HW03 Report

1. Introduction

Homework 3 is to implement the automatic panoramic image. It involves detecting feature points, describing them, and then matching features between two images to align them correctly. This task uses SIFT (Scale-Invariant Feature Transform) for feature detection and description, followed by feature matching with ratio to filter out low-quality matches. After matching features, we use RANSAC to compute a homography matrix that minimizes errors caused by outliers. This matrix allows us to warp one image to align with the other, setting the foundation for creating an automatic panoramic image through image blending.

2. Implementation

a. Interest points detection & feature description by SIFT

We use the SIFT (Scale-Invariant Feature Transform) algorithm to detect key interest points within the images. SIFT is a robust feature detection method that is both scale and rotation-invariant, making it suitable for detecting consistent image features under varying scales and orientations. These keypoints capture distinctive local features in the image, while the descriptors record pixel information around each keypoint.

For this part, we adjusted the following parameters:

nfeatures: Specifies the maximum number of features to retain.

contrastThreshold: Filters out low-contrast keypoints.

edgeThreshold: Ignores points with high edge responses, reducing-edge effects.

sigma: Controls the standard deviation for Gaussian blur in constructing the scale space.

b. **Feature matching by SIFT features**

We perform feature matching to find similar features between the two images. To improve the accuracy of the matches, we apply a ratio test.

In the ratio test, we compare the distances between the nearest neighbor and the second nearest neighbor of a feature. The filtering formula is as follows:

$$Ratio = \frac{||f1-f2||}{||f1-f'2||}$$

The principle of the ratio test is that if the ratio $\frac{||f1-f2||}{||f1-f'2||}$ is smaller than a predefined threshold, we consider the match reliable and retain it; otherwise, we filter it out. This effectively reduces false matches and keeps only high-quality feature pairs.

c. **RANSAC to find homography matrix H**

The RANSAC function we wrote takes 5 parameters as input: **points1**, **points2**, **samples**, **iteration**, and **threshold**.

points1 and **points2** are lists of points of the correspondences found in feature matching, in the two images, respectively.

samples is the number of correspondences we sample in each iteration.

iteration is the number of iterations we run for.

threshold is the criteria we use to determine if the result of applying the homography matrix is satisfying enough.

For the process, in each iteration, we first randomly sample a number of correspondences from **points1** and **points2**, then use these points to calculate the homography matrix. Afterwards, we apply the homography matrix to all points in **points1** to get a new list of points called **result_points**. We then compare every point in **result_points** with points in **points2**, and calculate the number of distances between those two points that is smaller than **threshold**, as the score of this homography matrix. Finally, we return the homography that has the highest score.

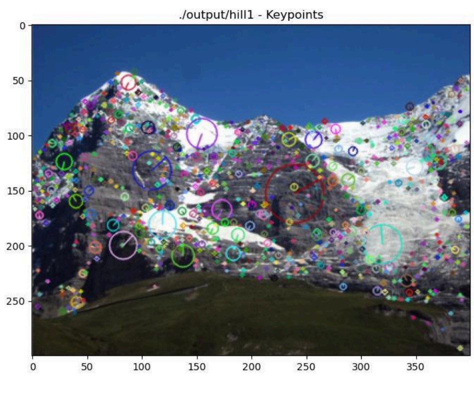
d. **Warp image to create panoramic image**

After obtaining the homography matrix, H, we can warp the pixels from image2's coordinate system to image1's coordinate system by the homography matrix. The warp function calculates the transformation for the image2 corners first so that we can know the size of the automatic panoramic image. We copy the image1 value to the corresponding pixels of the automatic panoramic image. Next, we apply the inverse of the homography matrix to map other pixels from the image1's coordinate system to image2's coordinates and use interpolation to obtain the corresponding color value. In the overlapping region, a blending ratio is applied to smoothly merge the two images to avoid the smear frames condition.

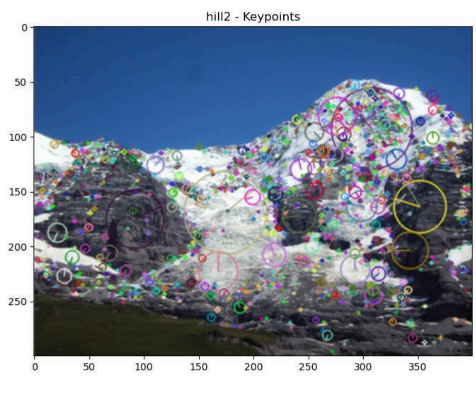
3. Experimental result

The process was tested on three pairs of provided images and one pair of our images with varying characteristics:

