1 Numerical Exercises

1. Which of the following points written in homogeneous coordinates represent the same point in $\mathbb{P}^2$ as $[6, 3, 2]$?

   (a) $[18, 9, 6]$ (b) $[12, -6, 4]$ (c) $\left[1, \frac{1}{2}, \frac{1}{3}\right]$ (d) $[1, 2, 3]$

   Solution
   In order for two points to be equivalent in $\mathbb{P}^2$, there must be a $\lambda \in \mathbb{R}$, such that the first is written as a λ-multiple of the second.

   (a) $[18, 9, 6] = 3 \cdot [6, 3, 2] \Rightarrow$ Points are equivalent

   (b) If there exists such λ, it should hold that: $12 = 6\lambda, -6 = 3\lambda, 4 = 2\lambda$. This system of equations is inconsistent $\Rightarrow$ Points are not equivalent

   (c) $\left[1, \frac{1}{2}, \frac{1}{3}\right] = \frac{1}{6} \cdot [6, 3, 2] \Rightarrow$ Points are equivalent

   (d) The system of equations $1 = 6\lambda, 2 = 3\lambda, 3 = 2\lambda$ is inconsistent $\Rightarrow$ Points are not equivalent

2. Consider the following projective transformation $\lambda \tilde{\mathbf{q}} = H\, {}_W\mathbf{P}$, which maps a world point ${}_W\mathbf{P}$ to the homogeneous image point $\tilde{\mathbf{q}}$ with

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

   Compute the vanishing points for lines in the XZ plane in the world frame that are (1) parallel to the Z-axis, and (2) at 45° to the Z_W axis.

   Solution
   As seen in the lecture, the coordinates of a 3D line with direction $[l, m, n]$ passing through the point $[X_0, Y_0, Z_0]$ can be parametrized with $s \in \mathbb{R}$

$$\begin{aligned} X_W &= X_0 + sl \\ Y_W &= Y_0 + sm \\ Z_W &= Z_0 + sn \end{aligned}$$

   Given the projection matrix H, a 3D line can be projected to the image coordinates u and v

$$\begin{aligned} \lambda u &= X_W \\ \lambda v &= Y_W \\ \lambda &= Z_W \end{aligned}$$

$$u = \frac{X_W}{Z_W}, \qquad v = \frac{Y_W}{Z_W}$$

   In order to find the vanishing points, we need to compute the limit for $s \to \infty$

$$\lim_{s\to\infty} \frac{X_0 + sl}{Z_0 + sn} = \frac{l}{n} \qquad \lim_{s\to\infty} \frac{Y_0 + sm}{Z_0 + sn} = \frac{m}{n}$$

   (1) A line in the $X_W Z_W$ plane parallel to the Z_W-axis has the direction $[0, 0, 1]$. Thus, the corresponding vanishing point is $(0, 0)$
   (2) A line in the $X_W Z_W$ plane at 45° to the Z_W axis has the direction $[1, 0, 1]$. Thus, the corresponding vanishing point is $(1, 0)$

3. Consider a camera with the following camera matrix

$$K = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

State the modified camera matrix $\hat{K}$ if the image resolution is increased by a factor of two while the sensor size remains the same.

Solution
An increased resolution affects focal length f and the principal point (u_0, v_0).
Since the number of the pixels is doubled, the vertical and horizontal number of pixels is increased by a factor of $\sqrt{2}$. Thus, both offsets are multiplied by the same factor.
In addition, the focal length is multiplied by the same factor of $\sqrt{2}$ because the number of pixels along both axes is increased by a factor of $\sqrt{2}$ while the sensor size remains constant.
Thus, the modified camera matrix $\hat{K}$ is

$$\hat{K} = \begin{bmatrix} \sqrt{2}f & 0 & \sqrt{2}u_0 \\ 0 & \sqrt{2}f & \sqrt{2}v_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

