UNDERSTANDING OF SCALE INVARIANT FEATURE TRANSFORM METHOD (SIFT) ASSIGNMENT 1

Nivin Pradeep Kumar
S4035593
University of Groningen

Cendur Vishal Karthik Khanna
S3878732
University of Groningen

1. INTRODUCTION

Over the years object recognition has been gaining popularity. Object recognition in cluttered real-world scenes requires local image features that are unaffected by nearby clutter or partial occlusion. One of the algorithm used for object recognition is Scale-Invariant Feature Transform. The scale-invariant feature transform (SIFT) is a feature detection algorithm in computer vision to detect and describe local features in images. SIFT transforms an image into large collection of local feature vectors. In this report we take a look into SIFT method and it's properties. [1]

2. EXERCISE 1

In this exercise we compute the number of keypoints in *scene.pgm* image using the *sift()* command in MATLAB it is then used to compare it with the paper but there is no mention of the number of keypoints in the paper. The features include the key point features and the location and orientation of the key points. **1021 keypoints were found in scene.pgm**.
Next the dimensions of the array keys were computed using the *whos* command. The result of which is shown in Figure 1 which is obtained through matlab code. There is few difference we found out like how the number of bins is different, in the paper it was mentioned of 160 bins whereas when we implemented it in Matlab we got 128 bins. Moreover in the paper Hough transform was used to identify the rectangular feature and the keypoints where contained in it.

Then we inspect the key points using the *showkeys* command. The resultant image is shown in the Figure 2.

Name	Size	Bytes	Class
I	384x512	196608	uint8
keys	1021x128	1045504	double
loc	1021x4	32672	double

Fig. 1. Dimension of array keys

Finally we match the book.pgm image to scene.pgm using the command *match()*.The output for the matching is shown in Figure 3. **Here we found 882 keypoints in the new image and 98 matches were found.**. From this we where able to figure out that only 9.59% matching occurs with the scene. Now we match the image basmati.pgm to scene.pgm and try to find the keypoint matches. **Here we found 579 keypoints in the new image and 34 matches were found.** From this we could figure out that not all the matches where reasonable. There where a few keypoints which where getting matched between the chair and the edge of the book.

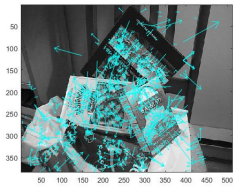

Fig. 2. Keypoints in scene.pgm

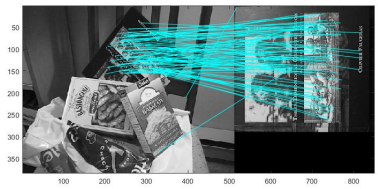

Fig. 3. Matching book.pgm to scene.pgm

2.1. Matching each of the detail images to street.png and then with streetlarge.png

We also used SIFT algorithm to compare the robustness of the algorithm with images of different resolutions.

2.1.1. Matching to lower resolution image - street.png

A lower resolution image *street.png* was used to compare with different detail images to get the number of keypoints it matches correctly. The result of the exercise is tabulated in Table 1. The Figure 4 shows the resultant image after matching detail1.png to street.png.

We have found 1452 keypoints in the *street.png* image and 1592 keypoints in the *detail1.png*. A total of 25 matches where found between detail1 and street.png.

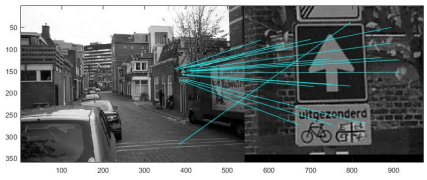

Fig. 4. Matching detail1.png to street.png

For the next image we found 46 keypoints in the *detail2* image and 6 matches were found between the *detail2* and *street.png*.

For the next image comparison we obtained over 588 key-

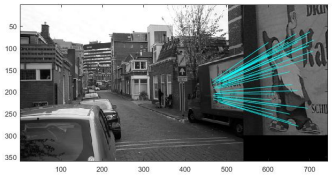

Fig. 5. Matching detail3.png to street.png

points in the *detail3* image and 37 matches were found between them. The Figure 5 shows the resultant image after matching *detail3.png* to *street.png*.

For the next image we where able to find 368 keypoints in the detail4 image and 20 matches were found between the *detail4* and *street.png*.

In the last low resolution comparison we obtained 126 key-

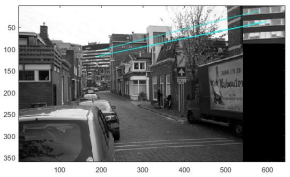

Fig. 6. Matching detail5.png to street.png

points in the detail5 image and 6 matches were found between the *detail5* and *street.png*. The Figure 6 shows the resultant image after matching *detail4.png* to *street.png*. We have summarized all the result and shown in the table 1.

