CSE344: Computer Vision

Assignment-2

Divyajeet Singh (2021529)

March 12, 2024

Question 1.

1. The given transformation is a composition of the following three transformations (in order):

   (a) Rotation by $\frac{\pi}{2}$ about the Y-axis

   (b) Rotation by $\frac{-\pi}{2}$ about the X-axis

   (c) Translation by $t = \begin{bmatrix}-1 & 3 & 2\end{bmatrix}^\top$

   The coordinate transformation matrices (using 3-dimensional homogeneous coordinates), for the three transformations are

$$R_y\left(\frac{\pi}{2}\right) = \begin{bmatrix} \cos\left(\frac{\pi}{2}\right) & 0 & \sin\left(\frac{\pi}{2}\right) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\left(\frac{\pi}{2}\right) & 0 & \cos\left(\frac{\pi}{2}\right) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_x\left(\frac{-\pi}{2}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\left(\frac{-\pi}{2}\right) & -\sin\left(\frac{-\pi}{2}\right) & 0 \\ 0 & \sin\left(\frac{-\pi}{2}\right) & \cos\left(\frac{-\pi}{2}\right) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T(t) = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

   The coordinate transformation matrix for the full transformation is then given by

$$\begin{aligned} T &= T(t) \circ R_x\left(\frac{-\pi}{2}\right) \circ R_y\left(\frac{\pi}{2}\right) \\ &= T(t) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 3 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

   since rotations and translations are linear transformations in homogeneous coordinate systems, and hence composition of transformations is equivalent to multiplication of the transformation matrices.

2. Now, we find the new coordinates of a given vector, $v = \begin{bmatrix}2 & 5 & 1\end{bmatrix}^\top$, after the transformation. The new coordinates, say v', are given by simply applying the transformation matrix to v in homogeneous coordinates, which gives

$$v' = Tv = \begin{bmatrix} 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 3 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -3 \\ 1 \end{bmatrix}$$

So, the new coordinates of the given vector are $v' = \begin{bmatrix} 0 & 1 & -3 \end{bmatrix}^\top$. We also find the point that the origin of the initial frame of reference gets mapped to. For this, we simply apply the transformation matrix T to the origin $\mathbf{0}$ in homogeneous coordinates, which gives

$$T\mathbf{0} = \begin{bmatrix} 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 3 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 2 \\ 1 \end{bmatrix}$$

So, the origin gets mapped to the point $\begin{bmatrix} -1 & 3 & 2 \end{bmatrix}^\top$.

3. We now have the combined rotation matrix, $\mathbf{R}$ (written without homogeneous coordinates), as

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

Using Rodrigues formula, we can obtain the direction of the axis of this combined rotation in the original frame of reference and the angle of rotation about this axis. According to the Rodrigues formula, the angle of rotation is given by

$$\theta = \cos^{-1}\left(\frac{\textsc{Trace}(\mathbf{R}) - 1}{2}\right) = \cos^{-1}\left(\frac{-1}{2}\right) = \frac{2\pi}{3}$$

The axis of rotation, $\hat{\mathbf{n}}$, is given by

$$\hat{\mathbf{n}} = \frac{1}{2\sin\theta} \begin{bmatrix} \mathbf{R}_{32} - \mathbf{R}_{23} \\ \mathbf{R}_{13} - \mathbf{R}_{31} \\ \mathbf{R}_{21} - \mathbf{R}_{12} \end{bmatrix} = \frac{1}{2\sin\frac{2\pi}{3}} \begin{bmatrix} -1 - 0 \\ 1 - 0 \\ -1 - 0 \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

Therefore, the combined rotation is a rotation of $\frac{2\pi}{3}$ about the axis $\frac{1}{\sqrt{3}}\begin{bmatrix} -1 & 1 & -1 \end{bmatrix}^\top$, i.e.

$$\mathbf{R} \equiv R\left(\frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}, \frac{2\pi}{3}\right)$$

4. We now use the above axis $\hat{\mathbf{n}}$ and angle θ to calculate the the rotation matrix (say) $\mathbf{R}'$ for the rotation, and show that it is the same as matrix $\mathbf{R}$ that we obtained through sequentially applying the two given rotations. Using the Rodrigues formula, the rotation matrix $\mathbf{R}'$ is given by

$$\mathbf{R}' = \mathbf{I} + \sin\theta\mathbf{N} + (1 - \cos\theta)\mathbf{N}^2 \quad \text{where} \quad \mathbf{N} = \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix} \quad (n_i \text{ represent the components of } \hat{\mathbf{n}})$$

