ICBV231 HW 1

Release Date: 23/11/2022

Submission Deadline: 06/12/2022, 23:59

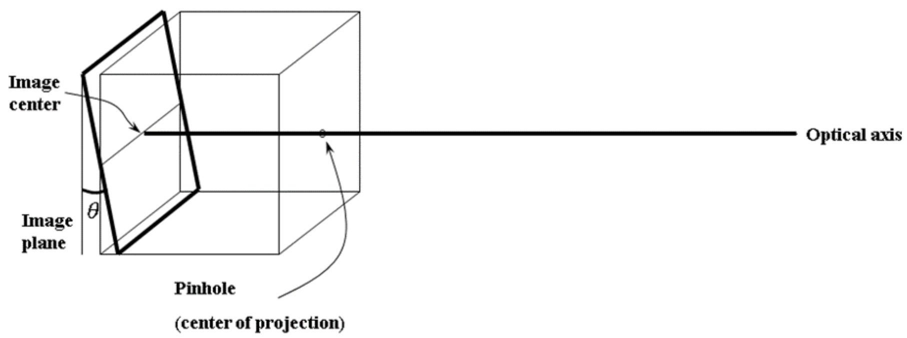
Instructions

Please read and follow these instructions carefully.

- To be able to submit your work you must first enroll in a group on Moodle.
- The assignment contains both programming tasks and a written part. The written part needs to be submitted in a single PDF file and the programming tasks must be answered in the provided Jupyter notebook. You need to submit a single zip file where:
  - The name of the file is "Assignment1-Group#.zip", where # should be replaced by your group number.
  - In the zip file put the written answers as a PDF file and the Jupyter notebook (.ipynb file) along with any additional files used by your code.
- The written assignment must be typed on the computer. Handwritten scanned documents **will not** be accepted.
- Make sure that the Jupyter notebook can be ran. Before submitting, restart runtime and run all cells. In Google Colab: 1) *Runtime > Restart and run all;* 2) *File > Download .ipynb*. In addition, make sure you submit any additional files used by your code (usually image files).
- Both the PDF containing the written answers and the Jupyter notebook must contain the id numbers and the names of the submitting students.

Question 1 – Perspective Projection

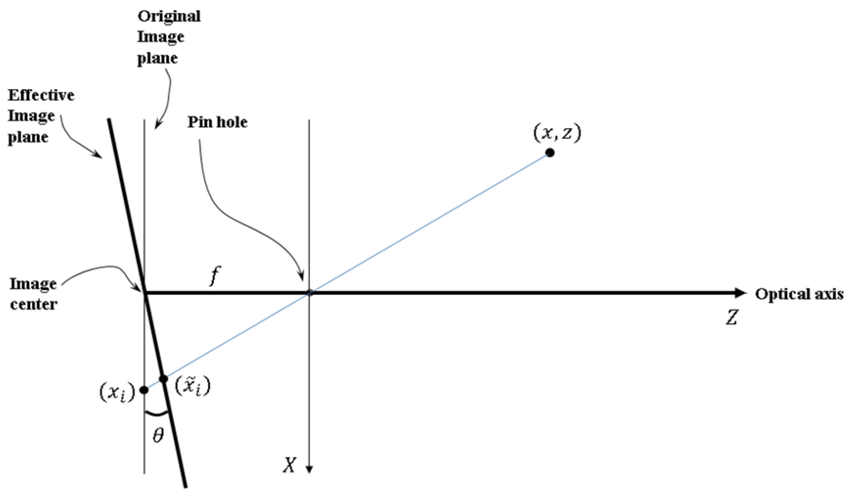
A malfunction occurred in the production line of the famous pinhole camera manufacturer HoleCam Ltd. As a result, the image planes of all the company cameras suffer from a θ degree deviation from the perpendicular direction of the optical axis, as shown in the diagram below.



The disappointed customers soon discovered that the perspective projection is not valid for their camera. Instead of asking the customers to return the defective cameras, the company decided to publish a new projection function which is suitable for the defective cameras.

<u>Your task</u> is to help the company engineer find this projection function.

To facilitate the analysis we will be discussing a 1D camera, that instead of an image plane has an 'image line' on which the projection x_i of a 2D world point (x, z) appears. The following diagram describes the cameras original state and the state after the malfunction:



Due to the malfunction, the world point (x, z) is projected onto a different point $\tilde{x}_i$ in the shifted 'image line'. <u>Your task</u> is to find the new projection function from (x, z) to $\tilde{x}_i$, and <u>elaborate</u> on how and why you developed this projection function. Assume that the origin of the 'image line' (i.e., coordinate x) is at its intersection point with the optical axis, and that the focal length (i.e., the length of the optical axis from the pinhole to its intersection with the 'image line') is f. In addition, assume θ is known.

