Assignment 4 - Report

EE5178: Modern Computer Vision

Vishnu Vinod

Contents

1 Introduction

Image stitching is the process of combining multiple images with overlapping fields of view to form a segmented panorama. In this particular assignment we work with image stitching using the RANSAC algorithm to estimating a parametric model from noisy observations of various features. To do this, we will compute SURF/SIFT/ORB features in both images and match them to obtain correspondences. We will then estimate a homography from these correspondences, and we'll use it to stitch the two images together in a common coordinate system. The two images we will stitch together are shown below:

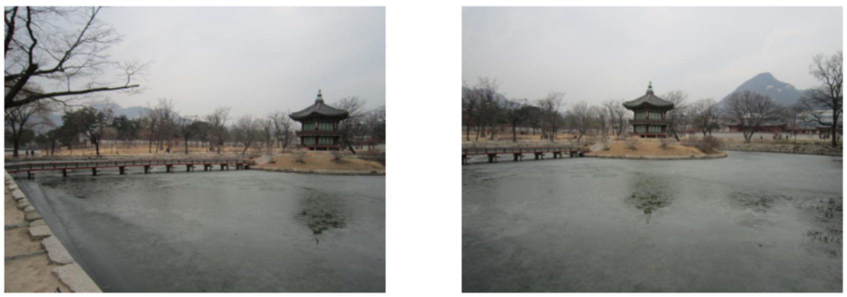

Figure 1: Images for Stitching Panorama

2 Matching Features

2.1 Finding SIFT Keypoints

As the first part of the assignment we obtain the **SIFT features** (keypoints and descriptors) for both the images. We use feature extractors which are usually invariant to scaling and rotation to obtain the features to be used for mapping. The features obtained thus are shown below:

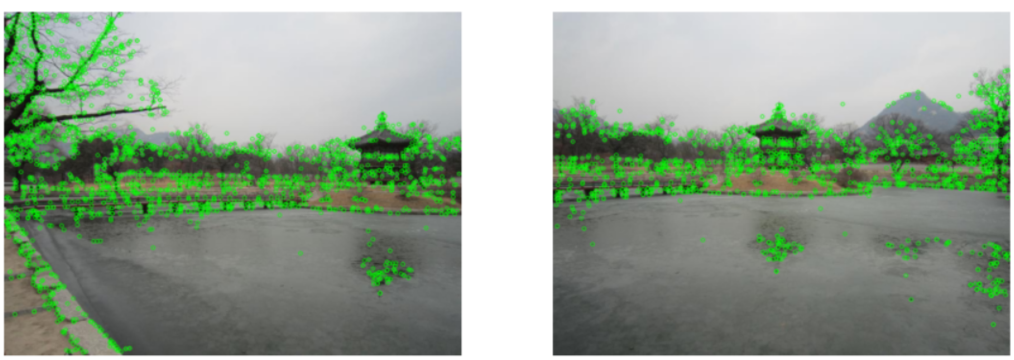

Figure 2: Keypoints for both Images

2.2 Feature Matching

We can obtain a large number of features using the feature extractor. Now we run a **brute-force matching** algorithm on these features with a KNN match which returns the k best matches for every feature. In our code we have set $k = 2$. Now we use these matches and find out which of them pass the **ratio test** proposed by D. Lowe in his SIFT paper. The matches obtained thus (sampled one of every four matches for representability purposes) are given below:

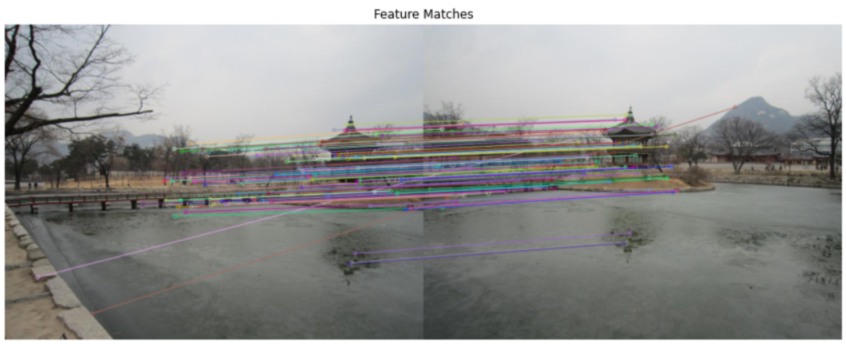

Figure 3: Feature Matches amongst the images

3 Estimating Homography using DLT

The homography matrix H for transforming geometrically between the first and the second images is calculated using the **DLT** (**Direct Linear Transform**) method. Here we use 4 point correspondences to set up a 8×9 matrix A of the form:

$$A_{2i} = [x_s, y_s, 1, 0, 0, 0, -x_t x_s, -x_t y_s, -x_t]$$

$$A_{2i+1} = [0, 0, 0, x_s, y_s, 1, -y_t x_s, -y_t y_s, -y_t]$$

Here the points (x_s, y_s) and (x_t, y_t) belong to the first and second image respectively. Now carry out the singular value decomposition of A to get:

$$A = U \cdot S \cdot V$$

Now the matrix V is a 9×9 matrix whose columns correspond to the singular values in the diagonal matrix S. Then the column corresponding to the smallest singular value is the solution to the homography vector h. Rearranging the 9 elements into a 3×3 matrix we get the solution to the homography matrix. The homography matrix for the two images given above is:

$$H = \begin{bmatrix} 1.850382 & -0.001449 & -533.994936 \\ 0.277501 & 1.627836 & -197.965394 \\ 0.001025 & 0.000172 & 1 \end{bmatrix}$$

