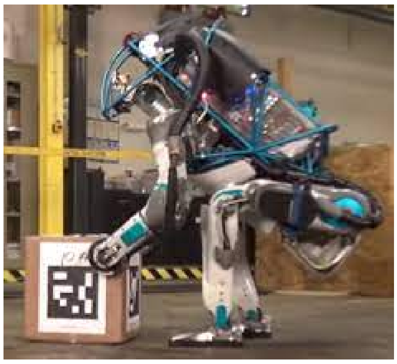
Exercise Image Processing
Sample Solution

Prof. Dr.-Ing. Volker Willert

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Sheet 2



In many applications of robotics and industrial image processing, the position and orientation of a camera relative to a known object should be determined from an image. For example, 3D coordinates of a workpiece may be known and the position of the camera is to be determined on the basis of projections of these points in the camera image. Other examples can be found in robot navigation. For example, the Boston Dynamics robot dog Spot recognizes its relative position to the charging station, or the humanoid robot Atlas knows how its hands are oriented to a box if the geometry of the box is known.

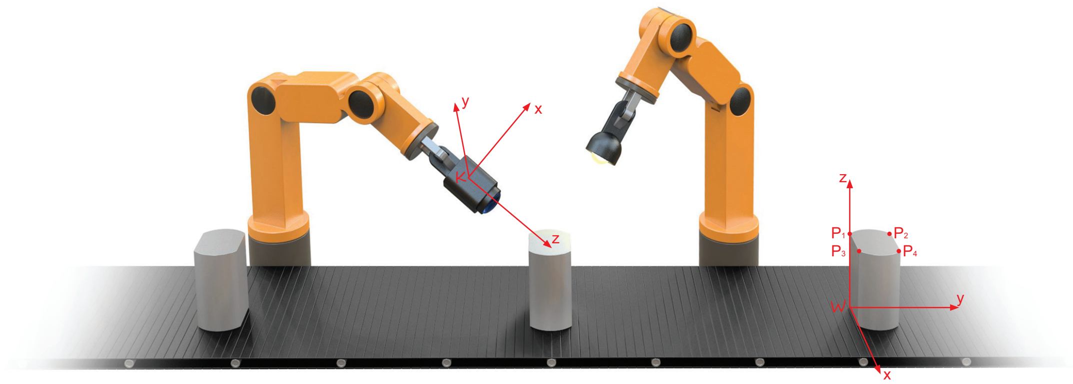


Figure 1: Where is the camera located in relation to the workpiece?

In this exercise we deal with the determination of the position of a geometrically calibrated camera in relation to an object, of which we know the coordinates of four points on a surface on the object. As an example we use the same scenario as in exercise sheet 1. The coordinates of the four vertices of the section with respect to the workpiece coordinate system are in millimeters: $\mathbf{p}_1 = [0, 0, 50]^T$, $\mathbf{p}_2 = [0, 20, 50]^T$, $\mathbf{p}_3 = [20, 0, 50]^T$ and $\mathbf{p}_4 = [20, 20, 50]^T$. The camera is calibrated and we have written an image processing algorithm that can detect the workpiece in the image and determine the projections of the vertices $\mathbf{x}'_1$, $\mathbf{x}'_2$, $\mathbf{x}'_3$ and $\mathbf{x}'_4$ with pixel accuracy.

Task 2.1: Determining the Pose of a Calibrated Camera

Calibration of the camera revealed the following intrinsic parameters:

$$\begin{aligned}
\text{principal point}: &\quad o_x = 600\,[\text{px}]\,, \quad o_y = 800\,[\text{px}]\,,\\
\text{scale factors}: &\quad s_x = 400\,[\text{px/mm}]\,, \quad s_y = 400\,[\text{px/mm}]\,,\\
\text{camera constant}: &\quad c = 5[\text{mm}]\,.
\end{aligned}$$

The image processing algorithm measures the following coordinates with pixel accuracy:

$$\begin{aligned}
\mathbf{x}_1' =& [489, 689]^\top\\
\mathbf{x}_2' =& [732, 405]^\top\\
\mathbf{x}_3' =& [803, 1003]^\top\\
\mathbf{x}_4' =& [1105, 777]^\top
\end{aligned}$$

The measurement error is therefore in the range of $\Delta x, \Delta y = [0; 0.5]$ pixels.

