# CS-E4850 Computer Vision
# Exercise Round 2

The problems should be solved before the exercise session and solutions returned via the MyCourses page in a single PDF file. Handwritten solutions are fine if they are scanned to PDF format and clear to read.

There are many exercises in this round but also partial solutions will be rewarded. Exercises 1-4 are a recap of the pinhole camera model presented during the first lecture. Problem 3 is based on 2.a but otherwise the problems can be solved in any order. A comprehensive explanation of the pinhole model is given in Chapter 6 of the book by Hartley & Zisserman.

**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.
(Hint: Use similar triangles and remember that the image plane is located at a distance $f$ from the projection center and is perpendicular to the optical axis, i.e. the z-axis of the camera coordinate frame.)

**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.
*b)* $u$ axis is parallel to $x$ axis and the angle between $u$ and $v$ axis is $\theta$.

**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.

**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 $3 \times 4$ 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 $\mathbf{R}, \mathbf{t}$.

**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 $3 \times 3$ matrix $\mathbf{R}$ be a 3-D rotation matrix, which rotates a vector $\mathbf{x}$ by the angle $\theta$ 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.

*b)* Derive the expressions for the elements of $\mathbf{R}$ as a function of $\theta$ and the elements of $\mathbf{u}$. (Hint: Write down the cross product $\mathbf{u} \times \mathbf{x}$ and the scalar product $\mathbf{u} \cdot \mathbf{x}$ in terms of the elements of vectors $\mathbf{u}$ and $\mathbf{x}$. Use the notations $\mathbf{u} = (u_1, u_2, u_3)^\top$ and $\mathbf{x} = (x_1, x_2, x_3)^\top$. You may also consult literature or public sources like Wikipedia.)