

ENG301 - Computer Vision

Assignment 1

Demonstration deadline: 13 March 2025

Report submission deadline: 11:59 pm, 16 March 2025

Plagiarism note: This is an **individual** assignment. *Students are free to discuss among the groups. However, their programs should be written by themselves. If similarity is detected in their report or programs, penalties will be applied.*

Theory – 3 points

In many computer vision applications, the camera is not attached to a fixed frame but to a moving platform such as smartphones, cars, robots, drones, airplanes, etc. It is then important to calculate the relative position of the camera with respect to its surrounding objects as in the following question:

A workstation that includes a robotic arm and a camera has been set up as in Figure 1. The camera can see both the origin of the base coordinate system of a robot arm and the centre of the object to be manipulated by the robot.

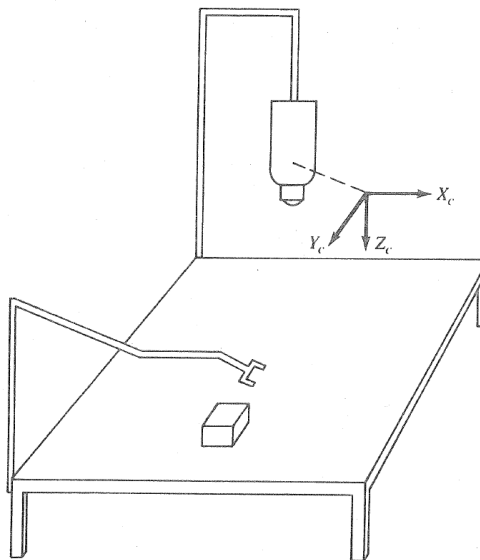


Figure 1: A workstation with a camera, a robot arm and an object.

If a coordinate system has been established at the centre of the object, then the object as seen by the camera can be expressed by the homogenous transformation matrix T_1 .

$$T_1 = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 10 \\ 0 & 0 & -1 & 9 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Similarly, the origin of the robot base coordinate system as seen by the camera can be represented by the homogenous transformation matrix T_2 .

$$T_2 = \begin{bmatrix} 1 & 0 & 0 & -10 \\ 0 & -1 & 0 & 20 \\ 0 & 0 & -1 & 10 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- After the equipment has been setup and these coordinate systems have been established, someone rotated the camera 90° about the z axis of the camera. What is the position and orientation of the camera with respect to the robot's base coordinate system?
- After the rotation happened in a), the object is also rotated by 90° about the x axis of the object and translated by 4 units along the rotated y axis. What is the position and orientation of the object with respect to the robot's base coordinate system? What is the position and orientation of the object with respect to the rotated camera coordinate system?

Note: The answer can be written in the **homogenous transformation matrix** form since retrieving the angles from a homogenous transformation matrix is not a trivial problem (see more: https://en.wikipedia.org/wiki/Rotation_matrix)

Programming – 7 points

Task 1 – 2 points

Write a program to carry out the calibration and print out the camera, distortion, rotation, and translation matrices for the webcam of your laptop/tablet. Contact the instructor if you need to borrow a camera.

Note: You will need to print out a chessboard and take several pictures of it from the camera being calibrated.

Task 2 - 1 points

Develop an augmented reality application using the ArUco markers. The suggested steps are as follows:

- Generate and print out the ArUco markers.
- Stick them to the corners of an object of your choice (e.g., a picture frame).
- Record a short video of the object with markers.
- Write a program to replace the area within the marker with a video of your choice. The result is something similar to my example at class.

Task 3 – 2 points

Write a program to track the eye of a user sitting in front of the computer. The program should have at least the following features:

- Track the pupil center of the eye.
- Show the frame rate.

Note:

- Using software libraries like MediaPipe is not allowed
- The program does not have to perfectly track the pupil center. It is accepted if the program cannot detect the pupil in certain frames or circumstances.

Task 4 – 2 points

In some applications, a camera that includes both color and depth sensors called an RGB-D camera can be used to capture both color and depth information of the scene. Since the color and depth sensors are physically separated, the color and depth frames are not aligned. Therefore, an alignment process using the intrinsic and extrinsic parameters of the camera is needed.

Given the intrinsic and extrinsic parameters of a color-depth camera in the camera_params.json file. Write a program to carry out the following tasks:

- Align the color and depth frames (color.png, depth.png) of the camera
- Create a pointcloud from the aligned depth and color points and visualize it. The result would look like the figure below:

