Multiple View Geometry: Exercise 3

Dr. Haoang Li, Daniil Sinitsyn, Sergei Solonets, Viktoria Ehm
Computer Vision Group, TU Munich
`Wednesdays 16:00-18:15 at Hörsaal 2, "Interims I" (5620.01.102), and on RBG Live`

Exercise: May 24, 2023

Image Formation

We are looking at the formation of an image in camera coordinates $\mathbf{X} = (X\ Y\ Z\ 1)^\top$. In the lecture, you learned the following relation of homogeneous pixel coordinates $\mathbf{x}'$ and $\mathbf{X}$:

$$\lambda \mathbf{x}' = K\Pi_0\mathbf{X} \tag{1}$$

with the intrinsic camera matrix K. To clearly differentiate between camera coordinates and pixel coordinates, call the pixel coordinates u and v: $\mathbf{x}' = (u\ v\ 1)^\top$. Furthermore, let the non-homogeneous camera coordinates be $\tilde{\mathbf{X}} := \Pi_0\mathbf{X} = (X\ Y\ Z)^\top$. (1) is then equivalent to

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K\tilde{\mathbf{X}}\,. \tag{2}$$

Let $s_x = s_y = 1$ and $s_\theta = 0$ in the intrinsic camera matrix.

1. Compute λ and show that (2) is equivalent to

$$u = \frac{fX}{Z} + o_x\,, \quad v = \frac{fY}{Z} + o_y\,. \tag{3}$$

2. A classic ambiguity of the perspective projection is that one cannot tell an object from another object that is exactly *twice as big but twice as far*. Explain why this is true.

3. For a camera with $f = 540$, $o_x = 320$ and $o_y = 240$, compute the pixel coordinates u and v of a point $\tilde{\mathbf{X}} = (60\ 100\ 180)^\top$. Explain with the help of (b) why the units of $\tilde{\mathbf{X}}$ are not needed for this task. Will the projected point be in the image if it has dimensions 640×480?

We define the generic projection π of $\tilde{\mathbf{X}}$ to 2D coordinates as follows:

$$\pi(\tilde{\mathbf{X}}) := \begin{pmatrix} X/Z \\ Y/Z \end{pmatrix} \tag{4}$$

4. Using the generic projection π, show that (3) — and therefore also (1) and (2) — is equivalent to

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \begin{pmatrix} \pi(\tilde{\mathbf{X}}) \\ 1 \end{pmatrix}\,. \tag{5}$$