Computer Vision 1: Homework 1

- Prepare your solutions and bring them with you to the exercise class.
- There will be 1 programming task per sheet, and on average 3 other tasks per sheet.
- To pass the exercise class, you must solve the programming task in at least 50% of the homework sheets, **and** at least 50% of all other (pen & paper) homework tasks.
- A problem is only counted as solved if you come to the exercise, mark it solved, and present your solution if requested. No submissions by email, no handing in a solution on paper – these will not be accepted.
- Programming homeworks will be presented using a computer (show your code and run it). You can use your own laptop or the computer in the class.
- Other homework problems will be presented on the blackboard.
- If you cannot present a solution to a problem you marked as solved, or cannot explain your solution appropriately, you lose all marks for the current homework sheet. If this happens a second time, you fail the exercise class.

Programming task. Using Python and NumPy, solve the following tasks from the exercise sheet.

- Create a vector of size 10 with all elements equal to zero.
- Create a 3-by-3 matrix with the numbers from 0 to 8. Multiply all elements in the matrix by 3.
- Calculate the matrix product $A^T \cdot B$ where $A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$[1].
- Add the row vector $\begin{bmatrix} -1 & 1 \end{bmatrix}$ to each row of $A^T \cdot B$.
- Create a 3-by-3 matrix X with random values in the range $[0, 1)$. Create a binary matrix C of size 3-by-3 where the value of an element is `False` if the value of the corresponding element in X is less than 0.5, and `True` otherwise. Set all values in X at indices where C is `True` to -1.

Other tasks. For the two first tasks, there is no one correct answer. It is sufficient to present a justified argument for what you think.

1. The handwritten digit recognizer demo at `http://myselph.de/neuralNet.html` lets an user draw a digit, and then applies a neural network to recognize which digit was drawn. Visit the URL and click on "Display Preprocessing" so you will see what is provided as an input to the NN. Draw a digit and see if the system recognizes it.

   Now, try to draw a digit that a human can easily recognize, but which the neural network cannot correctly recognize. Why do you think the digit cannot be recognized?

[1] The correct answer should be $A^T \cdot B = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$.

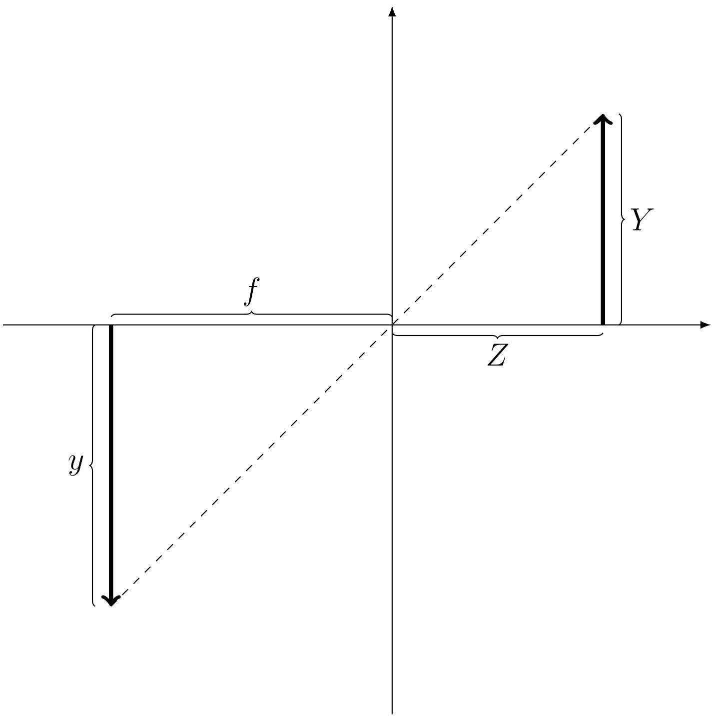

Figure 1: Pinhole camera. The pinhole is at the origin. An object with height Y is at distance Z from the pinhole. The image of the object appears on the image plane, which is at distance f from the pinhole. The dashed line shows a ray of light traversing from the object through the pinhole to the image plane.

2. Visit the Wolfram image identification project at `https://www.imageidentify.com/`. Find an image that gives a false recognition result when you upload it. Do you think this recognition problem is easier or harder than the digit recognition problem above? Why?

3. Consider the pinhole camera model shown in Figure 1. The pinhole where light passes through is located at the origin of the coordinate system. On the right hand side is an object being imaged. The object is at distance Z from the pinhole, and has a height Y. On the left hand side, at distance f from the pinhole, is the back of the camera with the image plane where the (upside-down) image will appear.

   Solve for the coordinate y on the image plane where the image of the object appears.

   Hint: relationship of the two right angle triangles formed by considering the ray traversing from the object through the pinhole to the image plane.