1. Explain why the stitching from Set 2 fails:

Set 1:

Set 2:

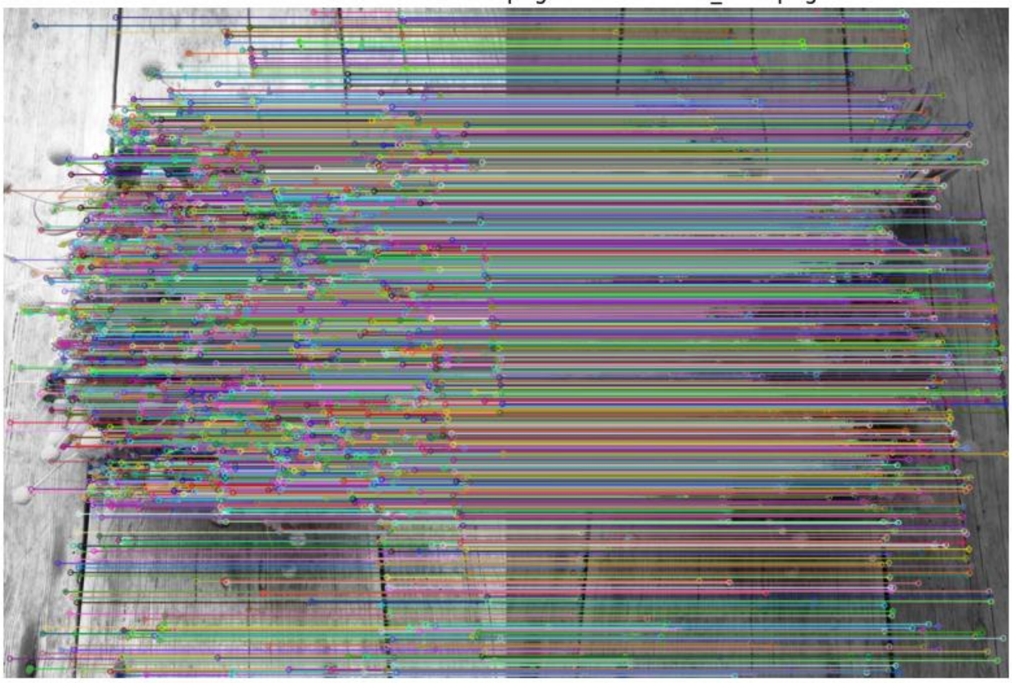
A, B, D and E are very close to each other on the Y axis. Therefore, noise greatly affects the homography computed suing these keypoints, and cause significant vertical distortions.

2a. 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.

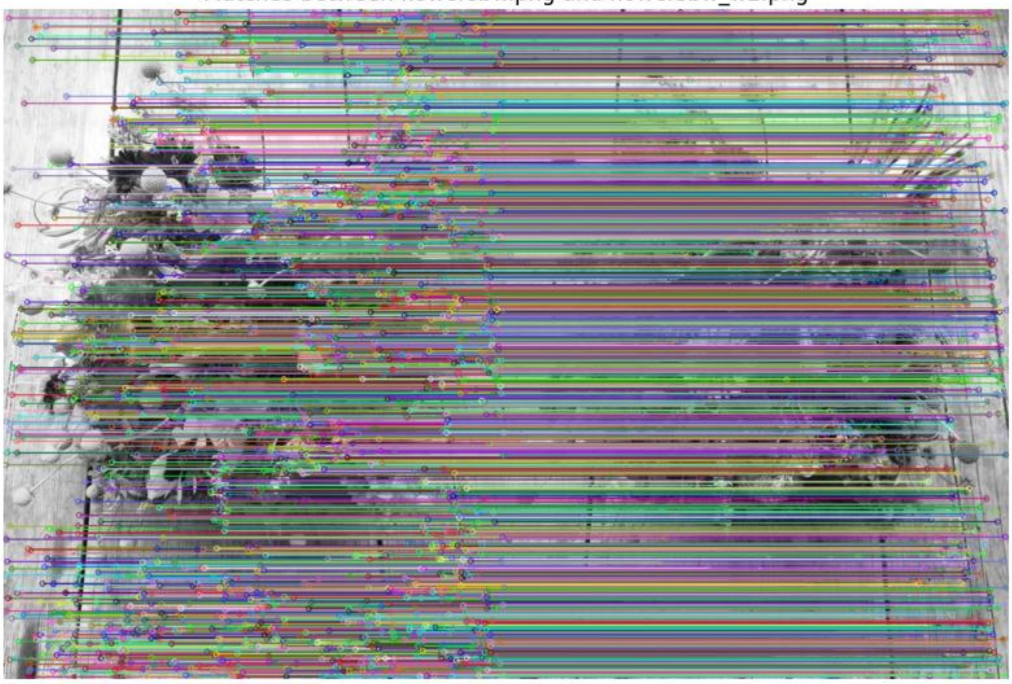
(c) has less keypoints detected than (b), while (d) has roughly the same number of keypoints detected as (b). The keypoints also appear in different locations. When the intensity is reduced like in (c), the keypoints change their location because the contrast in the image is reduced. However, this effect is small and most keypoints stay the same or shift slightly.

