CS-E4850 Computer Vision
Exercise Round 2

Alex Herrero Pons

Spring Term Course 2020-2021

Exercise 1. Pinhole camera.
The perspective projection equations for a pinhole camera are

$$x_p = f\frac{x_c}{z_c}, \qquad y_p = f\frac{y_c}{z_c}, \tag{1}$$

where $\mathbf{x}_p = [x_p, y_p]^\top$ are the projected coordinates on the image plane, $\mathbf{x}_c = [x_c, y_c, z_c]^\top$ is the imaged point in the camera coordinate frame and f is the focal length. Give a geometric justification for the perspective projection equations.

Solution. Having that the image plane is perpendicular to the optical axis, that the point $\mathbf{o}$ is the camera origin, $\mathbf{p}$ is the principal point in the image plain and $\mathbf{p}'$ the point where the optical axis pierces the the plain that contains $\mathbf{x}_c$ and is parallel to the image plane. We know that the triangle $A = (\mathbf{o}, \mathbf{p}, \mathbf{x}_p)$ is similar to the triangle $B = (\mathbf{o}, \mathbf{p}', \mathbf{x}_c)$. Hence, knowing that $dist(\mathbf{o}, \mathbf{p}) = f$ and $dist(\mathbf{o}, \mathbf{p}') = z_c$ we can make a simple trigonometry calculus to obtain that the equation (1) is in fact true.

□

Exercise 2. Pixel coordinate frame.
The image coordinates x_p and y_p given by the perspective projection equations (1) above are not in pixel units. The x_p and y_p coordinates have the same unit as distance f (typically millimeters) and the origin of the coordinate frame is the principal point (the point where the optical axis pierces the image plane). Now, give a formula which transforms the point $\mathbf{x}_p$ to its pixel coordinates $\mathbf{p} = [u, v]^\top$ when the number of pixels per unit distance in u and v directions are m_u and m_v, respectively, the pixel coordinates of the principal point are (u_0, v_0) and
a) u and v axis are parallel to x and y axis, respectively.

Solution. In this case we only need to convert the coordinate units into pixel coordinates and then translate from the principal point as follows,

$$u = u_0 + m_u x_p, \qquad v = v_0 + m_v y_p.$$

b) u axis is parallel to x axis and the angle between u and v axis is θ.

Solution. Now we have that the v is not necessarily parallel to the y axis. So for the coordinate u we have to add the incidence that has y on u. And for the coordinate v we have to make the correspondent trigonometry transformation. Let u_y be the incidence that has y on u. We know that $\cot\theta = u_y/y_p$ therefore we have that $u_y = y_p\cot\theta$. For the coordinate v we know that $\sin\theta = y_p/v$ so $v = y_p/\sin\theta$. Hence, applying the correspondent translation and scaling,

$$u = u_0 + m_u x_p - m_v y_p \cot\theta, \qquad v = v_0 + m_v \frac{y_p}{\sin\theta}.$$

Exercise 3. Intrinsic camera calibration matrix.
Use homogeneous coordinates to represent case (2.a) above with a matrix $\mathbf{K}_{3\times3}$, also known as the camera's intrinsic calibration matrix, so that $\tilde{\mathbf{p}} = \mathbf{K}\mathbf{x}_c$. Where $\tilde{\mathbf{p}}$ is $\mathbf{p}$ in homogeneous coordinates.

Solution.

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

Proof.

$$\mathbf{x}_c = [x_c, y_c, z_c]^\top = \left[\frac{x_c}{z_c}, \frac{y_c}{z_c}, 1\right]^\top.$$

$$\tilde{\mathbf{p}} = \mathbf{K}\mathbf{x}_c = \begin{bmatrix} m_u f & 0 & u_0 \\ 0 & m_v f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{x_c}{z_c} \\ \frac{y_c}{z_c} \\ 1 \end{bmatrix} = \begin{bmatrix} m_u f \frac{x_c}{z_c} + u_0 \\ m_v f \frac{y_c}{z_c} + v_0 \\ 1 \end{bmatrix} = \begin{bmatrix} m_u x_p + u_0 \\ m_v y_p + v_0 \\ 1 \end{bmatrix}.$$

$\square$

Exercise 4. Camera projection matrix.
Imaged points are often expressed in an arbitrary frame of reference called the world coordinate frame. The mapping from the world frame to the camera coordinate frame is a rigid transformation consisting of a 3D rotation $\mathbf{R}$ and translation $\mathbf{t}$:

$$\mathbf{x}_c = \mathbf{R}\mathbf{x}_w + \mathbf{t}.$$

Use homogeneous coordinates and the result of the exercise 3 above, to write down the 3×4 camera projection matrix $\mathbf{P}$ that projects a point form world coordinates $\mathbf{x}_w$ to pixel coordinates. That is, represent $\mathbf{P}$ as a function of the internal camera parameters $\mathbf{K}$ and the external camera parameters $\mathbf{R}$,$\mathbf{t}$.

