Instructions

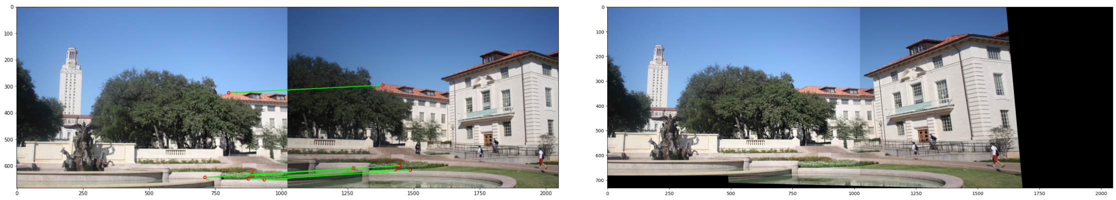
For each question, explain your motivations, the impact of various operations, and visualize the intermediate outputs. If you are unable to solve a question, show your attempt and explain your motivation. You may use any built-in APIs of OpenCV. Results are graded qualitatively and do not need to be reproductions of example outputs.

Kai Xu is the TA in charge. Post questions on Canvas or attend the FAQ during the Lab on Apr. 11 or 12.

Please Hand In to Canvas > Assignments > Assignment2:

- A pdf report with your answers and visualizations of the image outputs embedded in the report. Name your report Assignment2Report_AXXX.pdf where AXXX is your student number.
- Your source code in the form of a python notebook and intermediate and final image files a .zip named Assignment2Code_AXXX.zip, where AXXX is your student number.

Part 1: Keypoint Matching & Homography



(a) `building_left.jpg` (b) `building_right.jpg` (c) stitched image

Figure 1. The left and right images can be stitched into (c) based on detected feature matches.

Keypoint Match	A	B	C	D	E
Left Image Coordinates	(936, 652),	(718, 646),	(870, 349),	(889, 634),	(894, 623),
Right Image Coordinates	(464, 616),	(254, 611),	(424, 325),	(422, 599),	(430, 588),

Table 1. 5 Keypoint matches, with coordinates given in (column, row).

1. Consider the keypoint matches in Table 1. Compute and compare the homographies and visualize the stitched results based on two keypoint sets: Set 1, with keypoints {A, B, C, D} and Set 2 with {A, B, D, E}. Explain why the stitching from Set 2 fails. **(2 marks)**

2. Standard SIFT is applied to greyscale images. Consider the colour image in Fig.2(a) and three greyscale versions in Fig.2(b) – (d). Fig.2(b) uses the conventional weighting $[W_R, W_G, W_B]$ = [0.299, 0.587, 0.114]; Fig.2(c) scales the intensity of Fig.2(b) by a factor of 0.5; Fig.2(d) uses $[W_R, W_G, W_B]$ = [0.114, 0.299, 0.587].

   a. For the greyscale images (b), (c) and (d), detect SIFT keypoints and compute SIFT descriptors using the default settings of OpenCV. Visualize the keypoints as a scatterplot, similar to the sample in Fig. 2(e). Using the results of (b) as a reference, compare and comment on the number of keypoints and their locations for image (c) and (d). Explain any differences that arise. **(2 marks)**

   b. Consider image pair 1, from images (b) & (c) and image pair 2, from (b) & (d). Match the keypoints using `BFSMatcher.knnMatch()` [1] and a ratio of 0.5 for the ratio test. Visualize the *matched* keypoint locations, again similar to Fig. 2(e). Compare and comment on the number of matched keypoints and their locations for the two image pairs. Explain any differences that arise. **(2 marks)**

(a) `flowers.png` (b) `flowersbw.png` (c) `flowersbw_dark.png` (d) `flowersbw_w2.png` (e) keypoint visualization

Figure 2. SIFT matching on different greyscale image pairs.

[1] https://docs.opencv.org/4.x/dc/dc3/tutorial_py_matcher.html

Part 2: Tracking

Consider the sequence of pedestrian images in *data/Crowd_PETS09*:

(a) frame 0 (b)frame 25 (c) frame 50

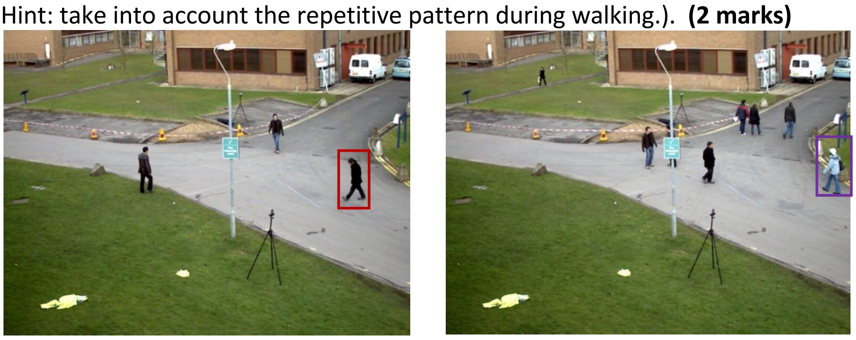
Figure 3. Image Samples for Crowd_PETS09 dataset

4. Track each person in frame 0 throughout the sequence using two off-the-shelf trackers[2] and answer the following questions. **(2 marks)**
   a. For your selected trackers, how are they initialized if a new person were to enter the scene?
   b. Visualize the resulting tracks for the two trackers using the provided code in `visualization.py` for the three pedestrians from frame 0. Comment on qualitative differences between the tracks.
   c. Multi-object tracking precision (MOPT) and multi-object tracking accuracy (MOTA) are two commonly used evaluation measures for tracking[3]. Compare the MOPT and MOTA of the two trackers based on the ground truth given in *data/gt.csv* and *data/gt_format.txt*.
   d. For each tracker, identify a failure case and analyse the potential cause(s) of the failure.

5. Figure 4 illustrates an ID shift failure, i.e., the tracker jumps to a different person when they cross paths. ID shifts occur when the two tracked individuals are similar in appearance. Propose a method to prevent ID shift. State the assumptions, describe the algorithm steps, and discuss possible drawbacks. **(1 mark)**

Figure 4. Shift of the Tracking Box

6. Estimate the number of steps taken by the individual marked with the red box in Figure 5a between frames 0 and 150. (Hint: take into account the repetitive pattern during walking.). **(2 marks)**



(a) Frame 0: Step counting (b) Frame 117: Distance estimation

Figure 5. Subjects for step counting and distance estimation.

7. Estimate the physical distance of the trajectory walked by the woman in the light blue jacket (Fig. 5b purple bounding box) between frames 117 to 221. Solve this problem by projecting her tracked trajectory onto a virtual ground plane, where 1 pixel represents a physical distance of 10cm. Use the following homography for mapping the image plane to the ground plane: **(2 marks)**

$$H = \begin{bmatrix} 7.358 & 21.55 & -5886 \\ -1.861 & 22.33 & -2400 \\ 0.0012 & 0.0207 & 1 \end{bmatrix}.$$

[2] *Hint*: You may find [A Complete Review of the OpenCV Object Tracking Algorithms]() useful.

[3] See Section 2.2 in Bernardin & Stiefelhagen, "Evaluating Multiple Object Tracking Performance" [[link]()].