PROBLEM SET 5

16822 GEOMETRY-BASED METHODS IN VISION (FALL 2024)
https://piazza.com/cmu/fall2024/16822b

OUT: Nov. 12, 2024
DUE: Nov. 19, 2024 11:59 PM
Instructor: Shubham Tulsiani
TAs: Easton Potokar, Jianjin Xu

START HERE: Instructions

- **Collaboration policy:** All are encouraged to work together BUT you must do your own work (code and write up). If you work with someone, please include their name in your write up and cite any code that has been discussed. If we find highly identical write-ups or code without proper accreditation of collaborators, we will take action according to university policies, i.e. you will likely fail the course. See the Academic Integrity Section detailed in the initial lecture for more information.
- **Submitting your work:**
  - We will be using Gradescope (https://gradescope.com/) to submit the Problem Sets. Please use the provided template. Submissions can be written in LaTeX, or submitted as a scanned PDF. Regrade requests can be made, however this gives the TA the opportunity to regrade your entire paper, meaning if additional mistakes are found then points will be deducted. Each derivation/proof should be completed on a separate page. For short answer questions you **should** include your work in your solution.
  - For questions where you must fill in a blank, please make sure your final answer is fully included in the given space. You may cross out answers or parts of answers, but the final answer must still be within the given space. We accept either LaTex pdfs or scanned documents as long as the location of each question is annotated properly.
- **Materials:** The data that you will need in order to complete this assignment is posted along with the writeup and template on Piazza.

1 Parametrizing Rotations [20 pts]

1. **[6 pts]** Are these statements true or false?

   (a) Given a quaternion q, $-q$ represents its inverse rotation.

   (b) Given quaternions q_1 and q_2, the composition of these 2 rotations is $q_1 + q_2$.

   (c) Given a unit quaternion q such that $||q|| = 1$, there is a unique rotation it corresponds to.

2. **[4 pts]** Given a quaternion $q = [\sqrt{3}/2, 1/2, 0, 0]$, convert it to the axis-angle representation.

3. **[4 pts]** In Lecture 20, we introduced a parameterization of rotation that $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3] = F(\mathbf{v}_1, \mathbf{v}_2)$. Specifically,

$$\mathbf{r}_1 = \text{normalize}(\mathbf{v}_1), \qquad \mathbf{r}_2 = \text{normalize}(\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{r}_1)\mathbf{r}_1), \qquad \mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2.$$

given $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$. What's the output $\mathbf{R}' = F(\mathbf{v}_1', \mathbf{v}_2')$ if $\mathbf{v}_1' = \mathbf{v}_1, \mathbf{v}_2' = -\mathbf{v}_2 + \mathbf{v}_1$? Represent the output $\mathbf{r}_1', \mathbf{r}_2', \mathbf{r}_3'$ using $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$.

4. **[6 pts]** Given corresponding 3D points $\{(\mathbf{p}_i, \mathbf{q}_i)\} \in (\mathbb{R}^3, \mathbb{R}^3)$, we want to find the optimal rigid transform. Assume the points are already mean-centered so we only need to find the rotation. For every pair of points $(\mathbf{p}_i, \mathbf{q}_i)$, both points are located on one of the x, y and z axes. Namely,

$$\begin{aligned} \forall i, \quad & \mathbf{p}_i = [u_i, 0, 0], \quad \mathbf{q}_i = [v_i, 0, 0] \\ \text{or} \quad & \mathbf{p}_i = [0, u_i, 0], \quad \mathbf{q}_i = [0, v_i, 0] \\ \text{or} \quad & \mathbf{p}_i = [0, 0, u_i], \quad \mathbf{q}_i = [0, 0, v_i] \end{aligned}$$

(a) **[2 pts]** What are the possible rotations intuitively?

(b) **[4 pts]** Prove they are the **only** answers mathematically.

2 MVS and Learning-based SfM [20pts]

5. **[4 pts]** Assume two cameras on a stereo rig, separated by a baseline of 20cm, are looking at an object from 5 meters away. Suppose that for a specific 3D point on the object, the disparity between its pixel coordinates for the two images is 5 pixels. How does the disparity change if:

   (a) the baseline between the two cameras is increased to 40 cm?

   (b) the camera rig is set to 10 meters away from the object but the baseline between the two cameras remains 20 cm?

6. **[4 pts]** How many potential keypoints does superpoint predict **at most** from a 128×128 image?

7. **[6 pts]** Are these statements true or false?

   (a) A larger window size in SSD always yields better disparity estimates.

   (b) When comparing images with varying lighting, sum of Absolute differences should be preferred over normalized cross correlation for accurate disparity estimates.

   (c) In fronto-parallel plane sweep stereo, we need the neighboring cameras to have the same orientation (Rotation) as the camera for which depth is being estimate.

8. **[6 pts]** In fronto-parallel plane sweep stereo, a key step is to compute the homography that warps pixels from a neighboring image onto the current image under an assumed planar geometry for the scene. Assume that the current image has a camera matrix $K[I|0]$, and its neighbor has a camera matrix $K'[R|\mathbf{t}]$. Derive the formula for the planar homography corresponding to the plane $z = d$ (i.e. a fronto-parallel plane a distance d away from the first camera).

Attendance Question: Among Lectures 20-22, how many did you attend in person?

Collaboration Questions Please answer the following:

1. Did you receive any help whatsoever from anyone in solving this assignment?

   ◯ Yes

   ◯ No

   - If you answered 'Yes', give full details:
   - (e.g. Jane Doe explained to me what is asked in Question 3.4)

2. Did you give any help whatsoever to anyone in solving this assignment?

   ◯ Yes

   ◯ No

   - If you answered 'Yes', give full details:
   - (e.g. I pointed Joe Smith to section 2.3 since he didnt know how to proceed with Question 2)

3. Did you find or come across code that implements any part of this assignment ?

   ◯ Yes

   ◯ No

   - If you answered 'Yes', give full details (book & page, URL & location within the page, etc.).