Homework 2 - Computer Vision, 2018 Spring
Team35

Teammates:0416088 鄭甯遠,0650249 劉岑陽子,0660819 司菲 Siffi Singh

The task is to perform automatic panoramic image stitching, on the dataset collected by our team. The tasks can be further broken down into following subtasks:
1. Interest Point detection and Feature description by SIFT.
2. Feature matching by SIFT features
3. RANSAC to find homography matrix 'H'
4. Warp image to create panoramic image

In this report, we will be explaining each part in the following way:
1. Steps to run the file.
2. Explanation of Code Logic
3. Results
4. Inferences from Result

I. Interest Point detection and Feature description by SIFT

So, in this homework we had to take our own photos and stitch those pictures together. We need to see how to blend the overlapping area. For the first part, we need to find the interest point, the interest points are well-defined positions in space. The local image structure around the interest point is rich in terms of local information contents (e.g.: significant 2D texture), such that the use of interest points simplify further processing in the vision system. It is stable under local and global perturbations in the image domain as illumination/brightness variations, such that the interest points can be reliably computed with high degree of repeatability.

II. Feature matching by SIFT features

Features from Scale-Invariant Keypoints, which extract keypoints and compute its descriptors. We have used a popular local feature descriptor called SIFT to extract some interesting points from images and describe them in a standard way. Once we have these local features and their descriptions, we can match local features to each other and therefore create a stitched target image.

III. RANSAC to find homography matrix 'H'

We have two images x,x'. And each represents one 3D position(the relationship between 2D and 3D can be formulated b different in 2 imageselow). Because there is only rotation in our cases(for HW2), only R matrix is different in two images. And the size of R is 3x3.

x = K [R t] X

x' = K' [R' t'] X'

Here we simplify the problem to x'=Hx, and H is the Homography matrix which can project one image to the other image coordinate. H can be reshape to vector with dimension 1.

$$\mathbf{x}' = \mathbf{H}\mathbf{x} \qquad \mathbf{x}' = \begin{bmatrix} w'u' \\ w'v' \\ w' \end{bmatrix} \qquad \mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}$$

$$\begin{bmatrix} -u & -v & -1 & 0 & 0 & 0 & uu' & vu' & u' \\ 0 & 0 & 0 & -u & -v & -1 & uv' & vv' & v' \end{bmatrix} \mathbf{h} = \mathbf{0} \qquad \mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix}$$

We can find homography matrix H using DLT(Direct Linear Transform). DLT algorithm is shown below.

1. Apply SVD decomposition to find V matrix.
2. Assign the vector with the smallest singular value to h.
3. Get H from reshaping h to a 3x3 matrix.

$$\begin{bmatrix} -u_1 & -v_1 & -1 & 0 & 0 & 0 & u_1u_1' & v_1u_1' & u_1' \\ 0 & 0 & 0 & -u_1 & -v_1 & -1 & u_1v_1' & v_1v_1' & v_1' \\ & & & & \vdots & & & & \\ 0 & 0 & 0 & -u_n & -v_n & -1 & u_nv_n' & v_nv_n' & v_n' \end{bmatrix} \mathbf{h} = \mathbf{0} \Rightarrow \mathbf{A}\mathbf{h} = \mathbf{0}$$

- Apply SVD: $\boldsymbol{UDV}^T = \boldsymbol{A}$
- $\boldsymbol{h} = \boldsymbol{V}_{\text{smallest}}$ (column of $\boldsymbol{V}$ corr. to smallest singular value)

How we choose the best H matrix is decided by **RANSAC algorithm.**

1. At first, we choose 4 random potential matches(the matches key points).
2. And then compute H using normalized DLT.
3. Project points from x to x’ for each matching pair.
4. Count points with projected distance smaller than a threshold.
5. Repeat steps 1-4 for N times.

IV. Warp image to create panoramic image

Given the homography on the third part, it is easy to stitch the image together. One simply use the matrix to do a perspective transform and stitch them according to the coordinate. However, without adjusting the color, when one inspect carefully, their would be a cut line due to minor RGB difference in the same area. One can improve it by analyzing the color distribution and normalize it.