Since all points lie on a plane, we can determine the pose of the camera by determining the homography between the bolt cross-sectional area and the image plane. We proceed as follows:

a) Define the coordinate system of the bolt such that the X and Y components do not change and the Z components all add up to zero.

   Answer: The coordinate system is shifted along the Z axis by 50mm. This results in the new coordinates of the four corner points of the bolt cross-section surface to: $\mathbf{p}_1 = [0, 0, 0]^T$, $\mathbf{p}_2 = [0, 20, 0]^T$, $\mathbf{p}_3 = [20, 0, 0]^T$ and $\mathbf{p}_4 = [20, 20, 0]^T$ in millimeter $[mm]$.

b) Compute the inverse calibration matrix $\mathbf{K}^{-1}$ to determine the normalized image coordinates $\mathbf{x}_1$ to $\mathbf{x}_4$ of points p_1 to p_4.

   Answer: The relationship between normalized image coordinates $\mathbf{x}$ and measured image coordinates $\mathbf{x}'$ is given by the calibration matrix as follows:

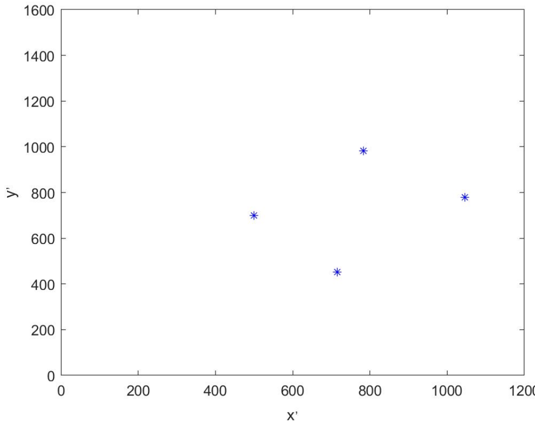
$$\begin{aligned}
\mathbf{x}' &= \mathbf{K}\mathbf{x}\,,\\
\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} &= \begin{bmatrix} cs_x & 0 & o_x \\ 0 & cs_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 2000[px] & 0 & 600[px] \\ 0 & 2000[px] & 800[px] \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}\,,
\end{aligned}$$

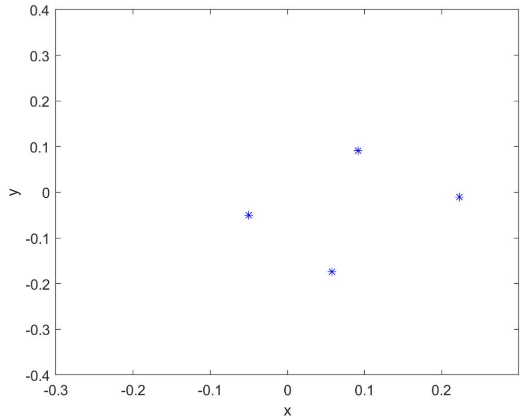
   where the shear factor $s_\theta = 0$ is set to zero. By converting the individual equations and solving to $\mathbf{x}$, the inverse calibration matrix results in:

$$\begin{aligned}
\mathbf{x} &= \mathbf{K}^{-1}\mathbf{x}'\,,\\
\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} &= \begin{bmatrix} \frac{1}{cs_x} & 0 & -\frac{o_x}{cs_x} \\ 0 & \frac{1}{cs_y} & -\frac{o_y}{cs_y} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 5\cdot 10^{-4}[px^{-1}] & 0 & -0.3 \\ 0 & 5\cdot 10^{-4}[px^{-1}] & -0.4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}\,,
\end{aligned}$$

