Computer Vision 2024
(CSE344/ CSE544/ ECE344/ ECE544)
Assignment-2

Max Marks (UG/PG): 110 / 110 **Due Date**: 24/03/2024, 11:59 PM

Instructions

- Keep collaborations at high-level discussions. Copying/plagiarism will be dealt with strictly.
- Your submission should be a single zip file **Roll_Number_HW[n].zip**. Include only the **relevant files** arranged with proper names. A single **.pdf report** explaining your codes with relevant graphs, visualization and solution to theory questions.
- Remember to **turn in** after uploading on Google Classroom. No justifications would be taken regarding this after the deadline.
- Start the assignment early. Resolve all your doubts from TAs during their office hours **two days before the deadline.**
- Kindly **document** your code. Don't forget to include all the necessary plots in your report.
- All [**PG**] questions, if any, are **optional for UG** students but are **mandatory for PG** students. UG students will get BONUS marks for solving that question.
- All [**BONUS**] questions, if any, are optional for all the students. As the name suggests, BONUS marks will be awarded to all the students who solve these questions.
- Your submission **must include a single python (*.py*) file for each question**. You can submit *.ipynb* along with the *.py* files. Failing to follow the naming convention or not submitting the python files will incur a **penalty**.

---

1. (15 points) Consider a vector $(2, 5, 1)^T$, which is rotated by $\pi/2$ about the Y-axis, followed by a rotation about X-axis by $-\pi/2$ and finally translated by $(-1, 3, 2)^T$.
    1. (4 points) What is the coordinate transformation matrix in the case?
    2. (3 points) Find the new coordinates of this vector. Where does the origin of the initial frame of reference get mapped to?
    3. (4 points) What is the direction of the axis of the combined rotation in the original frame of reference and what is the angle of rotation about this axis?
    4. (4 points) Using Rodrigues formula, show that you achieve the same rotation matrix as you sequentially apply the two rotations.

2. (5 points) Consider a rotation $\mathbf{R}$ of angle θ about the axis $\mathbf{u}$ (a unit vector). Show that $\mathbf{Rx} = \cos\theta\mathbf{x} + \sin\theta(\mathbf{u}\times\mathbf{x}) + (1-\cos\theta)(\mathbf{u}^\top\mathbf{x})\mathbf{u}$

3. (10 points) [PG]

   The image formation process can be summarized in the equation $\mathbf{x} = \mathbf{K}[\mathbf{R}|\mathbf{t}]\mathbf{X}$, where $\mathbf{K}$ is the intrinsic parameter matrix, $[\mathbf{R}|\mathbf{t}]$ are the extrinsic parameters, $\mathbf{X}$ is the 3D point and $\mathbf{x}$ is the image point in the homogeneous coordinate system. Consider a scenario where there are two cameras ($\mathbf{C}_1\&\mathbf{C}_2$) with intrinsic parameters $\mathbf{K}_1\&\mathbf{K}_2$ and corresponding image points $\mathbf{x}_1\&\mathbf{x}_2$ respectively. Assume that the first camera frame of reference is known and is used as the world coordinate frame. The second camera orientation is obtained by a pure 3D rotation $\mathbf{R}$ applied to the first camera's orientation. Show that the homogeneous coordinate representation of image points $\mathbf{x}_1$ and $\mathbf{x}_2$ of $\mathbf{C}_1$ and $\mathbf{C}_2$ respectively, are related by an equation $\mathbf{x}_1 = \mathbf{H}\mathbf{x}_2$, where $\mathbf{H}$ is an invertible 3×3 matrix. Find $\mathbf{H}$ in terms of $\mathbf{K}_1, \mathbf{K}_2$ & $\mathbf{R}$.

4. (40 points) **Camera Calibration:**

   Refer to the following tutorials on camera calibration: Link1 and Link2. Place your camera (laptop or mobile phone) stationary on a table. Take the printout of a chessboard calibration pattern as shown in the links above and stick it on a hard, planar surface. Click $\sim$ 25 pictures of this chessboard pattern in many different orientations. Be sure to cover *all degrees of freedom* across the different orientations and positions of the calibration pattern. Make sure that each image *fully* contains the chessboard pattern. Additionally, the corners in the chessboard pattern should be detected automatically and correctly using appropriate functions in the OpenCV library. Include the final set of images that you use for the calibration in your report.

