Homework 1

Kinematics and Camera Calibration

Due October 17 end of day

1 Forward and Inverse Kinematics

We have designed a 3 degree of freedom robot that is planar, i.e., it moves in the xy-plane, as shown in the figure below. The robot has three controllable axes q_0, q_1, q_2. The first link is 5 cm from the world coordinate reference, and each link is 10 cm long. Each degree of freedom has a range of $\pm 90°$ such that the arm forms a straight line in the $+y$ direction when all joints are at $0°$.

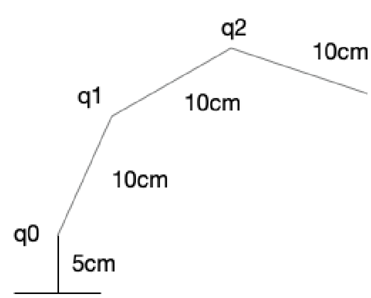

Figure 1: Planar 3DoF robot arm

a. **Forward Kinematics.** Derive the position of the end effector $\boldsymbol{x} = [x, y]^\top$ as a function of joint angles $\boldsymbol{q} = [q_0, q_1, q_2]^\top$ and its motion model $\dot{\boldsymbol{x}} = T\dot{\boldsymbol{q}}$. Draw out the world and joint reference frames in Figure 1.

   You are permitted to use symbolic solvers (e.g. MATLAB Symbolic Toolbox, Mathmatica, sympy) in your derivation, but your report must clearly explain how you got your solution. If you choose to derive by hand, you might find the following trig identities useful when simplifying your expressions:

$$\cos(\alpha)\cos(\beta) \pm \sin(\alpha)\sin(\beta) = \cos(\alpha \mp \beta)$$
$$\cos(\alpha)\sin(\beta) \pm \sin(\alpha)\cos(\beta) = \sin(\alpha \pm \beta)$$

b. **Inverse Kinematics.** Determine a joint angle configuration that positions the robot's end-effector at the point $(10, 15)$ (hint: the null space of T might be handy).

For those looking for guidance, we found H. Harry Asada's Planar Kinematics notes [1] to be particularly useful. (Note that you **do not** need to derive the general, algebraic IK formula for this problem, you only need to find a single valid solution.)

2 Camera Calibration

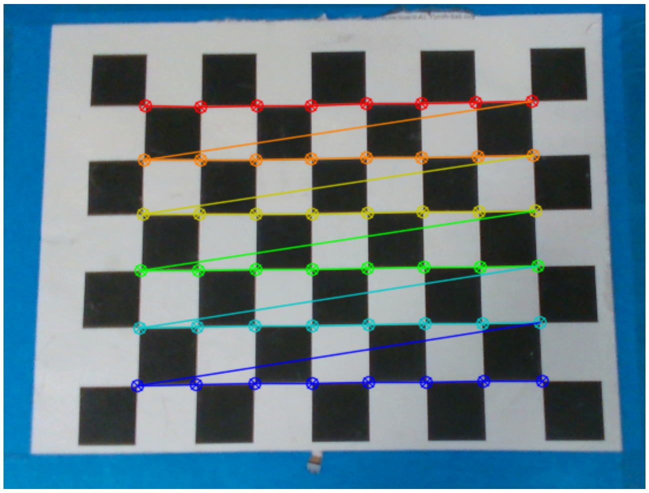

Figure 2: Checkerboard with colored corners

Checkerboard calibration is the standard camera calibration method(Fig. 2) We will model our camera with an idealized pinhole model, assuming no skew or radial distortion to our lens. To determine the 3D world coordinates $\boldsymbol{x}$ from a reference point $\boldsymbol{u}$, we need the extrinsic matrix $[\mathbf{R} \,|\, \boldsymbol{t}]$ that maps the world reference frame to the camera reference frame and the the intrinsic matrix $\mathbf{A}$ that projects it onto the 2D image plane, such that $\boldsymbol{u} = \mathbf{A}\,[\mathbf{R} \,|\, \boldsymbol{t}]\,\boldsymbol{x}$ with $\boldsymbol{x}$. This can be expanded (using homogeneous coordinates) as

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

where f_x, f_y are the camera focal lengths and c_x, c_y is the image center. s is an arbitrary scale factor that is often omitted for brevity.

The goal of this problem is to find the values of the intrinsic and extrinsic matrices for a set of calibration images and estimate the trajectory of the camera throughout the calibration. We will be following Zhang's algorithm [2], which can be accessed at the URL listed in the bibliography. It is highly recommended to reference this paper throughout the course of this problem. The relevant information can be found primarily in Section 2-3.1 and Appendices A-C, although it is encouraged to skim the rest for full comprehension.

Those looking for further guidance are encouraged to reference the tutorial by Wilhelm Berger [3], starting with Section 3. You are not expected to do any nonlinear optimization or accounting for lens distortion for this assignment, so you can skip any discussion of these techniques.