   Here you can see again clearly that the measured image matrix coordinates $\mathbf{x}'$ are measured in pixels $[px]$ and the normalized coordinates $\mathbf{x}$ are dimensionless. The normalized coordinates of the vertices result to:

$$\begin{aligned}
\mathbf{x}_1 =& [-0.05, -0.05]^\top\\
\mathbf{x}_2 =& [0.058, -0.1745]^\top\\
\mathbf{x}_3 =& [0.0915, 0.0915]^\top\\
\mathbf{x}_4 =& [0.2230, -0.0100]^\top
\end{aligned}$$





c) Use the projection equation for points $\mathbf{X}_i$ on a plane $\mathbf{x}_i = \mathbf{H}\mathbf{X}_i$ to set up a homogeneous linear system of equations using the Direct Linear Transformation (DLT). What is the minimum number of points you need to uniquely determine the homography $\mathbf{H}$?

Answer: To determine the homography one needs at least 4 points. The homography $\mathbf{H}$ describes the projection from the vertices $\mathbf{X}_i = [X_i, Y_i, 1]^\top$ on the 3D section plane of the bolt in space to the 2D points $\mathbf{x}_i = [x_i, y_i, 1]^\top$ in the image plane:

$$\lambda \mathbf{x}_i = \lambda \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \underbrace{\mathbf{K}\,[\mathbf{r}_1, \mathbf{r}_2, \mathbf{T}]}_{\mathbf{H}} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} = \mathbf{H}\mathbf{X}_i \,.$$

Applying the direct linear transformation

$$\widehat{\mathbf{x}}_i \mathbf{H}\mathbf{X}_i = 0\,, \quad \text{wobei} \quad \mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]\,, \quad \text{und} \quad \widehat{\mathbf{x}}_i = \begin{bmatrix} 0 & -1 & y_i \\ 1 & 0 & -x_i \\ -y_i & x_i & 0 \end{bmatrix},$$

$$\rightarrow \quad \widehat{\mathbf{x}}_i \otimes \mathbf{X}_i^\top \mathbf{h} = \mathbf{0}\,, \quad \text{mit} \quad \mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{bmatrix},$$

$$\rightarrow \quad \underbrace{\begin{bmatrix} \mathbf{0}^\top & -\mathbf{X}_i^\top & y_i \mathbf{X}_i^\top \\ \mathbf{X}_i^\top & \mathbf{0}^\top & -x_i \mathbf{X}_i^\top \end{bmatrix}}_{\mathbf{M}_i \ 2\times 9} \mathbf{h} = \mathbf{0}\,,$$

then for each pair of points $(\mathbf{x}_i, \mathbf{X}_i)$ three equations of a homogeneous system of equations result, where only the first two equations are linearly independent. These two equations form a part of the measurement matrix $\mathbf{M}_i$. For four points the complete measurement matrix and the homogeneous equation system to be solved results to:

$$\mathbf{M} = [\mathbf{M}_1^\top, \mathbf{M}_2^\top, \mathbf{M}_3^\top, \mathbf{M}_4^\top]^\top\,, \quad \mathbf{M}\mathbf{h} = \mathbf{0}\,.$$

d) Reconstruct the pose $(\mathbf{R}, \mathbf{T})$ of the camera from the homography $\mathbf{H}$. Check if the rotation can be read directly? Are the conditions for a rotation matrix satisfied?

Answer: About a singular value decomposition

$$\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$$

of the measurement matrix $\mathbf{M}$ of size 8×9, the solution of the system of equations $\mathbf{h} = \mathbf{v}_9$ is obtained, where $\mathbf{v}_9$ corresponds to the right-hand singular vector associated to the smallest singular value s_9. The scaling of the solution of a homogeneous system of equations is arbitrary. When solving via singular value decomposition, the magnitude of $||\mathbf{h}|| = 1$ is normalized to one. The correct scaling is obtained via the additional knowledge of the three columns. The last three elements of $\mathbf{v}_9$ correspond to the third column $\mathbf{h}_3$ of the homography matrix and contain a scaled estimate of the translation vector $\mathbf{T}$. The first six elements of $\mathbf{v}_9$ correspond to the first two columns $\mathbf{h}_1$ and $\mathbf{h}_2$ of the homography matrix and contain a scaled estimate of the first two columns of the rotation matrix $\mathbf{r}_1$ and $\mathbf{r}_2$. The sums of the columns of a rotation matrix must all add up to one and be orthogonal to each other. We can take advantage of this to get the correct scaling of the translation vector. For this we decompose the first two columns of the homography matrix by a singular value decomposition