Solution.

$$\mathbf{P}_{3\times 4} = \mathbf{K}[\mathbf{R}|\mathbf{t}] = \begin{bmatrix} m_u f & 0 & u_0 \\ 0 & m_v f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} =$$

$$= \begin{bmatrix} m_u f r_{11} + u_0 r_{31} & m_u f r_{12} + u_0 r_{32} & m_u f r_{13} + u_0 r_{33} & m_u f t_1 + u_0 t_3 \\ m_v f r_{21} + v_0 r_{31} & m_v f r_{22} + v_0 r_{32} & m_v f r_{23} + v_0 r_{33} & m_v f t_2 + v_0 t_3 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix}.$$

Exercise 5. Rotation matrix.
A rigid coordinate transformation can be represented with a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{t}$, which transform a point $\mathbf{x}$ to $\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{t}$. Now, let the 3×3 matrix $\mathbf{R}$ be a 3-D rotation matrix, which rotates a vector $\mathbf{x}$ by the angle θ about the axis $\mathbf{u}$ (a unit vector). According to the Rodrigues formula it holds that

$$\mathbf{R}\mathbf{x} = \cos\theta\,\mathbf{x} + \sin\theta\,\mathbf{u} \times \mathbf{x} + (1 - \cos\theta)(\mathbf{u} \cdot \mathbf{x})\mathbf{u}.$$

a) Give a geometric justification (i.e. derivation) for the Rodrigues formula.

Solution.
(In: Wikipedia Rodriges' rotation formula - https://bit.ly/3muAOkB)

b) Derive the expressions for the elements of $\mathbf{R}$ as a function of θ and the elements of $\mathbf{u}$.

Solution.

$$\mathbf{u} \times \mathbf{x} = \begin{bmatrix} u_2x_3 - u_3x_2 \\ u_3x_1 - u_1x_3 \\ u_1x_2 - u_2x_1 \end{bmatrix},$$

$$\mathbf{u} \cdot \mathbf{x} = (u_1x_1 + u_2x_2 + u_3x_3),$$

$$\mathbf{R}\mathbf{x} = \cos\theta\,\mathbf{x} + \sin\theta\,\mathbf{u} \times \mathbf{x} + (1 - \cos\theta)(\mathbf{u} \cdot \mathbf{x})\mathbf{u} =$$

$$= \cos\theta \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \sin\theta \begin{bmatrix} u_2x_3 - u_3x_2 \\ u_3x_1 - u_1x_3 \\ u_1x_2 - u_2x_1 \end{bmatrix} + (1 - \cos\theta)(u_1x_1 + u_2x_2 + u_3x_3) \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\theta x_1 + \sin\theta u_2x_3 - \sin\theta u_3x_2 + (1 - \cos\theta)(u_1x_1u_1 + u_2x_2u_1 + u_3x_3u_1) \\ \cos\theta x_2 + \sin\theta u_3x_1 - \sin\theta u_1x_3 + (1 - \cos\theta)(u_1x_1u_2 + u_2x_2u_2 + u_3x_3u_2) \\ \cos\theta x_3 + \sin\theta u_1x_2 - \sin\theta u_2x_1 + (1 - \cos\theta)(u_1x_1u_3 + u_2x_2u_3 + u_3x_3u_3) \end{bmatrix}$$

So the matrix $\mathbf{R}$ would be,

$$\mathbf{R} = \begin{bmatrix} \cos\theta + (1 - \cos\theta)u_1u_1 & -\sin\theta u_3 + (1 - \cos\theta)u_2u_1 & \sin\theta u_2x_3 + (1 - \cos\theta)u_3u_1 \\ \sin\theta u_3 + (1 - \cos\theta)u_1u_2 & \cos\theta + (1 - \cos\theta)u_2u_2 & -\sin\theta u_1 + (1 - \cos\theta)u_3u_2 \\ -\sin\theta u_2 + (1 - \cos\theta)u_1u_3 & \sin\theta u_1 + (1 - \cos\theta)u_2u_3 & \cos\theta + (1 - \cos\theta)u_3u_3 \end{bmatrix}.$$