1 Numerical Exercises

1. Which of the following points written in homogeneous coordinats 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 $\lambda q = HP$, which maps a world point P to the image point 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 (1) parallel to the Z-axis and (2) at 45° to the Z axis.

   Solution
   As seen in the lecture, the coordinates of a 3D line with direction $[l, m, n]$ passing though the point $[X_0, Y_0, Z_0]$ can be parametrized with $s \in \mathbb{R}$
$$\begin{aligned} X &= X_0 + sl \\ Y &= Y_0 + sm \\ Z &= Z_0 + sn \end{aligned}$$
   Given the projection matrix H, a 3D line can be projected to the image coordinates u and v
$$\begin{aligned} \lambda u &= X \\ \lambda v &= Y \\ \lambda &= Z \end{aligned}$$
$$u = \frac{X}{Z}, \qquad v = \frac{Y}{Z}$$
   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 XZ plane parallel to the Z-axis has the direction $[0, 0, 1]$. Thus, the corresponding vanishing points is $(0, 0)$
   (2) A line in the XZ plane at 45° to the Z axis has the direction $[1, 0, 1]$. Thus, the corresponding vanishing points is $(1, 0)$

3. Consider a camera with the following camera matrix

$$K = \begin{bmatrix} f & 0 & u_x \\ 0 & f & u_y \\ 0 & 0 & 1 \end{bmatrix}.$$

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

Solution

An increased resolution affects focal length f and the principal points u_x and u_y.
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.
The modified focal length is multipled by the same factor of $\sqrt{2}$ as the conversion factor from meters to pixel is increased as well.
Thus, the modified camera matrix $\hat{K}$ is

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

4. Compute the image point q in pixels corresponding to the point $P = [2, 5, 5]$ expressed in the world frame in meters. Consider the following camera extrinsics: rotation matrix R and translation vector T (in meters) and camera matrix K

$$R = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad T = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad K = \begin{bmatrix} 420 & 0 & 355 \\ 0 & 420 & 250 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution

$$\lambda q = K\,[R|T]\,P = \begin{bmatrix} 420 & 0 & 355 \\ 0 & 420 & 250 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4230 \\ 2760 \\ 6 \end{bmatrix} \Rightarrow q = \begin{bmatrix} 705 \\ 460 \\ 1 \end{bmatrix}$$

5. Find the rigid body transformation matrix (specifying the mapping from world to camera coordinates) for the arrangement in Figure 1

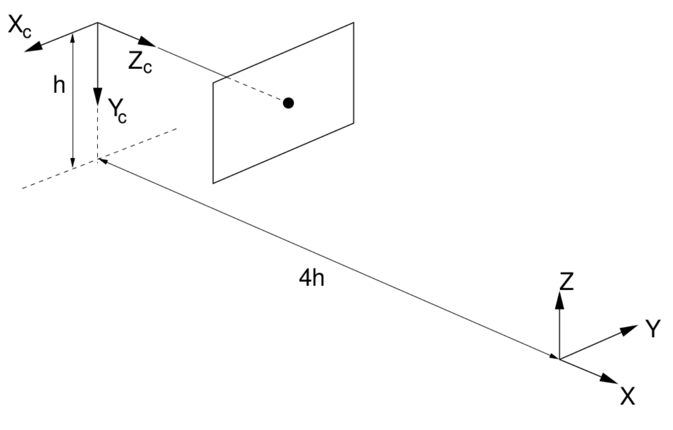

Figure 1: An example for camera transformation.

Solution

Inspecting the figure reveals that

$$X_c = -Y, \quad Y_c = -Z + h, \quad Z_c = X + 4h$$

The rigid body transformation between world and camera coordinates is therefore

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & h \\ 1 & 0 & 0 & 4h \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

6. A camera is used to photograph a 3D scene such that the relationship between world and image pixel coordinates (in homogeneous coordinates) is given by:

$$\tilde{\boldsymbol{w}} = \mathbf{K} \left[\begin{array}{c|c} \mathbf{I} & \mathbf{O} \end{array} \right] \tilde{\boldsymbol{X}}$$

If the camera is rotated about the optical centre, show that all the image points, independent of depth, undergo a transformation given by:

$$\tilde{\boldsymbol{w}}' = \mathbf{K}\mathbf{R}\mathbf{K}^{-1}\tilde{\boldsymbol{w}}$$

Solution
Before the camera is rotated, the camera is aligned with the world coordinate system and hence

$$\tilde{\boldsymbol{w}} = \mathbf{K} \left[\begin{array}{c|c} \mathbf{I} & \mathbf{O} \end{array} \right] \tilde{\boldsymbol{X}} = \mathbf{K} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{K}\mathbf{X}$$

It follows that

$$\mathbf{X} = \mathbf{K}^{-1}\tilde{\boldsymbol{w}}$$

After rotating by $\mathbf{R}$ about the optical centre, the same world point $\mathbf{X}$ projects to a different image point $\tilde{\boldsymbol{w}}'$ as follows:

$$\tilde{\boldsymbol{w}}' = \mathbf{K} \left[\begin{array}{c|c} \mathbf{R} & \mathbf{O} \end{array} \right] \tilde{\boldsymbol{X}} = \mathbf{K}\mathbf{R} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{K}\mathbf{R}\mathbf{X} = \mathbf{K}\mathbf{R}\mathbf{K}^{-1}\tilde{\boldsymbol{w}}$$

Hence the relationship between points in the original image and corresponding points in the second image 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.

   (a) Derive the homogeneous relationship between a point $(X_c, Y_c, Z_c, 1)$ and its image (su_A, sv_A, s) 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:

$$\begin{bmatrix} su_A \\ sv_A \\ s \end{bmatrix} = \begin{bmatrix} k_u f & 0 & 0 & u_0 Z_A \\ 0 & k_v f & 0 & v_0 Z_A \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

(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.