4 Implementing RANSAC

RANSAC or (**RAndom SAmple Consensus**) is a method of iteratively estimating a mathematical model from a dataset that contains a large number of outliers. It works by identifying datapoints which are outliers and excluding them from the estimation of a mathematical model. It is a non-deterministic algorithm which shows progressively better results as the number of iterations increases. It uses **repeated random sub-sampling**. The algorithm is detailed below.

4.1 Algorithm

- Randomly select a subgroup of the input datapoints with cardinality being the minimum required to estimate the mathematical model.
  For estimating the homography transform the minimum required point correspondences are 4.
- Estimate the mathematical model using this random minimal subgroup.
  The Homography matrix for the set of 4 points is estimated using the DLT method discussed in the previous section
- All other data is tested on this fitted mathematical model. The measure of how well the model fits the data uses a model-specific loss function.
  The Euclidean distance between a transformed - datapoint and its untransformed match is used as a measure of loss.
- A consensus set of all points fitting the model is formed.
- The above steps are repeated with repeated random sub-sampling of the initial minimal input subgroup to find the best estimated model and the set of consensus points (or inliers).
- After a specific number of iterations (around 1000) the best model and its corresponding consensus set is taken.
- The consensus set is now used to re-estimate the model.

4.2 Applying RANSAC

Upon using the above RANSAC algorithm with the two images we can estimate the homography transform and warp the images accordingly as well.

Figure 4: Warped Image

The keypoints in the image and the feature matchings have also been shown here so as to give an idea of how the homography affects the matchings.

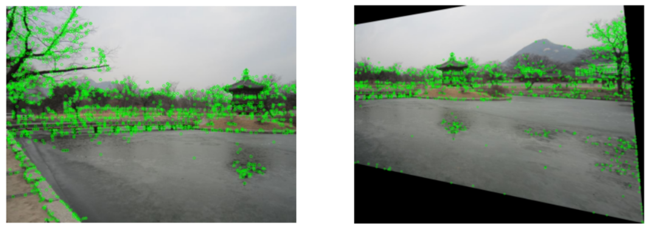

Figure 5: Keypoints in the warped images

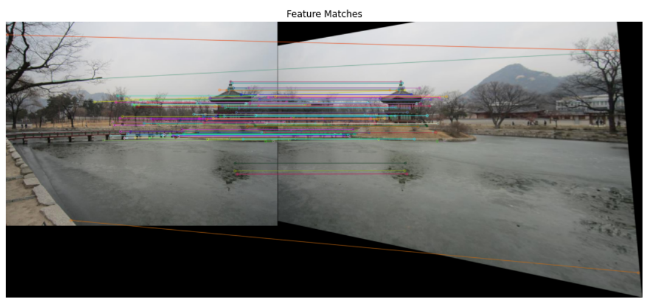

Figure 6: Feature matchings between the warped images

5 Panorama Stitching

The panorama stitching algorithm simply involves the two images being warped using a `warpPerspective()` function. The warped images are then overlayed over one another such that the features match up correctly. The overlapping areas have their pixels calculated by taking the average of the values of both the images. The final stitched image is given below:

Figure 7: Final Panorama

6 Three-Image Panorama

The last task in this assignment is to stitch together three images to form a panorama. This is done in a similar fashion. The images are given below:

Figure 8: Images to stitch

6.1 Finding the SIFT keypoints

Once again we calculate the **SIFT features** (keypoints and descriptors) for both the images. We use feature extractors which are usually invariant to scaling and rotation to obtain the features to be used for mapping. The features obtained thus are shown below:

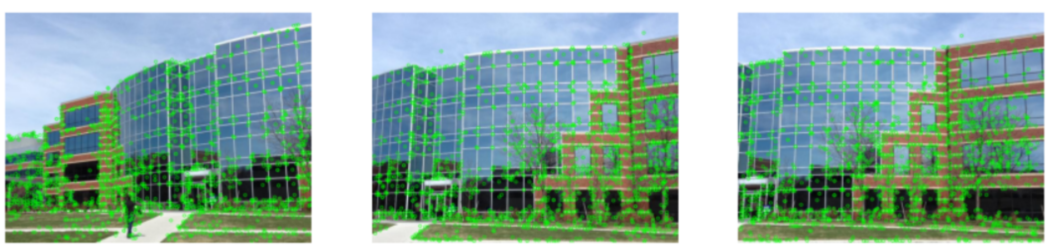

Figure 9: Image Keypoints

6.2 Finding the Feature Matches

Once again we run a **brute-force matching** algorithm on these features with a KNN match which returns the k best matches for every feature. In our code we have set $k = 2$. Now we use these matches and find out which of them pass the **ratio test** proposed by D. Lowe in his SIFT paper. The matches obtained thus (sampled one of every 2 matches for representability purposes) are given below for each pair of images:

