbf{X}_1}{d_1} \\
Z_2\mathbf{m}_2 &= Z_1\mathbf{R}\mathbf{m}_1 + Z_1\mathbf{t}\frac{\mathbf{n}_1^T\mathbf{m}_1}{d_1} \\
\mathbf{m}_2 &= \frac{Z_1}{Z_2}\left(\mathbf{R} + \mathbf{t}\frac{\mathbf{n}_1^T}{d_1}\right)\mathbf{m}_1
\end{aligned}$$

Given two (perspective) images of a world plane, the transformation relating points in two images which correspond to the same point $\mathbf{X}$ on the work plane is a homography.

(f) Given the normalized image coordinates $\mathbf{m}_1$ and $\mathbf{m}_2$ express the transformation of two points $\mathbf{p}_1$ and $\mathbf{p}_2$ in pixel coordinates with the help of the intrinsic matrix $\mathbf{K}$.

**Solution** We know that $\mathbf{p}_1 = \mathbf{K}_1\mathbf{m}_1$ and $\mathbf{p}_2 = \mathbf{K}_2\mathbf{m}_2$. Substituting this information we finally get a homograph mapping pixels to pixels:

$$\mathbf{p}_2 = \mathbf{K}_2\frac{Z_1}{Z_2}\left(\mathbf{R} + \mathbf{t}\frac{\mathbf{n}_1^T}{d_1}\right)\mathbf{K}_1^{-1}\mathbf{p}_1 \tag{1}$$

(g) Try to find a relationship between $\mathbf{m}_1$ and $\mathbf{m}_2$ without knowing the normal of the plane passing from $\mathcal{P}$.

**Solution** The normal vector of the plane can be derived from the cross product $\mathbf{t} \times \mathbf{X}_2 = [\mathbf{t}]_\times\mathbf{X}_2$, therefore we have

$$\begin{aligned}
\mathbf{X}_2 &= \mathbf{R}\mathbf{X}_1 + \mathbf{t} \\
[\mathbf{t}]_\times\mathbf{X}_2 &= [\mathbf{t}]_\times\mathbf{R}\mathbf{X}_1 + [\mathbf{t}]_\times\mathbf{t} \\
[\mathbf{t}]_\times\mathbf{X}_2 &= [\mathbf{t}]_\times\mathbf{R}\mathbf{X}_1
\end{aligned}$$

Since $\mathbf{X}_2$ lies on the plane, we have $\mathbf{X}_2^T[\mathbf{t}]_\times\mathbf{X}_2 = 0$, Replacing with the normalized image coordinates we have

$$\begin{aligned}
0 &= Z_1Z_2\mathbf{m}_2^T[\mathbf{t}]_\times\mathbf{R}\mathbf{m}_1 \\
0 &= \mathbf{m}_2^T\mathbf{E}\mathbf{m}_1
\end{aligned}$$

where $\mathbf{E}$ is the essential matrix.

(h) Give the relation between the points $\mathbf{p}_1$ and $\mathbf{p}_2$ in pixel coordinates.

**Solution** Using the intrinsic matrices $\mathbf{K}$ we have

$$\begin{aligned}
\mathbf{m}_2^T\mathbf{E}\mathbf{m}_1 &= 0\\
(\mathbf{K}_2^{-1}\mathbf{p}_2)^T\mathbf{E}(\mathbf{K}_1^{-1}\mathbf{p}_1) &= 0\\
\mathbf{p}_2^T(\mathbf{K}_2^{-T}\mathbf{E})\mathbf{K}_1^{-1}\mathbf{p}_1 &= 0\\
\mathbf{p}_2^T\mathbf{F}\mathbf{p}_1 &= 0
\end{aligned}$$

where $\mathbf{F}$ is the fundamental matrix.

2. The epipolar geometry is the intrinsic projective geometry between two views $I$ and $I'$. It depends only on the cameras' intrinsic parameters and their relative pose (rotation and translation between the camera centers). The **Fundamental Matrix F** is a $3 \times 3$ matrix with $\text{rank}(\mathbf{F}) = 2$.

(a) How is the fundamental matrix $\mathbf{F}$ related to pairs of corresponding points $x$, $x'$ in the two images?

**Solution** $\mathbf{x'}_i^T\mathbf{F}\mathbf{x}_i = 0$.

(b) How are the fundamental matrices $\mathbf{F}$, going from $I$ and $I'$, and $\mathbf{F'}$, going from $I'$ and $I$, related?

**Solution** $\mathbf{F'} = \mathbf{F}^T$.

(c) What is the geometric meaning of the epipoles $\mathbf{e}$ and $\mathbf{e'}$? How are they related to the fundamental matrix (algebraically)?

**Solution** Geometrically the epipoles are the points of the intersection of the line joining the camera centers (the baseline) with the image planes. Equivalently, an epipole is is the image in one view of the camera center of the other view. Algebraically, $\mathbf{e}$ is the right null space of $\mathbf{F}$, i.e. $\mathbf{F}\mathbf{e} = 0$. Similarly, $\mathbf{e'}$ is the left null space of $\mathbf{F}$, i.e. $\mathbf{e'}^T\mathbf{F} = 0$.

(d) Are the epipoles always visible in the two views?

**Solution** No, the intersection can happen outside the image or even at infinity, when the two image planes are parallel and the translation between their centers is perpendicular to the image planes.

(e) How can one determine the epipolar line $\mathbf{I'}$ passing trough a given point $\mathbf{x'}$?

**Solution** By finding the epipole and using $\mathbf{I'} = \mathbf{e'} \times \mathbf{x'}$.

(f) What is the effect of applying the fundamental matrix $\mathbf{F}$ to a point $\mathbf{x}$?

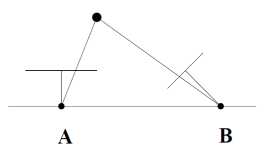
**Solution** The fundamental matrix maps a point in one image to a line in the second image. $\mathbf{I'} = \mathbf{Fx}$ and $\mathbf{I} = \mathbf{F}^T\mathbf{x'}$.

## 2 Calibrated Reconstruction

For a given essential matrix $E$ with singular value decomposition $E = U\mathrm{diag}(1,1,0)V^T$ and first camera matrix $P = [I \mid 0]$, there are four possible choices for the second camera matrix $P'$, namely

$$P' = [UWV^T \mid u_3] \quad \text{or} \quad [UWV^T \mid -u_3] \quad \text{or} \quad [UW^TV^T \mid u_3] \quad \text{or} \quad [UW^TV^T \mid -u_3],$$

where $W$ is a specific orthogonal matrix and $u_3$ is the third column of $U$.
The Figure below illustrates the geometric interpretation of the first solution. Draw the three missing configurations.




(a) (b)

(c) (d)

**Solution**
See page 260 of Hartley and Zisserman, Multiple View Geometry in Computer Vision, Second Edition, 2004.
`https://www.robots.ox.ac.uk/~vgg/hzbook/hzbook2/HZepipolar.pdf`
(a) points in front of both views; (b) points behind both A and B; (c) points in front of B but behind A; (d) points in front of A but behind B.
Author: Adrian Wälchli