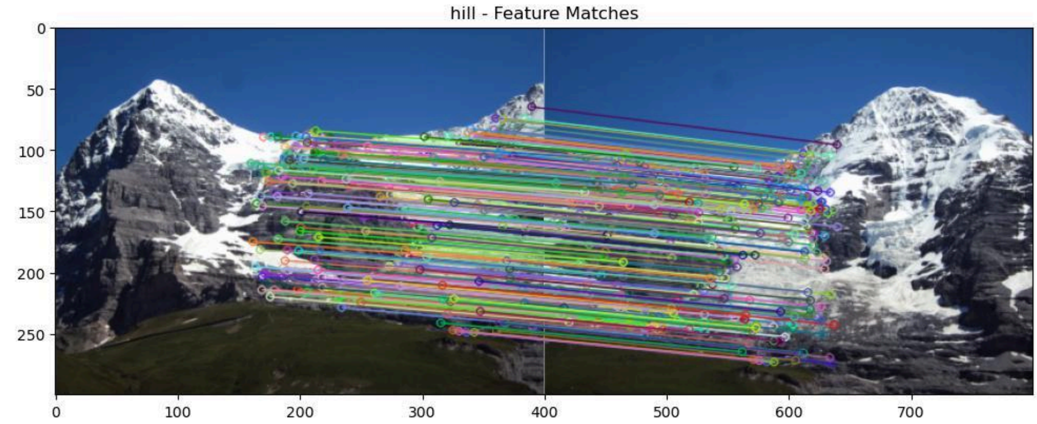
Mountain Images: The mountain images had distinct, non-repetitive patterns, which resulted in high-quality feature matches. The matching lines aligned with significant features like edges and crests.



hill1_keypoints.jpg



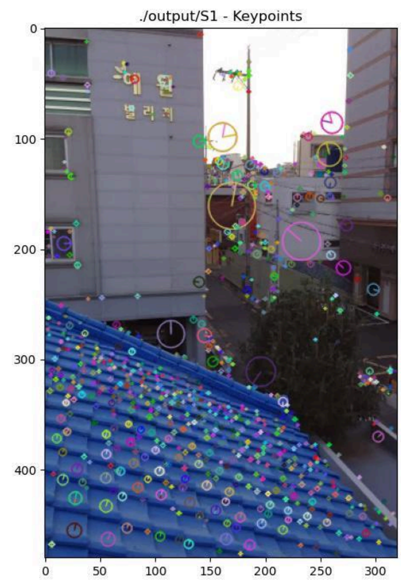
hill2_keypoints.jpg



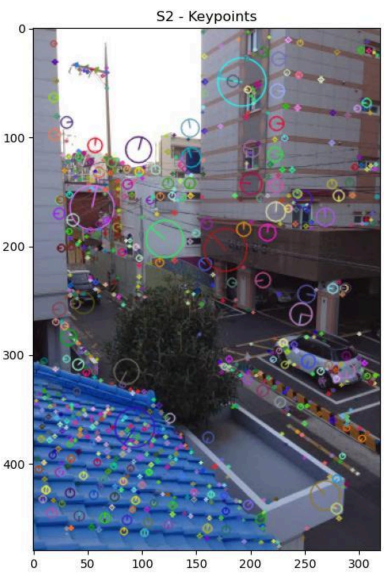
hill_feature_matching.jpg

hill_final_result.jpg

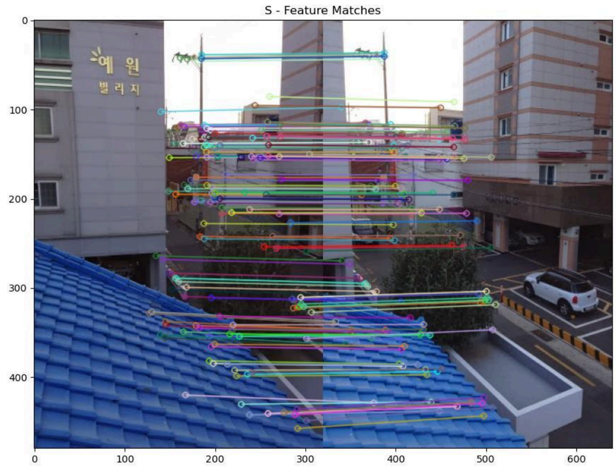
Building Images: The building images presented challenges due to the regular patterns in the tiles, which SIFT handled well with careful parameter tuning.