4. Compute the homogeneous image point $\tilde{\mathbf{q}}$ in pixels corresponding to the point ${}_W\mathbf{P} = [2, 5, 5]^\top$ (meters). Consider the extrinsics $\mathbf{R}_{CW}$ and ${}_C\mathbf{t}_{CW}$ (meters) and camera matrix K

$$\mathbf{R}_{CW} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad {}_C\mathbf{t}_{CW} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad K = \begin{bmatrix} 420 & 0 & 355 \\ 0 & 420 & 250 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution

$$\lambda\tilde{\mathbf{q}} = K\left[\mathbf{R}_{CW}\middle|{}_C\mathbf{t}_{CW}\right]{}_W\tilde{\mathbf{P}} = \begin{bmatrix} 420 & 0 & 355 \\ 0 & 420 & 250 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4230 \\ 2760 \\ 6 \end{bmatrix} \Rightarrow \tilde{\mathbf{q}} = \begin{bmatrix} 705 \\ 460 \\ 1 \end{bmatrix}$$

5. Find the world-to-camera rigid-body transformation $\mathbf{T}_{CW}$ for the arrangement in Figure 1. State both $\mathbf{R}_{CW}$ and ${}_C\mathbf{t}_{CW}$.

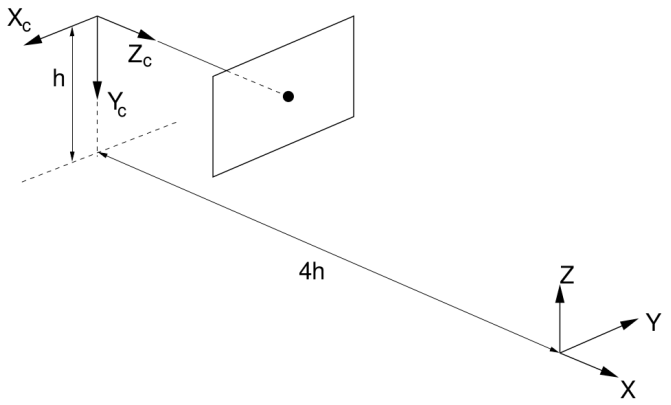


Figure 1: World frame W (axes X, Y, Z - subscript W omitted) and camera frame C (axes X_C, Y_C, Z_C). Distances are given in terms of h.

Solution
Inspecting the figure reveals that

$$X_C = -Y_W, \quad Y_C = -Z_W + h, \quad Z_C = X_W + 4h.$$

Thus, ${}_C\tilde{\mathbf{P}} = \mathbf{T}_{CW}\,{}_W\tilde{\mathbf{P}}$, with

$$\mathbf{T}_{CW} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & h \\ 1 & 0 & 0 & 4h \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{R}_{CW} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad {}_C\mathbf{t}_{CW} = \begin{bmatrix} 0 \\ h \\ 4h \end{bmatrix}.$$

6. A camera with frame C is initially aligned with the world frame W and located at the same origin. Therefore, a world point ${}_W\tilde{\mathbf{P}}$ projects to the homogeneous image point $\tilde{\mathbf{q}}$ according to

$$\lambda\tilde{\mathbf{q}} = K \left[\; I \;\middle|\; \mathbf{0} \;\right] {}_W\tilde{\mathbf{P}}.$$

The camera is then rotated about its optical centre, producing a new camera frame C'. Let $\mathbf{R}_{C'C}$ map vectors from C to C'. Show that all image points, independent of depth, undergo the projective transformation

$$\tilde{\mathbf{q}}' \sim K\mathbf{R}_{C'C}K^{-1}\tilde{\mathbf{q}},$$

where $\sim$ denotes equality up to a nonzero scale.

Solution
Before the camera is rotated, $\mathbf{R}_{CW} = I$ and ${}_C\mathbf{t}_{CW} = \mathbf{0}$. Hence,

$$\lambda\tilde{\mathbf{q}} = K\,{}_C\mathbf{P}, \qquad \text{so} \qquad {}_C\mathbf{P} = \lambda K^{-1}\tilde{\mathbf{q}}.$$

After the rotation, the same point is expressed in C' as

$${}_{C'}\mathbf{P} = \mathbf{R}_{C'C}\,{}_C\mathbf{P}.$$

It therefore projects to

$$\lambda'\tilde{\mathbf{q}}' = K\,{}_{C'}\mathbf{P} = \lambda K\mathbf{R}_{C'C}K^{-1}\tilde{\mathbf{q}} \quad \Longrightarrow \quad \tilde{\mathbf{q}}' \sim K\mathbf{R}_{C'C}K^{-1}\tilde{\mathbf{q}},$$

where $\sim$ denotes equality up to a nonzero scale. Hence, the relationship between corresponding points in the two images is a plane-to-plane projectivity.