We first find $\mathbf{N}$ and then use it to find $\mathbf{R}'$. We have

$$\mathbf{N} = \frac{1}{\sqrt{3}} \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \implies \mathbf{N}^2 = \frac{1}{3} \begin{bmatrix} -2 & -1 & 1 \\ -1 & -2 & -1 \\ 1 & -1 & -2 \end{bmatrix}$$

Using these, we find

$$\begin{aligned} \mathbf{R}' &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \frac{1}{\sqrt{3}} \sin\frac{2\pi}{3} \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} + \frac{1}{3}\left(1 - \cos\frac{2\pi}{3}\right) \begin{bmatrix} -2 & -1 & 1 \\ -1 & -2 & -1 \\ 1 & -1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} + \frac{1}{6} \begin{bmatrix} -2 & -1 & 1 \\ -1 & -2 & -1 \\ 1 & -1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} = \mathbf{R} \end{aligned}$$

Hence, we have shown that the rotation matrix $\mathbf{R}'$ obtained using the axis $\hat{\mathbf{n}}$ and angle θ is the same as the matrix $\mathbf{R}$ that we obtained through sequentially applying the two given rotations. This also proves the correctness of the axis and angle obtained using the Rodrigues formula.

Question 2.

By Rodrigues formula, we know that the rotated vector for the given rotation is given by

$$\begin{aligned}
\mathbf{R}\mathbf{x} &= \mathbf{x} + (\hat{\mathbf{u}} \times \mathbf{x}) \sin\theta + (1 - \cos\theta)\hat{\mathbf{u}} \times (\hat{\mathbf{u}} \times \mathbf{x}) \\
&= \mathbf{x} + (\hat{\mathbf{u}} \times \mathbf{x}) \sin\theta + (1 - \cos\theta) \left[(\hat{\mathbf{u}}^\top \mathbf{x})\hat{\mathbf{u}} - (\hat{\mathbf{u}}^\top \hat{\mathbf{u}})\mathbf{x}\right] \\
&= \mathbf{x} + (\hat{\mathbf{u}} \times \mathbf{x}) \sin\theta + (1 - \cos\theta)(\hat{\mathbf{u}}^\top \mathbf{x})\hat{\mathbf{u}} - (1 - \cos\theta)\mathbf{x} \\
&= \mathbf{x} - \mathbf{x} + \mathbf{x}\cos\theta + (\hat{\mathbf{u}} \times \mathbf{x}) \sin\theta + (1 - \cos\theta)(\hat{\mathbf{u}}^\top \mathbf{x})\hat{\mathbf{u}} \\
&= \mathbf{x}\cos\theta + (\hat{\mathbf{u}} \times \mathbf{x}) \sin\theta + (\hat{\mathbf{u}}^\top \mathbf{x})(1 - \cos\theta)\hat{\mathbf{u}} \qquad \square
\end{aligned}$$

which proves the result, using the vector triple product identity that states for any vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a}^\top \mathbf{c})\mathbf{b} - (\mathbf{a}^\top \mathbf{b})\mathbf{c}$$

and that $\hat{\mathbf{u}}^\top \hat{\mathbf{u}} = 1$ since $\hat{\mathbf{u}}$ is a unit vector.

Question 3.

We are given image formation equations of two cameras, $\mathbf{C}_1$ and $\mathbf{C}_2$, for the same 3D point $\mathbf{X}$ in the world in homogeneous coordinates, as follows

$$\begin{aligned}
\mathbf{x}_1 &= \mathbf{K}_1 \left[\mathbf{R}_1 \mid \mathbf{t}_1\right] \mathbf{X} \\
\mathbf{x}_2 &= \mathbf{K}_2 \left[\mathbf{R}_2 \mid \mathbf{t}_2\right] \mathbf{X}
\end{aligned}$$

where we assume that the extrinsic matrices $[\mathbf{R}_i \mid \mathbf{t}_i]_{3\times 4}$, include the perspective projection matrix $\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}$. We need to show that the image points are related by

$$\mathbf{x}_1 = \mathbf{H}\mathbf{x}_2$$

for some invertible $\mathbf{H}_{3\times 3}$ and find it in terms of $\mathbf{K}_1$, $\mathbf{K}_2$, and $\mathbf{R}$.
Given that the orientation of $\mathbf{C}_2$ is obtained by applying a 3D rotation $\mathbf{R}$ on $\mathbf{C}_1$, which means

$$[\mathbf{R}_2 \mid \mathbf{t}_2] = \mathbf{R}\,[\mathbf{R}_1 \mid \mathbf{t}_1]$$

