**Computer Vision 2024**
**(CSE344/ CSE544/ ECE344/ ECE544)**
**Assignment-3**

**Max Marks (UG/PG)**: 45/ 50 **Due Date**: 14/04/2024, 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. (10 points) Theory Questions on Epipolar Geometry

    1. (3 points) The Essential matrix $\mathbf{E} = [\mathbf{t}_\times]\mathbf{R}$, where $\mathbf{t}$ and $\mathbf{R}$ are the translation vector and the rotation matrix. The epipoles are the points at the intersection of the baseline (the line connecting the two camera centres) and the image planes. The epipoles also happen to be the null-space and the left null-space of $\mathbf{E}$. Find the epipoles in terms of $\mathbf{t}$ and $\mathbf{R}$. Show your work.
    2. (2 points) Assume that a stereo camera setup has no relative rotation (i.e., $\mathbf{R} = \mathbf{I}$) and the translation is purely horizontal (i.e., $\mathbf{t} = [t_x, 0, 0]^\top$. Construct the Essential matrix and show that the *corresponding* points (i.e., image points in the two cameras that are projections of the same 3D point) will always have the same $y$-coordinate.

3. [PG] (5 points) Stereo rectification is the process of transforming the two images obtained from an arbitrary stereo camera setup such that their epipolar lines are parallel *and* horizontal (equivalent to the two cameras having no relative rotation and translation only along the horizontal direction). This amounts to applying a rotation to the image points such that the epipoles get mapped to points at infinity along the x axis. Refer to this slide deck[1] and derive the expression for the rotation matrix ($\mathbf{R}_{rect}$) required for the stereo rectification.

2. (40 points) **Panorama Generation**
Download the dataset from link. The dataset contains set of images that will be used for panorama generation. For the steps 1 to 5 you are required to use only the first two images from the set.

   1. (5 points) Keypoint detection: Extract the keypoints and descriptors from the first two images using the SIFT algorithm. SIFT (Scale-Invariant Feature Transform) is a computer vision algorithm used for feature detection and description. After extracting the keypoints and descriptors, draw them overlaid on the original images to visualize and verify their correctness.

   2. (5 + 5 points) Feature matching: Match the extracted features using two different algorithms: BruteForce and FlannBased . BruteForce is a simple algorithm that matches features by comparing all the descriptors of one image with all the descriptors of the other image. FlannBased (Fast Library for Approximate Nearest Neighbors) is a more efficient algorithm that uses a hierarchical structure to speed up the matching process. After performing the matching, display the matched features by drawing lines between them.

   3. (5 points) Homography estimation: Compute Homography matrix using RANSAC. RANSAC (Random Sample Consensus) is an iterative algorithm used for robust estimation of parameters in a mathematical model. The homography matrix is used to align the two images so that they can be stitched together to form a panorama.

   4. (5 points)Perspective warping: Perspective warping is a process that transforms the perspective of an image so that it appears as if it was taken from a different viewpoint. Warp any pair of images (with overlapping field of view) using their respective homography matrices and display the left and right images side-by-side. These *warped* images will be part of your panorama. Display the images without cropping the images or stitching them (as asked in the next part).

   5. (5 points) Stitching: Two images need to be stitched together to form a panorama. Display the final panorama without any cropping or blending, along with the panorama obtained after cropping and blending.

   6. (10 points) Multi-Stitching: Perform multi stitching for all the images in the folder and display the final result. (Hint: Use the function implemented for Stitching).

[1] https://www.cs.cmu.edu/ 16385/s17/Slides/13.1_Stereo_Rectification.pdf