2b. Compare and comment on the number of matched keypoints and their locations for the two image pairs. Explain any differences that arise.

Matches between flowersbw.png and flowersbw_dark.png



Matches between flowersbw.png and flowersbw_w2.png

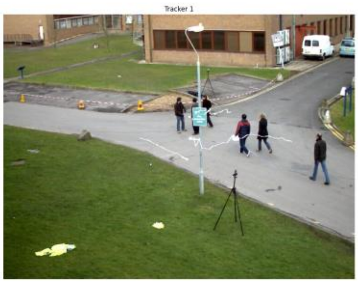


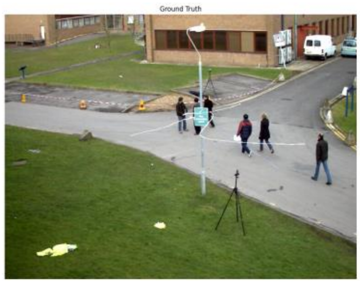
The matchings have roughly the same number of keypoint matches. However, the matchings between flowersbw.png and flowersbw_dark.png has less keypoints on the background of the image compared to the matchings between flowersbw.png and flowersbw_w2.png.

The SIFT feature detector looks for local features such as gradient intensity and direction in an image. flowersbw_dark.png has low brightness and contrast, so these features may be harder to detect in the background of the image, where there is lower variation in intensity.

4.

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.





The BOOSTING tracker is rather accurate, but much jerkier compared to the ground truth. The MIL tracker is completely off.

c. Compare the MOPT and MOTA of the two trackers based on the ground truth given in data/gt.csv and data/gt_format.txt.

```
BOOSTING tracker:
MOPT: 0.00
MOTA: -1.19
MIL tracker:
MOPT: 0.00
MOTA: -1.19
```

d. For each tracker, identify a failure case and analyse the potential cause(s) of the failure.

The trackers fail when the object is occluded by another object. This is because the tracker relies on the object being tracked to be visible in the image. When the object is occluded, the tracker cannot find the object in the image and cannot update the position of the object. If the occluding object is also moving, then the tracker often shifts to the tracking the occluding object instead.

5. Propose a method to prevent ID shift. State the assumptions, describe the algorithm steps, and discuss possible drawbacks.

In addition to motion-based methods, we can use appearance-based features to differentiate the objects. These features can include color histograms or texture features. The assumptions are that each person has a unique appearance that can be used to distinguish them from others, and that the appearance of the object do not change (by much). The algorithm steps are as follows:

1. Detect and track objects using a motion-based approach, such as a KCF tracker.
2. Extract features from the detected objects in the box.
3. When a tracked object moves from one frame to the next, for each box, calculate obtain a score based match it to the closest appearance model using a distance metric, such as Euclidean distance. If the distance between the tracked object and the appearance model is below a certain threshold, associate them with the same ID. Otherwise, assign a new ID.
4. Update the appearance model for each object based on the features measured in this frame.

Possible drawbacks of this approach include:

Computational complexity: The appearance-based methods can be computationally expensive, especially when dealing with large datasets.

Occlusions: Appearance-based methods may not work well when objects are occluded or partially hidden from view. When they come back into view, they will be assigned as a different object.

Changes in appearance: Appearance-based methods may not work well when there are significant changes in appearance, such as changes in pose, lighting, rotation, etc.

6. Estimate the number of steps taken by the individual marked with the red box in Figure 5a between frames 0 and 150.

Assume each step to be every time the box moves up and down alternatingly compared to the previous frame. The ID of the individual in the gt data is 19. The result is 42.

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.

The ID of the individual is 16. The physical distance walked by individual with id=9 between frames 117 to 221 is 116.99 m.