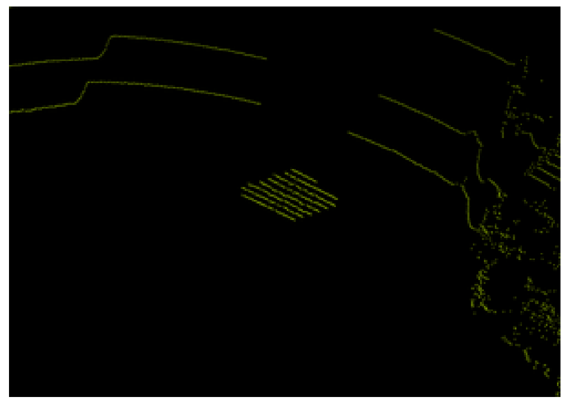
   1. (5 points) Report the estimated intrinsic camera parameters, i.e., focal length(s), skew parameter and principal point along with error estimates if available.
   2. (5 points) Report the estimated extrinsic camera parameters, i.e., rotation matrix and translation vector for each of the selected images.
   3. (5 points) Report the estimated radial distortion coefficients. Use the radial distortion coefficients to undistort 5 of the raw images and include them in your report. Observe how straight lines at the corner of the images change upon application of the distortion coefficients. Comment briefly on this observation.
   4. (5 points) Compute and report the re-projection error using the intrinsic and extrinsic camera parameters for each of the 25 selected images. Plot the error using a bar chart. Also report the mean and standard deviation of the re-projection error.
   5. (10 points) Plot figures showing corners detected in the image along with the corners after the re-projection onto the image for all the 25 images. Comment on how is the reprojection error computed.
   6. (10 points) Compute the checkerboard plane normals $\mathbf{n}_i^C,\ i \in \{1,..25\}$ for each of the 25 selected images in the camera coordinate frame of reference (O^c).

5. (40 points) **Camera-LIDAR Cross-Calibration:**

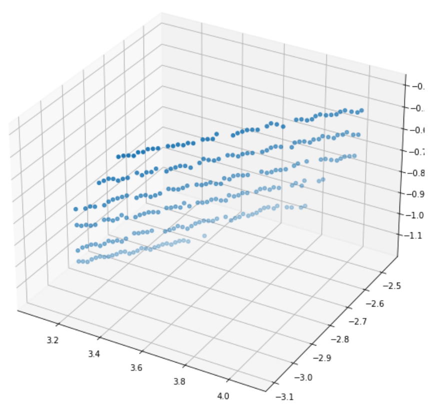
   Use this link to download the complete dataset. The dataset includes the RGB images and their corresponding LIDAR scans, the camera calibration parameters (intrinsic parameters, extrinsic parameters and distortion coefficients as returned by OpenCV) and the checkerboard plane normals in the camera reference frame, $\mathbf{n}_i^C$. The dataset is acquired by a monocular camera and a LIDAR sensor mounted on top of an autonomous vehicle. You can select any 25 corresponding images and LIDAR data points for this assignment. The size of the chessboard box is given to be 108mm. A sample image and the corresponding LIDAR scan are shown in Figure 1a and 1c. We have pre-processed the LIDAR scan points and extracted the LIDAR points on the chessboard as .pcd files. A sample of the extracted file is shown in Figure 1b. To cross-calibrate the LIDAR and the camera is to find the invertible Euclidean transformation (3D rotation and 3D translation) between the camera and LIDAR coordinate frames. Follow the steps below to complete the LIDAR-Camera cross-calibration.

   1. (10 points) Compute the chessboard plane normals $\mathbf{n}_i^L, \;\; i \in \{1, \ldots, 25\}$ and the corresponding offsets $d_i, \; i \in \{1, \ldots, 25\}$ using the planar LIDAR points in each of the pcd files. You can estimate these by making use of singular value decomposition.
   2. (10 points) Now that you have the plane normals $\mathbf{n}_i^C$ and $\mathbf{n}_i^L$ in camera and LIDAR frame of reference respectively for all the selected images, derive the set of equations that you would use to solve for estimating the transformation ${}^{\mathrm{C}}\mathbf{T}_{\mathrm{L}} = [{}^{\mathrm{C}}\mathbf{R}_{\mathrm{L}}|{}^{\mathrm{C}}\mathbf{t}_{\mathrm{L}}]$, i.e. the transformation from the LIDAR frame to the camera frame. Explain how the rotation and translation matrices are calculated. [*Hint*: You may refer to this thesis (Sec. 5) for deriving the necessary equations.]
   3. (5 points) Using the above equations, implement the function that estimates the transformation ${}^{\mathrm{C}}\widetilde{\mathbf{T}}_{\mathrm{L}}$. Recall that the rotation matrix has determinant +1.
   4. (5 points) Use the estimated transformation ${}^{\mathrm{C}}\widetilde{\mathbf{T}}_{\mathrm{L}}$ to map LIDAR points to the camera frame of reference, then project them to the image plane using the intrinsic camera parameters. Are all points within the checkerboard pattern's boundary in each image?
   5. (10 points) Plot the normal vectors $\mathbf{n}_i^L, \mathbf{n}_i^C, {}^C\mathbf{R}_L\mathbf{n}_i^L$ for any 5 image and LIDAR scan pairs. Compute the cosine distance between the camera normal $\mathbf{n}_i^C$ and the transformed LIDAR normal, ${}^C\mathbf{R}_L\mathbf{n}_i^L$ for *all* 38 image and LIDAR scan pairs, and plot the histogram of these errors. Report the average error with the standard deviation.

(a)



(b)



(c)

Figure 1: LIDAR-camera calibration. **(a)** Camera Frame. **(b)** Screenshot of LIDAR scan. The dense points in the center of the image are the points on the chessboard pattern. **(c)** Chessboard points extracted from the LIDAR Frame.