$$[\mathbf{h}_1 \mathbf{h}_2] = \tilde{\mathbf{U}}\tilde{\mathbf{S}}\tilde{\mathbf{V}}^\top$$

and divide all elements of the homography matrix by the mean of the first two singular values

$$\hat{\mathbf{H}} = \frac{1}{(\tilde{s}_1 + \tilde{s}_2)}\mathbf{H}\,.$$

By this normalization, the magnitudes of the first two columns are approximately one and the third column gives a good estimate of the translation vector with proper scaling:

$$\hat{\mathbf{T}} = \hat{\mathbf{h}}_3\,.$$

The first two columns $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ have only approximately the amount 1 and are also not exactly orthogonal to each other. To obtain from them an estimate of a rotation matrix with exact properties of a rotation matrix, the vector orthogonal to $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ is calculated via the cross product and the singular value decomposition is applied to the resulting inexact rotation matrix:

$$\left[\hat{\mathbf{h}}_1\,, \hat{\mathbf{h}}_2\,, \hat{\mathbf{h}}_1 \times \hat{\mathbf{h}}_2\right] = \overline{\mathbf{U}}\,\overline{\mathbf{S}}\,\overline{\mathbf{V}}^\top\,.$$

A good estimate of the rotation matrix is obtained by setting all singular values of the decomposition to one $\overline{\mathbf{S}} = \mathbf{I}$:

$$\hat{\mathbf{R}} = \overline{\mathbf{U}}\,\mathbf{I}\,\overline{\mathbf{V}}^\top\,.$$

Here $\mathbf{I}$ corresponds to a 3×3 unit matrix.

Task 2.2: Decomposition of a Projection Matrix

You have bought a new camera and in the data sheet the intrinsic parameters are not completely specified. Therefore, you perform a camera calibration and thereby determine the projection matrix:

$$\mathbf{P} = \begin{bmatrix} 2 & -\sqrt{3}/2 & 3/2 & 2+\sqrt{3} \\ 0 & \sqrt{3}/2 & 5/2 & 2+\sqrt{3} \\ 0 & -1/2 & \sqrt{3}/2 & 1 \end{bmatrix} = \mathbf{K}\left[\ \mathbf{R}|\mathbf{T}\ \right] = \left[\ \mathbf{KR}|\mathbf{KT}\ \right]$$

You must now decompose this into the individual calibration matrices of the intrinsics $\mathbf{K}$ and extrinsics $\mathbf{R}, \mathbf{T}$. To do this, perform the RQ decomposition (not QR decomposition!).

a) Let's assume a matrix equation $\mathbf{M} = \mathbf{RQ}$. What kind of decomposition results if we transpose the equation?

Answer: Transposing the matrix equation, we get a product of an orthogonal matrix $\mathbf{Q}^\top$ with a lower triangular matrix $\mathbf{R}^\top$, since the transposed of an orthogonal matrix remains orthogonal and the transposed of an upper triangular matrix results in a lower triangular matrix:

$$\mathbf{M}^\top = (\mathbf{RQ})^\top = \mathbf{Q}^\top\mathbf{R}^\top\,.$$

Thus, we do not arrive at a QR decomposition via the transpose.

b) Using a permutation matrix $\mathbf{\Gamma}$ which, multiplied from the left, reverses the order of the rows and, multiplied from the right, reverses the order of the columns of a (3x3) matrix, try to transform the matrix equation from a) in such a way that a QR decomposition is possible.