1. **Plane-to-Plane Homography**

   The file **imgpoints.txt** contains a list of checkerboard key points 2D pixel coordinates. These key points were extracted from 9 checkerboard images. Each image has 8 corners in x-axis and 6 corners in y-axis, totalling $8 \times 6 = 48$ key points. The resolution of each image is 640×480. Assuming each checkerboard square measures 25 mm per side, you will first derive the 2D coordinates of each reference point on the chessboard plane w.r.t an origin placed at the bottom left point. This will be coordinates in the World Frame.

   Because all of the reference points lie on a plane, we can describe the transformation from the checkerboard plane to the image plane as a 2D homography that acts on homogeneous coordinates. This homography $\mathbf{H}$ is a 3×3 matrix given by $\boldsymbol{u} = \mathbf{H}\boldsymbol{p}$ where $\boldsymbol{p} = (x, y, 1)$ is the checkerboard coordinate and $\boldsymbol{u} = (u, v, 1)$ is the image coordinate.

   a. Let $\boldsymbol{h}$ be the flattened 9-vector containing the elements of $\mathbf{H}$. Given a point correspondence $\boldsymbol{p}_i \mapsto \boldsymbol{u}_i$, rearrange $\boldsymbol{u}_i = \mathbf{H}\boldsymbol{p}_i$ into a system of linear equations of the form $\mathbf{M}_i\,\boldsymbol{h} = \mathbf{0}$. In the report, answer following questions:

      i. How many linear equations do we get from one point correspondence?

      ii. How many do we need to fully define $\mathbf{H}$?

      iii. Given the number of reference points on our chessboard, is each image's homography our system over-, under-, or fully-defined?

   b. Construct $\mathbf{M}$ by stacking $\mathbf{M_i}$ for each reference point within one image. Solve the resulting homogeneous linear least squares problem $\mathbf{M}\,\boldsymbol{h} = \mathbf{0}$ for $\boldsymbol{h}$ and record the homography matrix $\mathbf{H}$ for each image. for each homography matrix we can reproject the world coordinates back to the pixel coordinates, and calculate the reprojection error $u - Hp$ in euclidean distance. Report the total reprojection error for all images.

2. **Intrinsic Matrix** The intrinsic matrix $\mathbf{A}$ of the camera is related to the homography by the following constraints

$$\boldsymbol{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \boldsymbol{h}_2 = 0$$
$$\boldsymbol{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \boldsymbol{h}_1 = \boldsymbol{h}_2^T \mathbf{A}^{-T} \mathbf{A}^{-1} \boldsymbol{h}_2$$

where $\boldsymbol{h}_i$ is the ith column vector of $\mathbf{H}$ (see paper for derivation).

Let $B = \mathbf{A}^{-T} \mathbf{A}^{-1}$ be a 3×3 symmetric matrix whose elements define a 6-vector $\boldsymbol{b} = [b_{11}, b_{12}, b_{13}, b_{22}, b_{23}, b_{33}]$.

   a. Given a homography $\mathbf{H}_i$, use these constraints to form a system of linear equations of the form $\mathbf{L}_i \, \boldsymbol{b} = \mathbf{0}$. In the report, answer following questions:

      i. How many linear equations do we get from one homography?

      ii. How many images do we need to fully define $\mathbf{b}$?

      iii. Given the number of images provided, is $\mathbf{b}$ over-, under-, or fully-defined?

   b. Construct $\mathbf{L}$ by stacking $\mathbf{L}_i$ for each image in the reference set. Solve the resulting homogeneous linear least squares problem $\mathbf{L}_i \, \boldsymbol{b} = \mathbf{0}$ for $\mathbf{b}$.

   c. Using the equations found in Appendix B, calculate the intrinsic matrix $\mathbf{A}$.

   d. In report, briefly explain each entry you obtained for $\mathbf{A}$. Is this what you expected for a "perfect" camera? What are some possible reasons for the difference?

3. **Calculating Extrinsic Matrix from Intrinsic Matrix**

   a. Given homography $\mathbf{H}$ and intrinsic matrix $\mathbf{A}$, use the procedure outlined in Section 3.1, calculate the extrinsic matrix $[\mathbf{R} \,|\, \boldsymbol{t}]$ for each image.

   b. In the report, use each image's extrinsic matrix to plot the trajectory of the camera. (Hint: $\mathbf{R}\boldsymbol{C} + \boldsymbol{t} = 0$, where $\boldsymbol{C}$ is the camera coordinate)

References

[1] H. H. Asada, "Planar kinematics," course notes for Introduction to Robotics, Massachusetts Institute of Technology, Dept. of Mechanical Engineering, Cambridge, MA, USA, Fall 2005, https://ocw.mit.edu/courses/2-12-introduction-to-robotics-fall-2005/c8828a16e71c246b78461dd0596b983f_chapter4.pdf.

[2] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 1. IEEE, 1999, pp. 666–673, https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr98-71.pdf.

[3] W. Burger, "Zhang's camera calibration algorithm: In-depth tutorial and implementation," University of Applied Sciences Upper Austria, School of Informatics, Communications and Media, Dept. of Digital Media, Hagenberg, Austria, Tech. Rep. HGB16-05, May 2016, https://www.researchgate.net/publication/303233579_Zhangs_Camera_Calibration_Algorithm_In-Depth_Tutorial_and_Implementation.