Computer Vision Exercise 1

Summary

October 5th, 2017

Submited by:
Cheng-Kai Wang, M.Sc

1 Objective

In this exercise, you will first implement your own feature detector/descriptor and find best matching features between two images. In the second part, you will compare results against a reference implementation of the SIFT descriptor. Along with the template source code, a set of images is provided. However, you are encouraged to capture and use your own images.

2 Harris Corner Feature Matching

2.1 Input datasets

In this exercise, the image set presented below are used to test the results of the implemented feature matching algorithm.

Figure 1: Input image set

2.2 Feature Extraction

To obtain the Harris response for each pixel, the first step is to calculate the image gradients Ix and Iy in x and y-direction. This can be easily achieved by the Matlab built-in function $[Ix, Iy] = imgradientxy(inputImg)$. The input image here is first blurred with the Gaussian kernel in order to get rid of the noise. According to the following equation 1, we can obtain the Harris matrix. Be noted that in this implementation, the sum area is set to be a 3x3 window. However, alternative variations can be applied for different conditions. Follow equation 2, the Harris response for each pixel is given.

$$H = \sum_{3x3window} \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix} \tag{1}$$

$$K = det(H)/tract(H) \tag{2}$$

To find out representative key points. The Harris response should be larger than a certain threshold. Since the range of Harris response is different for different images. The implementation here allows the users to define the threshold of each image as a minimum value of the top few percent Harris response value. Left images of Figure 2 and Figure 3 present the extracted Harris corners which have Harris responses larger than the defined threshold (threshold = min(top 1 % of the Harris response in this case))

Non-maximum-suppression is then used to further filter the selected Harris corners. In this process, if pixels are not part of the local maxima, they will be discarded from the selected Harris corners. Right images of Figure 2 and Figure 3 show the results after non-maximum-suppression.

Figure 2: Extracted Harris corners before non-maximum suppression (left); after non-maximum suppression (right)

Figure 3: A closer look at two results. Before/After non-maximum suppression (left/right)

2.3 Feature Descriptor

The ideal matches of two interest points from both images will be the lowest sum of squared differences value (SSD). To do this calculation, we first need to extract feature descriptor, and it is defined as a image patch of 9x9 pixels centered around the point of interest. To simply further processing, each image patch is reshaped and stored as a 81x1 array

2.4 SSD Feature Matching

After the above processes, the number of extracted interest points will be different for both images. To start SSD feature matching, the image with fewer feature descriptors is defined as seed descriptors and the one with more descriptors is pool descriptors. Later on, the calculation of pair SSD

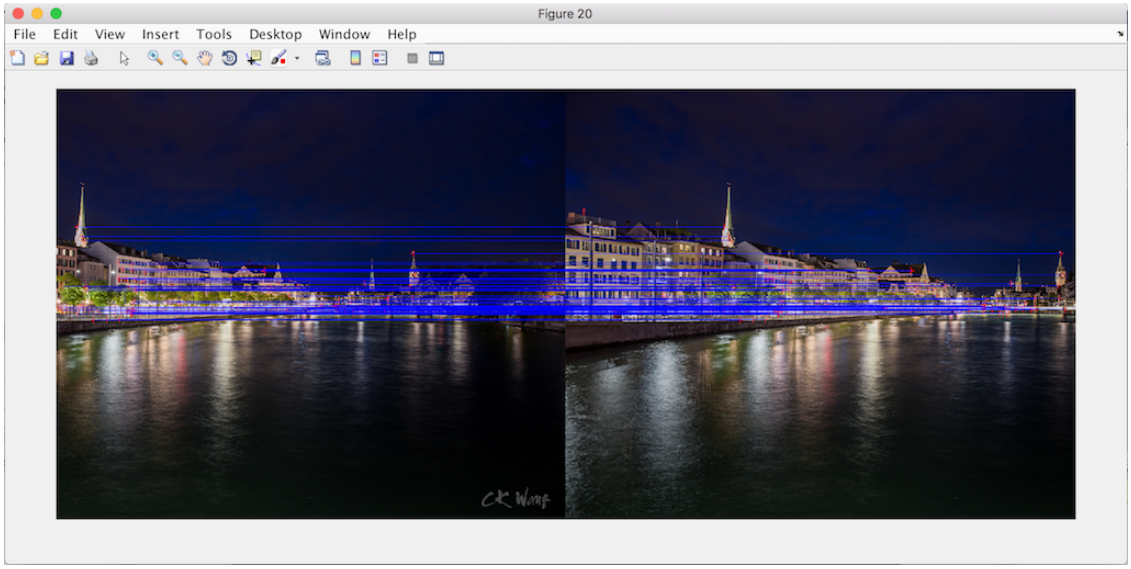

Figure 4: Matching result of the image set 01

value starts with each seed descriptor and iterate through pool descriptor. If the lowest pair SSD value in the loop is lower than the defined threshold, it is defined as a matched pair. The threshold value in this case is set to be 0.01 as an empirical result. The larger value indicates there will be more matches and it also means more errors. This threshold highly depends on the similarity of two images, to avoid manual testing, defining fixed number of matches for each image set is also a recommended process. Figure 4 shows a very good matching result.

3 SIFT Features

Going through the tutorial and executing SIFT algorithm. One can easily observe that the features extracted via SIFT is very different from the Harris corners method, which only extracts sharp edge as a descriptor, see Figure 5.

Figure 5: Applying SIFT and Harris corners algorithm to extract features on the same image.