**ENGN 2605 Image Understanding**

**Lab 05**

**Instructor: Prof. Benjamin Kimia**

**Name: Kuan-Min Lee**

**Student ID: 140481859**

**Problem 1 Feature Descriptors and Matching**

In this section, there are three different ways of features matching that have been tested: Sum of Squared Distance, Normalizaed Cross Correlation, and Chi-Square. For each of the method mentioned, four test images are given for testing each algorithm, and the best results among the algorithms will be shown.

The first image is the drawing wall, which represents the scenario of image being rotated. The best algorithm performed in this case is the chi-square algorithm with window size of 51\*51, and the accuracy is 11.38% for feature repeating:

A picture containing text, graffiti, decorated

Description automatically generated

Figure 1: Results of Chi-Square Algorithm for Image Pair 1

For the second image pair, the test image is a scenario of a ship featuring with different scale of taking photos. The best algorithm performed in this case is the normalized cross correlation with a window size of 19\*19, and the rate of repeating features is 1.9%:

A picture containing sky, outdoor, cloudy, day

Description automatically generated

Figure 2: Results of Normalized Cross Correlation Algorithm for Image Pair 2

For the third image pair, the scenario is regarding a view of a metropolis area featuring with different rotation of camera. In this scenario, the algorithm performs the best is the sum of squared distance with window size of 15\*15, and its rate of repeating features is 79.88%:

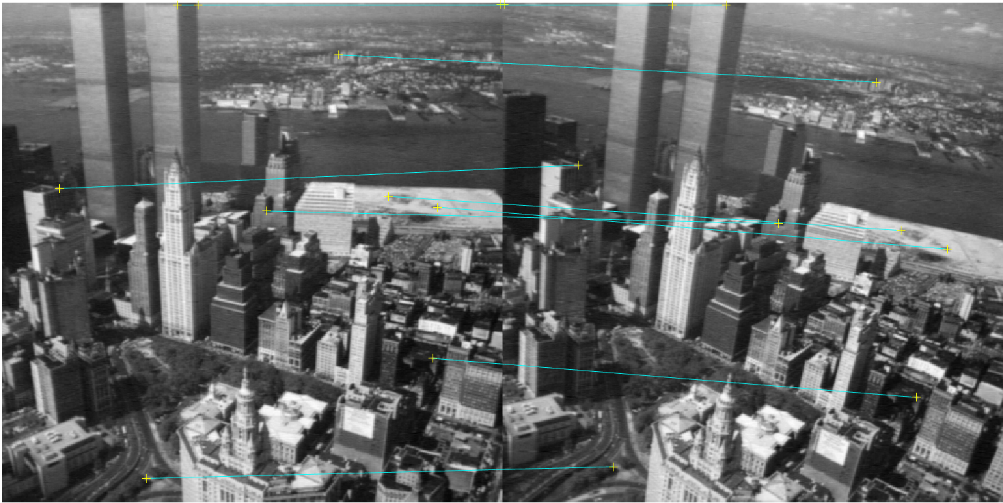


Figure 3: Results of Sum of Square Distance Algorithm for Image Pair 3

For the last test image, the scene is regarding as a parking lot view featuring with different illuminances. In this case, the algorithm performs the best among all is the normalized cross correlation algorithm with window size of 13\*13, and its rate of repeating features is 67.4%:



Figure 4: Results of Normalized Cross Correlation Algorithm for Image Pair 4

The later part of this section is utilizing SIFT features for the purpose of features matching. For this part, the peak threshold is set as 0.1 and the edge threshold is set as 5. The following are the results of the tested image pairs:



Figure 5: Results of SIFT for Image Pair 1

(Accuracy rate: 12.9%)

A picture containing sky, outdoor, grass, cloudy

Description automatically generated

Figure 6: Results of SIFT for Image Pair 2

(Accuracy rate: 37%)

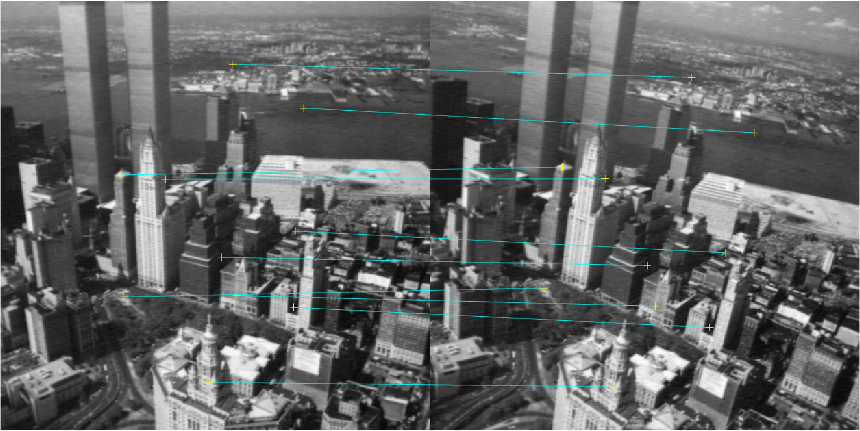


Figure 7: Results of SIFT for Image Pair 3

(Accuracy rate: 80.3%)

A screenshot of a video game

Description automatically generated with medium confidence

Figure 8: Results of SIFT for Image Pair 4

(Accuracy rate: 87.1%)

From the results shown above, it can be observed that the descriptors tend to fail more easily when the scale of the view is changed or when the view is rotated or distorted. SIFT and other algorithms all fails in the first and second image. In the first image, although the view is still typically the same but due to the rotation and distortion, the accuracy rate gets deteriorated, which none of them exceeds the rate of 50%. As for the second image, the view changes more drastically compared to the first image. Not only did it change the overall intensity, the change in the scale of the view also change the structure of the image. For example, neighborhood, the contour, etc. Therefore, for this scene, the descriptor also fails.

**Problem 2 Improve Feature Matching Accuracy**

In the first section of this experiment, the scene of mapping several features to one is observed. To eliminate this scenario, the bidirectional consistency and correspondence unique is considering ensuring that there will only be a 1 to 1 scenario in the feature matching. By implementing this, the accuracy rate is also expected to raise.

The following are the results of the same setting as the previous section for each testing image pair:



Figure 9: Results of SIFT for Image Pair 1

(Accuracy rate: 17%)

A picture containing sky, outdoor, grass, cloudy

Description automatically generated

Figure 10: Results of SIFT for Image Pair 2

(Accuracy rate: 5%)

A picture containing indoor

Description automatically generated

Figure 11: Results of SIFT for Image Pair 3

(Accuracy rate: 89%)

A picture containing text, building, road, outdoor

Description automatically generated

Figure 12: Results of SIFT for Image Pair 4

(Accuracy rate: 93%)

**Problem 3 Efficient Feature Matching**

To give a brief summary of BRIEF feature, the research implemented a creative feature of binary representation of the image patches. For each image patches, several numbers of sample pairs are extracted from the patch and for each of the sample pair, their magnitude are compared and the outcome of either is a 1 or 0 is given based on the testing results. For the features are a series of 1 and 0 numbers, which makes the entire features a binary number. That one thing that can save up the memory cause and also speed up the computation time due to its simplicity. Also, the hamming distance used in the research is also helping to reduce the computational time.