7. A weak perspective projection comprises an orthographic projection onto the plane $Z_C = Z_A$ followed by perspective projection onto the image plane. Here, Z_A is an assumed scene depth; the subscript A denotes the approximation and is not a coordinate frame.

   (a) Derive the homogeneous relationship between a camera-frame point ${}_C\tilde{\mathbf{P}} = (X_C, Y_C, Z_C, 1)^\top$ and its image $(su_A, sv_A, s)^\top$ under weak perspective projection.

   (b) Show that the error $(u - u_A, v - v_A)$ introduced by the weak perspective approximation is given by

$$\left(\frac{\Delta Z}{Z_A}(u - u_0), \frac{\Delta Z}{Z_A}(v - v_0)\right)$$

   where $\Delta Z \equiv Z_A - Z_C$.

   (c) Under what viewing conditions is weak perspective a good camera model? What are its advantages?

Solution

(a) Under weak perspective projection, we assume that all points lie at approximately the same depth Z_A from the camera. This allows the projection to be re-written as follows:

$$\begin{bmatrix} su_A \\ sv_A \\ s \end{bmatrix} = \begin{bmatrix} k_u f & 0 & 0 & u_0 Z_A \\ 0 & k_v f & 0 & v_0 Z_A \\ 0 & 0 & 0 & Z_A \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix}$$

Essentially, this projection matrix ignores the depth of the point and instead assumes each point is at depth Z_A.

(b) Under full perspective we have:

$$u = \frac{k_u f X_C + u_0 Z_C}{Z_C}.$$

Under weak perspective, we have:

$$u_A = \frac{k_u f X_C + u_0 Z_A}{Z_A}.$$

Thus, the error introduced by weak perspective is:

$$u - u_A = \left(\frac{k_u f X_C}{Z_C} - \frac{k_u f X_C}{Z_A}\right) + \cancel{\left(\frac{u_0 Z_C}{Z_C} - \frac{u_0 Z_A}{Z_A}\right)}.$$

The term simplifies as follows:

$$\frac{k_u f X_C}{Z_C} - \frac{k_u f X_C}{Z_A} = k_u f X_C \left(\frac{1}{Z_C} - \frac{1}{Z_A}\right),$$

which can also be written as:

$$u - u_A = k_u f X_C \left(\frac{Z_A - Z_C}{Z_A Z_C}\right).$$

We can factor out $\Delta Z = Z_A - Z_C$ and rewrite the expression by adding and subtracting $\frac{u_0 Z_C}{Z_C}$:

$$u - u_A = \frac{\Delta Z}{Z_A}\left(\underbrace{\frac{k_u f X_C}{Z_C} + \frac{u_0 Z_C}{Z_C}}_{=u} - \underbrace{\frac{u_0 Z_C}{Z_C}}_{u_0}\right).$$

Thus, the error in u introduced by the weak perspective approximation is:

$$u - u_A = \frac{\Delta Z}{Z_A}(u - u_0).$$

Similarly, for the v-coordinate, we have:

$$v - v_A = \frac{\Delta Z}{Z_A}(v - v_0).$$

Therefore, the weak perspective approximation is perfect at the centre of the image, but introduces an error that depends on the difference in depth $\Delta Z = Z_A - Z_C$ and increases as we move away from the principal point (u_0, v_0).

(c) Weak perspective is a good approximation when the depth range of objects in the scene is small compared with the viewing distance. A good rule of thumb is that the viewing distance should be at least ten times the depth range.
The main advantage of the weak perspective model is that it is easier to calibrate than the full perspective model. The calibration requires fewer points with known world position, and, since the model is linear, the calibration process is also better conditioned (less sensitive to noise) than the nonlinear full perspective calibration.