1. Steps to run file

- There is one file in PYTHON. The driver function is included in the same file.
- The file has several functions, separate functions for reading the file, finding key points, storing the interest and matched points between two images.
- To run the file, simply put the images in the same folder as the program files or specify the path of the two images.
- Run the, in the output you will get the stitched panoramic image.
- The final output is the stitched panoramic image of input Images taken by us.

2. Explanation of Code Logic

For finding the interest points, and matching features, we have several sub tasks. We have made use of OpenCV in python for leveraging the built-in functions that have made us to do our task easier. Following the some of the tasks:

- Reading the image files from the Data folder.
- Perform Gaussian smoothing of image, done to limit the interest points and filtering out the relevant interest points(kind of like thresholding).
- Converting the image to grayscale.
- Then we have used sift.detectandcompute function from openCV to perform feature extraction.
- After feature extraction, we need to store these key points for drawing the homography matrix.
- We have made use of cv2.BFMatcher function and passed the image descriptors as parameters to find the matching points in the two images.

The key points in using RANSAC to find homography matrix are:

RANSAC algorithm:

A learning technique to estimate parameters(here is the H matrix) of a model from some data(matched key points). Details will be shown below.

1. Select 4 random points to calculate H matrix.(the so-called inliners)
2. A H matrix is found to fit the four random points.
3. Check other key points if they can fit the estimated model well. A loss function *geometricDistance* is used to evaluate the distance between estimated point and original points(p2 and estimated p2).
4. If the distance is smaller than a threshold, we add it to the inliners set.
5. After we go through the whole matching key point set, we can check how many inliners we have. If the number is larger than the current maximum inliners, then replaced the H matrix with the new H. Otherwise, we just keep the best model and then do another iteration.
6. Repeat step 1~5 for 1000 iterations.
7. H matrix is found.

The key points in stitching images:

- Take image1, image2 and the homography matrix
- Map coordinate of image2 to image1 by perspective transform
- Stitch the transformed image and image1 together

3. Results

- Following is the output from the data that we collected, and used our algorithm to perform image stitching to create panoramic image.