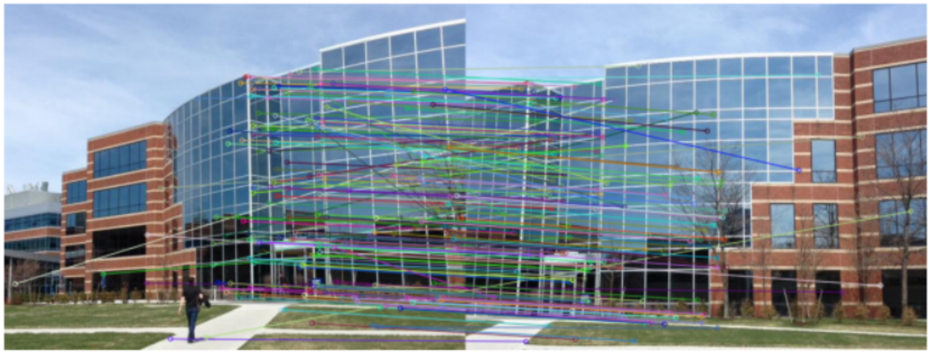

Figure 10: Feature Matches between Images 1 and 2

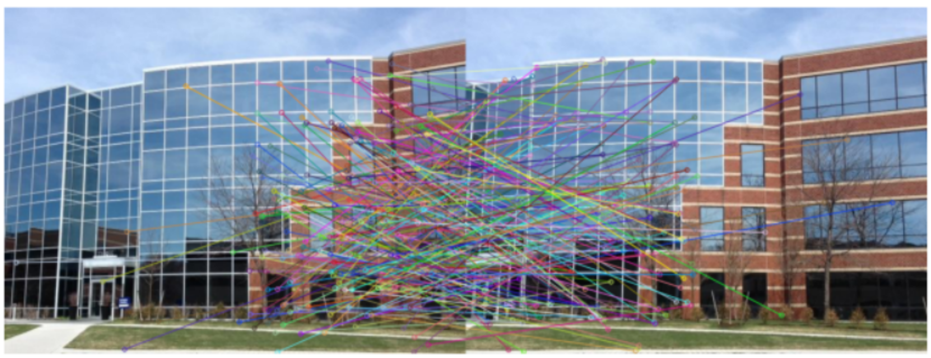

Figure 11: Feature Matches between Images 2 and 3

6.3 Calculating the Homography Matrices

The homography matrices for both these transformations are calculated using the previously discussed RANSAC method. The homography matrices are themselves calculated using the Direct Linear Transform (DLT) method which was also discussed earlier.

$$H_{12} = \begin{bmatrix} 1.262278 & -0.095288 & -315.375424 \\ 0.194759 & 1.170368 & -57.692845 \\ 0.000446 & -0.000048 & 1 \end{bmatrix}$$

$$H_{32} = \begin{bmatrix} 0.908935 & 0.013489 & 103.448963 \\ -0.052317 & 0.958936654 & 16.155917 \\ -0.000143 & -0.000004 & 1 \end{bmatrix}$$

Where the matrices H_{12} and H_{32} correspond to the homography transformations required to overlap images 1 and 3 onto the central image (ie. image 2).

6.4 Final Panorama

The three-image panorama can be obtained by repeatedly applying the two image panorama stitching operation on pairs of the images thereby stitching them together. The final stitched Panorama is given below:

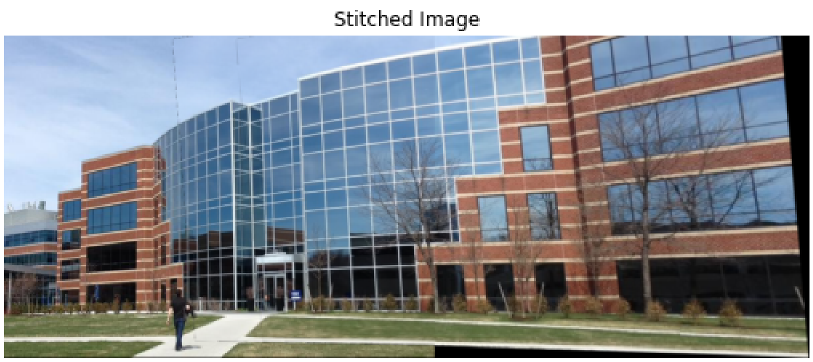

Figure 12: Final Panorama

THE END