S1_keypoints.jpg

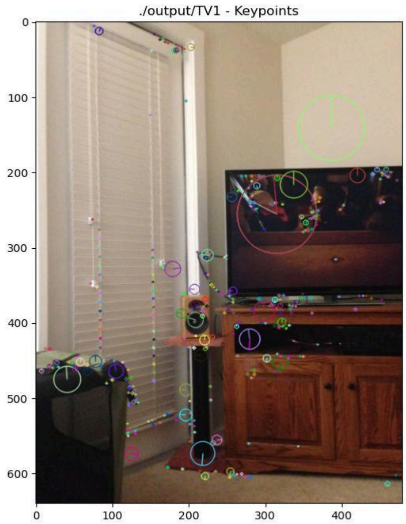


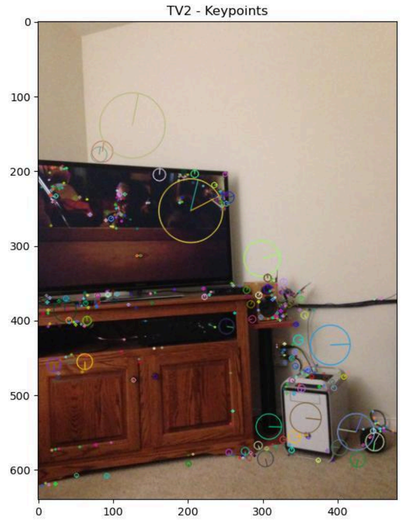
S2_keypoints.jpg



S_feature_matching.jpg

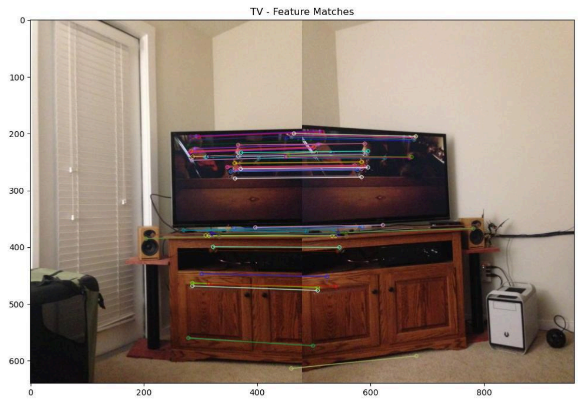
TV Images: This pair contained some repetitive textures, but the SIFT algorithm effectively distinguished the unique feature points around the edges of the TV and other distinct objects.



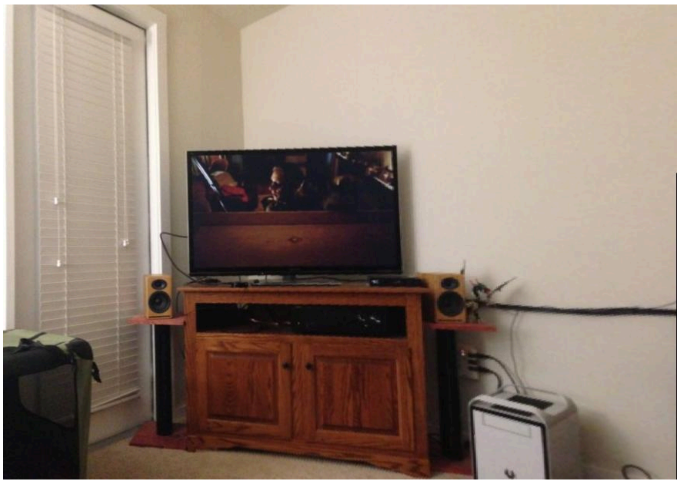


TV1_keypoints.jpg

TV2_keypoints.jpg



TV_feature_matching.jpg



TV_final_result.jpg

Our image

doll1_keypoints.jpg

doll2_keypoints.jpg

doll_feature_matching.jpg

doll_final_result.jpg

For each pair, the results included:

Step 1、2: Keypoint detection visualized as circles overlaid on the original images.

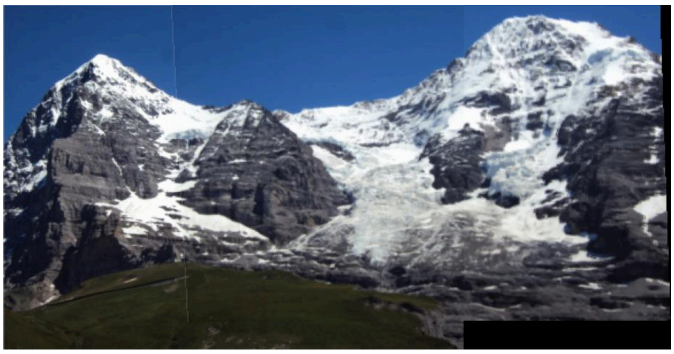
Step 3: Visualization of feature matching between the two images with lines connecting the matched points

and the panoramic image.