where the last equality follows as $\mathbf{R}$ is an orthogonal matrix. Then, we have

$$\begin{aligned}
\mathbf{x}_1 &= \mathbf{H}\mathbf{x}_2 \\
\mathbf{x}_1 &= \mathbf{H}\mathbf{K}_2 \left[\mathbf{R}_2 \mid \mathbf{t}_2\right] \mathbf{X} \\
\mathbf{K}_1 \left[\mathbf{R}_1 \mid \mathbf{t}_1\right] \mathbf{X} &= \mathbf{H}\mathbf{K}_2\mathbf{R} \left[\mathbf{R}_1 \mid \mathbf{t}_1\right] \mathbf{X} \\
\implies \mathbf{K}_1 \left[\mathbf{R}_1 \mid \mathbf{t}_1\right] &= \mathbf{H}\mathbf{K}_2\mathbf{R} \left[\mathbf{R}_1 \mid \mathbf{t}_1\right] \quad \text{since } \mathbf{X} \text{ is arbitrary}
\end{aligned}$$

At this point, it is important to consider the shape of the matrices involved. We have

$$\underbrace{\mathbf{K}_1}_{3\times 3}\underbrace{[\mathbf{R}_1 \mid \mathbf{t}_1]}_{3\times 4} = \underbrace{\mathbf{H}}_{3\times 3}\underbrace{\mathbf{K}_2}_{3\times 3}\underbrace{\mathbf{R}}_{3\times 3}\underbrace{[\mathbf{R}_1 \mid \mathbf{t}_1]}_{3\times 4}$$

Note that $[\mathbf{R}_1 \mid \mathbf{t}_1]$ is not invertible. However, we do not need to invert it. Since the operands to the right remain the same, it is sufficient to find $\mathbf{H}$ such that the

$$\begin{aligned}
\mathbf{K}_1 &= \mathbf{H}\mathbf{K}_2\mathbf{R} \\
\implies \mathbf{K}_1\mathbf{R}^T &= \mathbf{H}\mathbf{K}_2 \\
\implies \mathbf{K}_1\mathbf{R}^T\mathbf{K}_2^{-1} &= \mathbf{H}
\end{aligned}$$

Therefore, the matrix $\mathbf{H} = \mathbf{K}_1\mathbf{R}^T\mathbf{K}_2^{-1}$ gives the required relation between the image points. Finally, we show two properties of $\mathbf{H}$

1. $\mathbf{H}$ is of size 3×3. This is easy to see as

$$\underbrace{\mathbf{H}}_{3\times 3} = \underbrace{\mathbf{K}_1}_{3\times 3}\underbrace{\mathbf{R}^T}_{3\times 3}\underbrace{\mathbf{K}_2^{-1}}_{3\times 3}$$

2. $\mathbf{H}$ is invertible. This is also easy to see as

$$\mathbf{H}^{-1} = \left(\mathbf{K}_1 \mathbf{R}^T \mathbf{K}_2^{-1}\right)^{-1} = \mathbf{K}_2 \mathbf{R} \mathbf{K}_1^{-1}$$

where we crucially use the fact that $\mathbf{R}$ is an orthogonal matrix, and $\mathbf{K}_1$ and $\mathbf{K}_2$ are invertible since they are upper triangular matrices with non-zero entries (number of pixels) on their principal diagonals.

We can verify the correctness of the result by substituting the obtained $\mathbf{H}$ into the required relation. We have

$$\begin{aligned}
\mathbf{x}_1 &= \mathbf{H}\mathbf{x}_2 \\
&= \mathbf{K}_1 \mathbf{R}^T \mathbf{K}_2^{-1} \mathbf{x}_2 \\
&= \mathbf{K}_1 \mathbf{R}^T \mathbf{K}_2^{-1} \mathbf{K}_2 \left[\mathbf{R}_2 \mid \mathbf{t}_2\right] \mathbf{X} \\
&= \mathbf{K}_1 \mathbf{R}^T \left[\mathbf{R}_2 \mid \mathbf{t}_2\right] \mathbf{X} \\
&= \mathbf{K}_1 \left[\mathbf{R}_1 \mid \mathbf{t}_1\right] \mathbf{X} \\
&= \mathbf{x}_1
\end{aligned}$$

since $\left[\mathbf{R}_2 \mid \mathbf{t}_2\right] = \mathbf{R}\left[\mathbf{R}_1 \mid \mathbf{t}_1\right] \implies \mathbf{R}^T \left[\mathbf{R}_2 \mid \mathbf{t}_2\right] = \left[\mathbf{R}_1 \mid \mathbf{t}_1\right]$.