2.1.2. Matching to higher resolution image - streetlarge.png

A higher resolution image streetlarge.png was used to compare with different detail images to get the number of keypoints it matches correctly. The result of the exercise is tabulated in Table 2.

In the first image comparison between the high resolution im-

Image	Keypoints	Matches	Percentage
Detail 1	1592	25	1.57
Detail 2	46	6	13.04
Detail 3	588	37	6.29
Detail 4	368	20	5.43
Detail 5	126	6	4.76

Table 1. Summary of Matching of lower resolution image

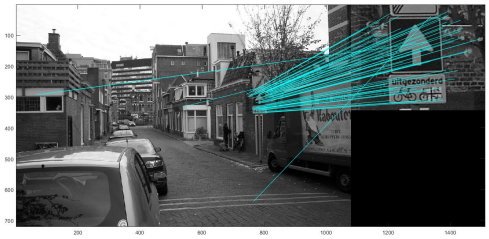

Fig. 7. Matching detail1.png to streetlarge.png

age and dettail1 we get 5316 keypoints in the *streetlarge.png* image and 1592 keypoints in the *detail1.png*.There were 82 matches found between detail1 and *streetlarge.png*. The Figure 7 shows the resultant image after matching *detail1.png* to *streetlarge.png*.

For the next image matching we obtained 46 keypoints in the detail2 image and 17 matches were found between the *detail2* and *streetlarge.png*.

For the next image matching we get 588 keypoints in the de-

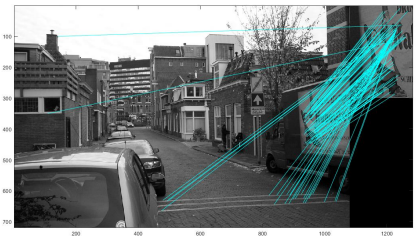

Fig. 8. Matching detail3.png to streetlarge.png

tail3 image and 110 matches were found between the detail3 and *streetlarge.png*. The Figure 8 shows the resultant image after matching *detail3.png* to *streetlarge.png*.

For the next image matching we where able to find 368 keypoints in the detail4 image and 73 matches were found between the detail4 and *streetlarge.png*.

In the last image matching between with higher resolution image we get 126 keypoints in the detail5 image and 17 matches were found between the detail5 and *streetlarge.png*. The Figure 6 shows the resultant image after matching *detail4.png* to *streetlarge.png*.We have summarized all the result

Fig. 9. Matching detail5.png to streetlarge.png

Image	Keypoints	Matches	Percentage
Detail 1	1592	82	5.15
Detail 2	46	17	36.95
Detail 3	588	110	18.70
Detail 4	368	73	19.83
Detail 5	126	17	13.49

Table 2. Summary of Matching of higher resolution image

and shown in the Table 2.

The results clearly shows that having a higher resolution will lead to better keypoints matching. This is justified by checking the images and the tables. When Gaussian filter is used in order for SIFT on the higher resolution image, it will perform well and have greater contrast so it provides better matching.

3. EXERCISE 2

In this exercise we match the detail images of the outdoor scene to the indoor scene and vice-versa.

3.1. Matching detail to scene

In this exercise we use *sift()* to compute the SIFT features in *scene.pgm* and then match the images *detail1.png* to *detail5.png* to the indoor scene image. We have found 1021 keypoints in the *scene.pgm* image and 1592 keypoints in the *detail1.png*. Only 1 match was found between *detail1.png* and *scene.pgm*. The Figure 10 shows the resultant image after matching *detail1.png* to *scene.pgm*.

For the next image matching we obtained 46 keypoints in the *detail2.png* and found 0 matches between *detail2,png* and *scene.pgm*. The Figure 11 shows the resultant image after matching *detail3.png* to *scene.pgm*.

For the next image matching we get 588 keypoints in the textitdetail3.png and found 0 matches between *detail3.png* and *scene.pgm*.

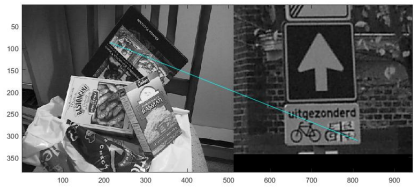

Fig. 10. Matching detail1.png to scene.pgm

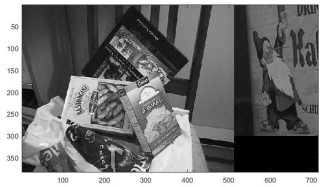

Fig. 11. Matching detail3.png to scene.pgm

For the next image matching we get 368 keypoints in the *detail4.png* and found 2 matches between *detail4.png* and *scene.pgm*. The Figure 12 shows the resultant image after matching *detail5.png* to *scene.pgm*.