4. Discussion

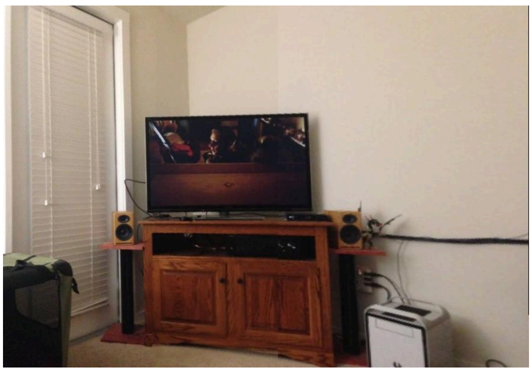
Compare the blending method

a. Fix the blending ratio



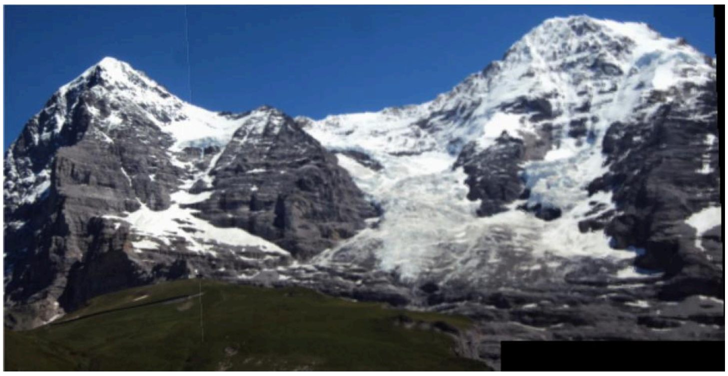
hill_fix_final_result.jpg

S_fix_final_result.jpg



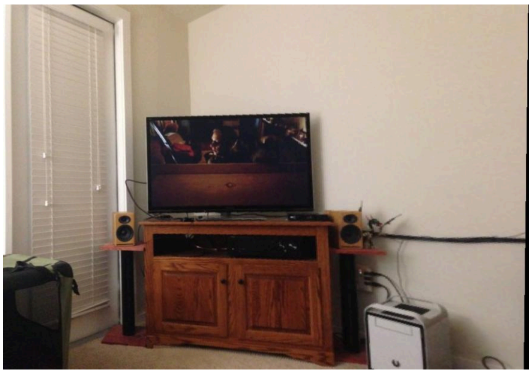
TV_fix_final_result.jpg

b. Linear blending ratio



hill_linear_final_result.jpg

S_linear_final_result.jpg

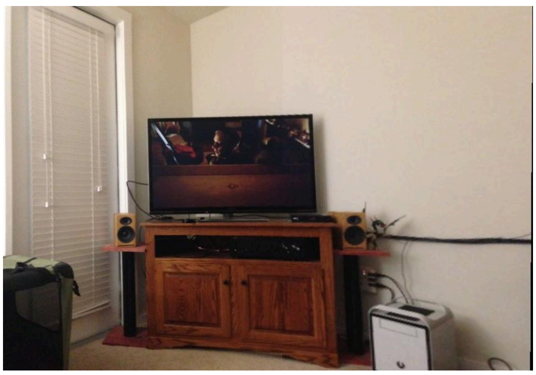


TV_linear_final_result.jpg

c. Sigmoid blending ratio

hill_sigmoid_final_result.jpg

S_sigmoid_final_result.jpg



TV_sigmoid_final_result.jpg

Compare different blending methods, we observe that the hill image has an obvious color variation boundary in the fixing and linear blending ratio. And the problem effectively improved by using the sigmoid blending ratio in the hill image. The S_linear_final_result.jpg and the TV image have the best performance by using the linear blending ratio.

Challenges Encountered:

- Overlapping and Ambiguous Matches: SIFT generated ambiguous matches in images with repetitive textures, like building tiles. Adjusting the ratio_test parameter to a lower value helped filter out many false positives.
- Parameter Tuning: The right values for contrastThreshold, edgeThreshold, and nfeatures required experimentation. Higher thresholds helped reduce noise, especially in images with busy backgrounds, but could also lead to missed keypoints in images with subtle details.
- Feature Matching Accuracy: Some matches were initially inaccurate, especially when using a higher ratio_test value.
- The black borders of the automatic panoramic image

Resolutions:

We adjusted edgeThreshold and contrastThreshold to reduce noise and focus on more prominent features, which improved the robustness of the feature detection process.

To reduce false matches in feature matching, we experimented with different values of ratio_test.

Removing the black borders of the automatic panoramic image is the problem that I have spent a lot of time trying to solve. However, I can't completely solve this problem. I could only reduce the black borders by limiting the image row.

5. Conclusion

Through parameter tuning and experimental analysis, we achieved a reliable alignment across a variety of images, overcoming challenges related to texture ambiguity and feature matching accuracy. The implementation demonstrates that SIFT, when combined with the proper matching techniques and parameter adjustments, can effectively stitch images with differing perspectives. The method effectively shows the way combined feature detection, matching, and alignment techniques produce the high-quality panoramic images.

6. Work assignment plan between team members.

彭宇楨:33% , 洪義翔:34% , 張宥楨: 33%