Question 2 – Camera Calibration

In this question, you will perform a calibration of a camera of your choice (e.g., your mobile phone or computer camera). Your goal is to find the camera matrix using pairs of corresponding world and image points, from which the camera parameters and camera matrix can be estimated.

Section A - Chessboard corners

Take 20 pictures of a chessboard and use OpenCV methods cv2.findChessboardCorners and cv2.drawChessboardCorners. Use the template code provided in the Jupyter notebook to view the results over your pictures. Add the required code in the provided Jupyter notebook. **Hint**: Finding the corners needs to be performed over a grayscale image, drawing the corners is over an RGB image.

Section B - Create matching 3D points

As shown in the practical session, create a set of 3D points matching your board dimension and square size. Add the required code in the provided Jupyter notebook.

Section C - Perform the calibration

Perform the camera calibration using OpenCV cv2.calibrateCamera. Add the required code in the provided Jupyter notebook.
The method returns the following variables as output:

- cameraMatrix
- distCoeffs
- rvecs
- tvecs

Explain shortly what each of them mean.

Section D - Camera Matrix

You saw the following definition in class:

$$\underbrace{K \cdot F}_{M}$$

1. What element of the formula does the output parameter "cameraMatrix" refer to?
2. Obtain the extrinsic transformation matrix for each pair in (rvecs, tvecs). Write the code in the Jupyter notebook.
3. Obtain a corresponding matrix M for each of the pairs. Write the code in the Jupyter notebook.

Section E

In this section you will project the world points you defined in Section B on top of one of the images you used for the calibration process. On top of the image you chose:

- Plot the chessboard corners from section A using the openCV cv2.drawChessboardCorners method.
- Plot a projection of the world points defined in section B on top of the same image. Perform the projection using:
    1. The camera matrix you got from the calibration process.
    2. The OpenCV method cv2.projectPoints.

Is there a difference between the projection using the camera matrix and the projection using the OpenCV method? If so, why? Please explain shortly.

Section F

Choose another image from the images used in the calibration process. Draw a cube on top the upper left black square of the chessboard present in the chosen image.

1. **Hint:** First find out the world points of the cube and project them on top of the image as shown in section E. You can draw the cube skeleton using Matplotlib line plotting methods or OpenCV cv2.line.
2. Rotate the cube points 30° clockwise around one of the axes and display the result.

Perform the drawing in the Jupyter notebook.

Section G

A student at the course didn't have access to a camera. Instead of taking pictures, the student obtained a collection of chessboard images from multiple sources on the internet and used them instead. Is there a problem with this approach? If so, explain what the student did wrong.

Question 3 – Reading Material

The following sections include questions regarding the reading material which accompanies the course lectures (see the Lecture Notes page in the course website). Answer the following sections in writing, and keep your answers short.

Section A

This section refers to *The forty fold path to enlightenment*, chapter 5 from Richard Dawkins' book “*Climbing mount improbable”*, 1996.

According to Dawkins, what is the dilemma that evolution had to solve during the development of the eyes, and what was created as a solution to it?

Section B

This section refers to *Introduction to computer vision*, chapter 1 from Nalwa's book “*A guided tour of computer vision”* , 1993.

According to Nalwa “the purpose of computer vision is to infer the state of the physical world from the inherently noisy and ambiguous images of the world”. Given what is described in this chapter about the human visual system and abilities, and considering the main purpose of computer vision, should computer vision scientists attempt to emulate human vision in your opinion? Or in Nalwa’s words, “do we really wish to make machines see as we do”? Answer the question and present arguments in support of your position (with information taken from the text).

Section C

From what is explained in the short video that compares CCD and CMOS technologies, describe 4 central differences between CCD and CMOS.

Section D

This section refers to Marr's philosophy to vision research, chapter 1 from David Marr's book “*Vision - A Computational Investigation into the Human Representation and Processing of Visual Information* “, 1983.

According to Marr, any object can be represented in several different ways. Pick an object of your choosing (e.g., it can be a household object, like water tap), and describe 2 different possible representations of this object. For each representation, explain which information becomes explicit (i.e., how easy it is to do different things with it) and which information is pushed into the background (and is hard to recover) when using this representation.