$$\mathbf{\Gamma} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ mit } \mathbf{\Gamma} = \mathbf{\Gamma}^T = \mathbf{\Gamma}^{-1}.$$

Answer: For the first three columns of $\mathbf{P}$ the following relation holds:

$$\mathbf{M} = \begin{bmatrix} 2 & -\frac{\sqrt{3}}{2} & \frac{3}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{5}{2} \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} = \mathbf{K}\mathbf{R}_{rot} = \mathbf{R}\mathbf{Q} \neq \mathbf{Q}\mathbf{R}.$$

It is obvious that a QR decomposition cannot be applied directly. The question is: How do you get from a RQ to a QR decomposition? The answer is: By a rotation of the matrix $\mathbf{M}$ by $90°$ in a clockwise direction. The original problem is reformulated as follows:

$$\bar{\mathbf{M}} = \mathbf{M}^T\mathbf{\Gamma} = (\mathbf{R}\mathbf{Q})^T\mathbf{\Gamma} = \mathbf{Q}^T\mathbf{R}^T\mathbf{\Gamma} \stackrel{!}{=} \bar{\mathbf{Q}}\bar{\mathbf{R}}\,,$$

where $\bar{\mathbf{R}}$ again is an upper triangular matrix and $\bar{\mathbf{Q}}$ must again result in an orthogonal matrix. This is achieved by a clever expansion with the unit matrix $\mathbf{I} = \mathbf{\Gamma}\mathbf{\Gamma}$:

$$\bar{\mathbf{M}} = \mathbf{M}^T\mathbf{\Gamma} = (\mathbf{R}\mathbf{Q})^T\mathbf{\Gamma} = \underbrace{\mathbf{Q}^T\mathbf{\Gamma}}_{\bar{\mathbf{Q}}}\underbrace{\mathbf{\Gamma}\mathbf{R}^T\mathbf{\Gamma}}_{\bar{\mathbf{R}}} = \bar{\mathbf{Q}}\bar{\mathbf{R}}\,,$$

This results in the transformed matrix $\bar{\mathbf{M}}$, which can now be decomposed using the QR method:

$$\mathbf{M}^T\mathbf{\Gamma} = \bar{\mathbf{Q}}\bar{\mathbf{R}} = \bar{\mathbf{M}} = \begin{bmatrix} 0 & 0 & 2 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{5}{2} & \frac{3}{2} \end{bmatrix}.$$

From the calculation of the QR decomposition of the matrix $\bar{\mathbf{M}}$, the matrices $\mathbf{K} = \mathbf{R} = \mathbf{\Gamma}\bar{\mathbf{R}}^T\mathbf{\Gamma}$ and $\mathbf{R}_{rot} = \mathbf{Q} = \mathbf{\Gamma}\bar{\mathbf{Q}}^T$ can then be determined.

c) Test your solution with `Matlab`.

Answer: A QR decomposition of $\bar{\mathbf{M}}$ via the Matlab function `qr` yields:

$$\bar{\mathbf{Q}}\bar{\mathbf{R}} = \begin{bmatrix} 0 & 0 & -1 \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -1 & -\sqrt{3} & -\sqrt{3} \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

By the back transformations one obtains:

$$\mathbf{K} = \mathbf{R} = \mathbf{\Gamma}\bar{\mathbf{R}}^T\mathbf{\Gamma} = \begin{bmatrix} -2 & 0 & -\sqrt{3} \\ 0 & -2 & -\sqrt{3} \\ 0 & 0 & -1 \end{bmatrix},$$

and

$$\mathbf{R}_{rot} = \mathbf{Q} = \mathbf{\Gamma}\bar{\mathbf{Q}}^T = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}.$$

The upper triangular matrix does not yet satisfy the condition of a calibration matrix, since all diagonal elements are negative, but they must all be positive. By extension with a $1 = (-1)(-1)$ we get the estimation results $\hat{\mathbf{K}}$ und $\hat{\mathbf{R}}_{rot}$:

$$\hat{\mathbf{K}}\hat{\mathbf{R}}_{rot} = (-\mathbf{K})(-\mathbf{R}_{rot}) = \begin{bmatrix} 2 & 0 & \sqrt{3} \\ 0 & 2 & \sqrt{3} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}.$$