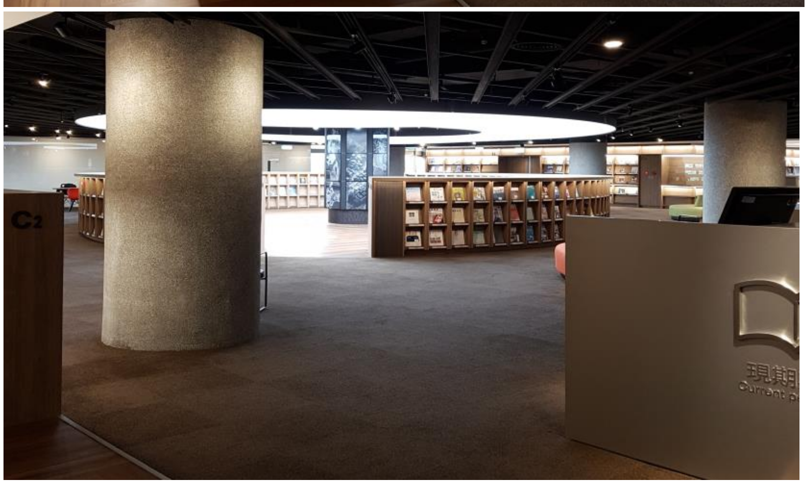
1. **Panoramic images - Input 1,2,3**

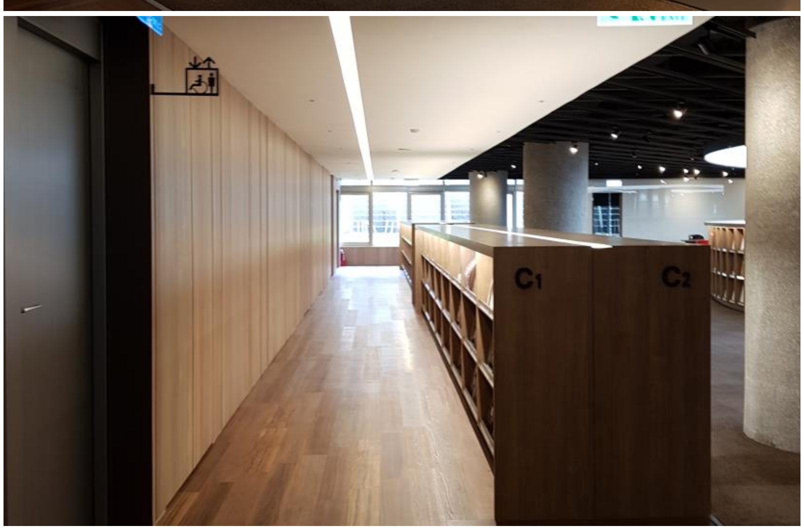
Output: Stitched Image from Image 1, 2, 3.

Matched Keypoints images(image 1and 2)

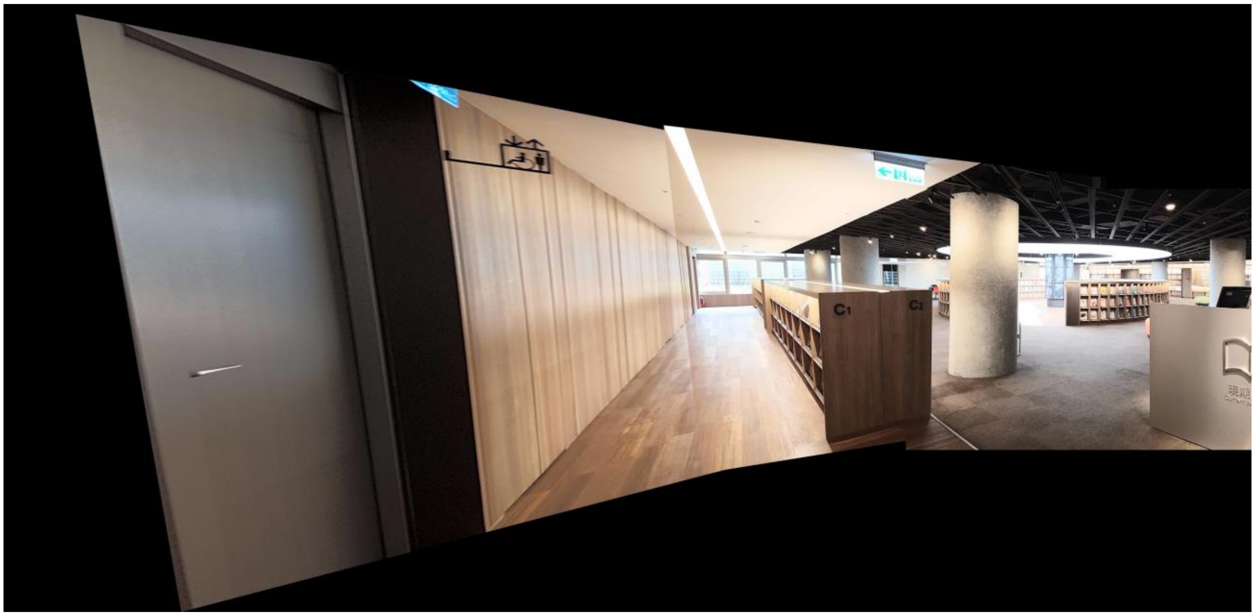
Matched Keypoints images(stitched_image1&2 and image3)

2. Panoramic images - Input 4, 5, 6





Output: Stitched Image from Image 4, 5, 6.



Matched Keypoints images(image 4 and 5)

Matched Keypoints images(stitched_image4&5 and image6)

4. Inferences from Result

- Not all interest points are relevant for image stitching, we must find the most important features, match them, create homography matrix and warp them to stitch them together to create panoramic image, which is why we have done image smoothening, that helps us to filter out important features to be matched.