For the final image matching we obtained 126 keypoints in

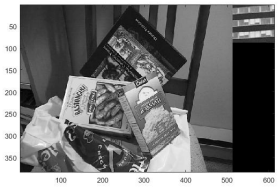

Fig. 12. Matching detail5.png to scene.pgm

the detail5.png and found 0 matches between *detail5.png* and *scene.pgm*.

3.2. Matching scene to detail

We use the SIFT algorithm to compute the SIFT features in *detail1.png, detail2.png, detail3.png, detail4.png, and detail5.png* and then match *scene.pgm* to each of the detail images.

In the first image matching 1592 keypoints in the *detail1.png* and 1021 keypoints in the *scene.pgm* image were found. There were 5 matches found between *scene.pgm* and *detail1.png*. The Figure 13 shows the resultant image after matching *scene.pgm* to *detail1.png*.

In the second image matching we obtained 46 keypoints in the *detail2.png* and 0 matches were found between *scene.pgm* and *detail2.png*.

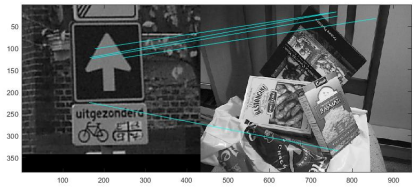

Fig. 13. Matching scene.pgm to detail1.png

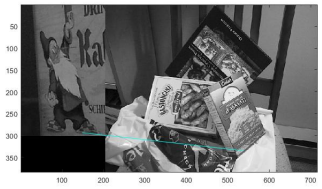

Fig. 14. Matching detail3.png to scene.pgm

In the third image matching we obtained 588 keypoints in the *detail3.png* and 1 match was found between *scene.pgm* and *detail3.png*. The Figure 14 shows the resultant image after matching *detail3.png to scene.pgm*.
In the fourth image matching we obtain 368 keypoints in the *detail4.png* and 3 matches were found between *scene.pgm* and *detail4.png*.

In the last image matching we obtain 126 keypoints in the

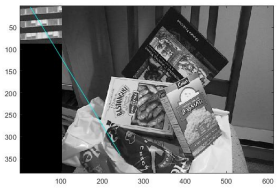

Fig. 15. Matching detail5.png to scene.pgm

detail5.png and 1 match was found between *scene.pgm* and *detail5.png*. The Figure 15 shows the resultant image after matching *detail5.png* to *scene.pgm*.

4. EXERCISE 3

In this exercise we get insight on how SIFT deals with rotation of an image while matching. In order to do that we perform the experiment with an image *tscene.pgm* which is rotated by an angle from 10 to 90,180 and 270 degree. From the Figure 16 we can infer that SIFT is invariant with rotation.

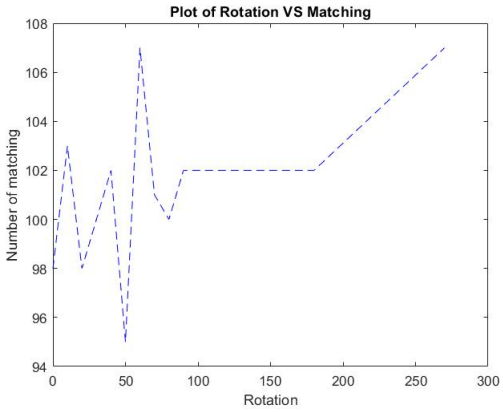

Fig. 16. Rotation VS Matching

5. EXERCISE 4

In this exercise we studied how the number of matches varies when the image is sheared. We perform the experiment by calciulating the matches when the image is sheared by 25, 50, 100, 200, 300, and 400 pixel. We observed that SIFT is not shear invariant. The Figure 17 shows the relation of matching with shearing. The matching decreases when we increase shearing pixel from an amount of 100.

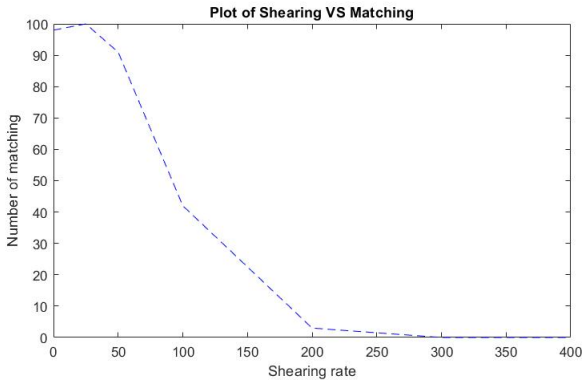

Fig. 17. Shearing VS Matching

6. REFERENCES

[1] D. G. Lowe, *Object Recognition from Local Scale-Invariant Features*. Proceedings of the Seventh IEEE International Conference on Computer Vision.