Multiple View Geometry: Solution Sheet 6

Prof. Dr. Florian Bernard, Florian Hofherr, Tarun Yenamandra
Computer Vision Group, TU Munich
`Link Zoom Room , Password: 307238`

Exercise: June 4th, 2020

Part I: Theory

1. (a) E is essential matrix $\quad\Rightarrow\quad \Sigma = \text{diag}\{\sigma, \sigma, 0\}$:

$$R_z(\pm\frac{\pi}{2})\Sigma = \begin{pmatrix} 0 & \mp 1 & 0 \\ \pm 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mp\sigma & 0 \\ \pm\sigma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = -(R_z(\pm\frac{\pi}{2})\Sigma)^\top$$

$$\begin{aligned} -\hat{T}^\top &= -(UR_z\Sigma U^\top)^\top \\ &= U(-R_z\Sigma)^\top U^\top \\ &= UR_z\Sigma U^\top \\ &= \hat{T} \end{aligned}$$

(b) Since U, V are orthogonal with determinant 1 (see lecture), they are rotation matrices. Since $\text{SO}(3)$ is a group and thus closed under multiplication, $R \in \text{SO}(3)$.

Alternative longer proof:

i. U, V are orthogonal matrices $\Rightarrow U^\top U = \mathbb{1}$ and $VV^\top = \mathbb{1}$
R_z is a rotation matrix $\Rightarrow R_z R_z^\top = \mathbb{1}$

$$\begin{aligned} R^\top R &= (UR_z^\top V^\top)^\top (UR_z^\top V^\top) \\ &= VR_z U^\top U R_z^\top V^\top \\ &= VR_z R_z^\top V^\top \\ &= VV^\top \\ &= \mathbb{1} \end{aligned}$$

ii. U and V are special orthogonal matrices with $\det(U) = \det(V^\top) = 1$ (Slide 9, Chapter 5).

$$\det(R) = \det(UR_z^\top V^\top) = \underbrace{\det(U)}_{1} \cdot \underbrace{\det(R_z^\top)}_{1} \cdot \underbrace{\det(V^\top)}_{1} = 1$$

2. (a) $H = R + Tu^\top \Leftrightarrow R = H - Tu^\top$.

$$\begin{aligned} E &= \hat{T}R \\ &= \hat{T}(H - Tu^\top) \\ &= \hat{T}H - \underbrace{\hat{T}T}_{=T\times T=0} u^\top \\ &= \hat{T}H \end{aligned}$$

(b)

$$\begin{aligned}
H^\top E + E^\top H &= H^\top(\hat{T}H) + (\hat{T}H)^\top H \\
&= H^\top(\hat{T}H) + H^\top \hat{T}^\top H \\
&= H^\top \hat{T} H - H^\top \hat{T} H \quad \text{(because } \hat{T} \text{ is skew-symmetric, i.e. } \hat{T}^\top = -\hat{T}\text{)} \\
&= 0
\end{aligned}$$

3. The notations below are as in Slide 6, Chapter 5. Note that the following slides deal with projected points in the normalized plane ($Z = 1$), whereas here we assume pixel coordinates. The case of normalized coordinates is then just a special case with $K = \mathbb{1}$.

   Rotation R and translation T are defined such that

   $$g_{21} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$$

   transforms a point from coordinate system 1 (CS1) to coordinate system 2 (CS2). This means that the inverse transformation (converting points from CS2 to CS1) is given by

   $$g_{12} = g_{21}^{-1} = \begin{bmatrix} R^\top & -R^\top T \\ 0 & 1 \end{bmatrix}.$$

   o_1 seen in CS1: $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\top$ (homogeneous coordinates)

   o_1 seen in CS2: $g_{21} \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\top = \begin{bmatrix} T \\ 1 \end{bmatrix}$

   e_2 are the pixel coordinates of o_1 projected into image 2:

   $$\lambda_1 e_2 = K\Pi_0 \begin{bmatrix} T \\ 1 \end{bmatrix} = KT$$

   o_2 seen in CS2: $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\top$

   o_2 seen in CS1: $g_{12} \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\top = \begin{bmatrix} -R^\top T \\ 1 \end{bmatrix}$

   e_1 are the pixel coordinates of o_2 projected into image 1:

   $$\lambda_2 e_1 = K\Pi_0 \begin{bmatrix} -R^\top T \\ 1 \end{bmatrix} = -KR^\top T$$

   $$\begin{aligned}
   Fe_1 &= (\underbrace{K^{-\top}\hat{T}RK^{-1}}_{F})\underbrace{(-\frac{1}{\lambda_2}KR^\top T)}_{e_1} \\
   &= -\frac{1}{\lambda_2}K^{-\top}\hat{T}R\underbrace{K^{-1}K}_{\mathbb{1}}R^\top T \\
   &= -\frac{1}{\lambda_2}K^{-\top}\hat{T}\underbrace{RR^\top}_{\mathbb{1}}T \\
   &= -\frac{1}{\lambda_2}K^{-\top}\underbrace{\hat{T}T}_{=T\times T=0} \\
   &= 0
   \end{aligned}$$

$$
\begin{aligned}
e_2^\top F &= (\underbrace{\frac{1}{\lambda_1} KT}_{e_2})^\top (\underbrace{K^{-\top} \hat{T} R K^{-1}}_{F}) \\
&= \frac{1}{\lambda_1} T^\top \underbrace{K^\top K^{-\top}}_{\mathbb{1}} \hat{T} R K^{-1} \\
&= \frac{1}{\lambda_1} T^\top \hat{T} R K^{-1} \\
&= \frac{1}{\lambda_1} (\hat{T}^\top T)^\top R K^{-1} \\
&= \frac{1}{\lambda_1} (-\hat{T} T)^\top R K^{-1} \\
&= -\frac{1}{\lambda_1} (T \times T)^\top R K^{-1} \\
&= -\frac{1}{\lambda_1} 0 R K^{-1} \\
&= 0
\end{aligned}
$$

Part II: Practical Exercises

Remarks

1. **Are there two or four possible solutions for R and T?**
   The answer is four. You may find it confusing since in Slide 9 you are told to have two sets of R and T, while in Slide 14 it says there are four possible solutions. Recall that the essential matrix E is calculated by solving the equation $\chi E^s = 0$ (at the bottom of Slide 11). Ideally the E^s you get should lie in the nullspace of χ. However, since there are always erros in the point pairs you've chosen to make χ, in practice it is very difficult to get an E^s that makes χE^s exactly 0. Instead we use the SVD of χ to get the E^s which minimizes $||\chi E^s||$. In other words, the E^s we get will give us $\chi E^s = \sigma$ with σ being some very small vector. Now think about this: what will happen if we turn the sign of the E^s? We will get $\chi(-E^s) = -\sigma$ which still gives us the smallest $||\chi E^s||$ ($||\chi E^s|| = ||\sigma|| = || - \sigma|| = ||\chi(-E^s)||$).

   Now we know that solving $\chi E^s = 0$ for E^s will always give us two possible solutions E and $-E$. From each of them you can get two possible sets of R and T using the equations in Slide 9. Altogether we get four possible solutions. In practice we do the calculation in a little different way. We usually calculate according to Slide 14 to get four solutions out of E. The two extra solutions we get are nothing but the ones we should have got from $-E$.

2. **How to get the correct R and T from the possible solutions?**
   The criterion you should use to rule out the incorrect solutions is that all the reconstructed 3D points should have positive depth seen from **both of the camera coordinate systems**. In other words, bothλ_1^j and λ_2^j in Slide 17 need to be positive.