The translation vector is given by:

$$\hat{\mathbf{T}} = \hat{\mathbf{K}}^{-1}\mathbf{p}_4 = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{\sqrt{3}}{2} \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2+\sqrt{3} \\ 2+\sqrt{3} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Task 2.3: Additional task: Determination of the world coordinates of the projection center via the projection matrix

In this task, the projection center in world coordinates corresponding to the null space of the projection matrix $\mathbf{\Pi}$ should be computed, which is formed by the following matrices $\mathbf{K}$, $\mathbf{R}$ and the vector $\mathbf{T}$, without decomposing the projection matrix into the individual matrices.

a) Use the singular value decomposition of $\mathbf{\Pi} = \mathbf{U}\mathbf{S}\mathbf{V}^\star$.

Answer: The null space of a matrix can be read from the $\mathbf{V}$ matrix of its singular value decomposition. It is spanned by the column vectors of $\mathbf{V}$ whose corresponding singular values are equal to zero, since the singular values σ_i and right singular vectors $\mathbf{v}_i$ respectively left singular vectors $\mathbf{u}_i$ must satisfy the following equations must be satisfied:

$$\mathbf{\Pi}\mathbf{v}_i = \sigma_i\mathbf{u}_i \quad \text{bzw.} \quad \mathbf{\Pi}^\top\mathbf{u}_i = \sigma_i\mathbf{v}_i\,.$$

Thus the null space of the matrix $\mathbf{\Pi}$ corresponds to the right singular vector $\mathbf{v}_i$ to the singular value $\sigma_i = 0$. The right singular vector to the singular value $\sigma_0 = 0$ is:

$$\mathbf{\Pi}\mathbf{v}_0 = \sigma_0\mathbf{u}_0 \quad \text{bzw.} \quad \mathbf{\Pi}^\top\mathbf{u}_0 = \sigma_0\mathbf{v}_0\,, \quad \rightarrow \quad (\mathbf{\Pi}^T\mathbf{\Pi} - \sigma_0^2\mathbf{I})\mathbf{v}_0 = \mathbf{\Pi}^T\mathbf{\Pi}\mathbf{v}_0 = \mathbf{0}\,, \quad \rightarrow \quad \mathbf{v}_0 = \begin{bmatrix} 0 \\ 0 \\ -\sqrt{2} \\ 1 \end{bmatrix}.$$

So this eigenvector corresponds to the homogeneous world coordinates of the projection center $\mathbf{v}_0 = \overline{\mathbf{o}}_{\mathcal{W}}$.

b) Check your solution using the single matrices $\mathbf{K}$, $\mathbf{R}$ and the vector $\mathbf{T}$ and the formula from the script.

Answer:

$$\text{Formula from the lecture:} \quad \mathbf{o}_{\mathcal{W}} = -\left[\ \mathbf{K}\mathbf{R}\ \right]^{-1}\mathbf{K}\mathbf{T} = \begin{bmatrix} 0 \\ 0 \\ -\sqrt{2} \end{bmatrix}.$$

With the intermediate results:

$$\mathbf{K}\mathbf{R} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \sqrt{2} \\ -\frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{bmatrix}, \quad [\mathbf{K}\mathbf{R}]^{-1} = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & -\sqrt{2} \\ 0 & 1 & -1 \\ \frac{\sqrt{2}}{2} & 0 & 0 \end{bmatrix}, \quad \mathbf{K}\mathbf{T} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

The matrices are given as follows:

$$\mathbf{\Pi} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \\ \pi_{21} & \pi_{22} & \pi_{23} & \pi_{24} \\ \pi_{31} & \pi_{32} & \pi_{33} & \pi_{34} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \sqrt{2} & 2 \\ -\frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} & 1 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 1 \end{bmatrix}$$

$$\mathbf{K} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$