

# Computervision Lab 1

Jan Kramer  
Klaas Kliffen

February 19, 2016

## Exercise 1



Figure 1: The features SIFT finds in “scene.pgm”

The features SIFT finds in “scene.pgm” are shown in Figure 1. Since SIFT uses the differences of gaussian pyramid technique, most keys can be found in regions where the intensities change. These regions can be small or large depending on where in the pyramid it was found. Running the `whos` command gives us the number of keys that are shown in this image, namely 1021 keys with 128 elements. The SIFT paper mentions that the number of expected keys is in the order of thousands for a 512 by 512 pixel image. Since our image is only 512 by 384 pixels, we can expect less. Hence the number that we found is probably in agreement with the paper. Also note

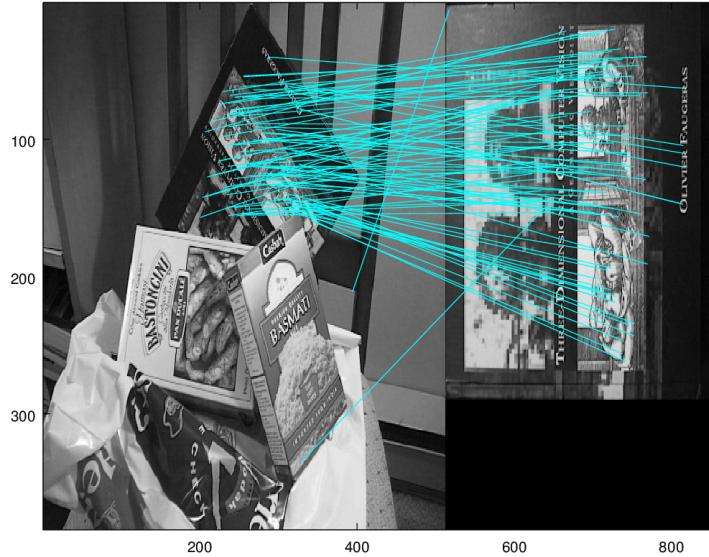


Figure 2: The matching features SIFT finds in “scene.pgm” with “book.pgm”

that the number of elements in the keys are different from the paper, keys in the paper have 160 instead of 128 elements. We assume the cause is that this implementation does not sample a higher plane in the image pyramid, since the difference in number of elements is in agreement with the number of samples taken in this plane.

The result of matching this scene with “book.pgm” is shown in Figure 2, it also shows the 98 matching keypoints with lines denoting matches. This amounts to 9.5% of the keypoints in “scene.pgm” and 11% of the keypoints in “book.pgm”. It also shows two mismatches. The first mismatch is between a corner from the book and a corner in the background of the scene. The second one is between part of the bowl of rice and part of the image of the book. It is understandable that these matches exist, since the features(colors/texture) are similar in the neighborhoods of the keypoints if you account for orientation.

The result of matching this scene with “basmati.pgm” is shown in Figure 3, it also shows the 34 matching keypoints with lines denoting matches. This amounts to 3.3% of the keypoints in “scene.pgm” and 5.9% of the keypoints in “basmati.pgm”. Only one mismatch exists in this case between the edge in the background of the scene and the edge of the basmati box. Again the mismatch is reasonable, since in the local neighborhoods of the matches are similar. The other mismatch from the book case does not exist, because that

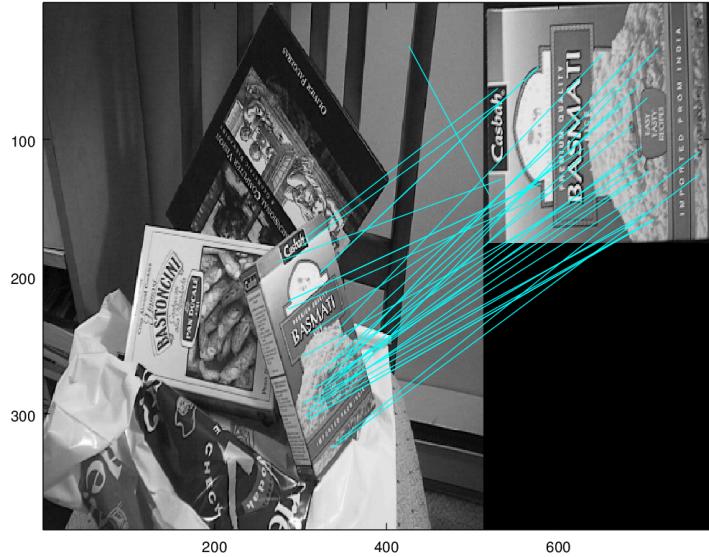


Figure 3: The matching features SIFT finds in “scene.pgm” with “basmati.pgm”

part of the book is occluded in the scene. Also note that in both case only a tiny percentage of the keys had to be matched in the scene to recognize the object.

Let us consider the effect of the resolution of the scene images with respect to the matched objects. Figure 4 and 5 show the results of SIFT matching. As can be seen in the images the number of matches is higher, but the number of mismatches also increased. If we ignore the mismatches for now, then we get that the number of matches increased approximately linearly with the number of keys in the scene image. Table 2 shows the number of matches per image combination, while Table 1 lists the number of keys per image. In this case the large image has about 3 to 4 times the amount of keys with respect to the normal image and the number of matches increased with a similar factor.

The number of mismatches seems to be related to the number of keys of the object relative to the number of keys of the scene. This is mainly based on the observation that in Figure 4 and 5 the larger street image seems to have more mismatches. Object images that are about have the size of the scene image seem to do ok, but smaller ones have plenty of mismatches. The extreme case is detail image 5 with the large street image, most of the matches are mismatches. So having higher resolution one image does not decrease the number of mismatches.

Normal street image



Large street image

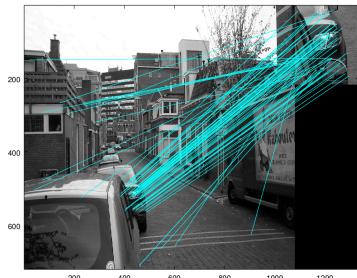
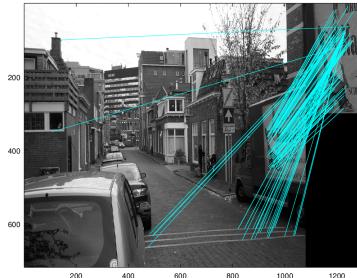
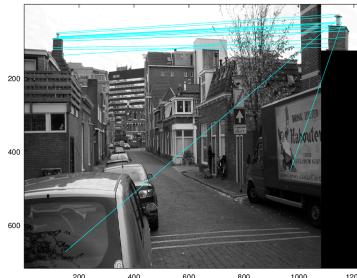
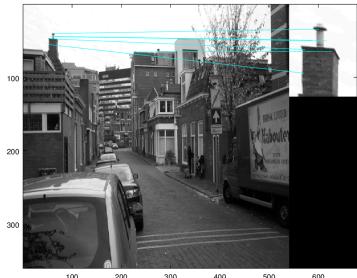
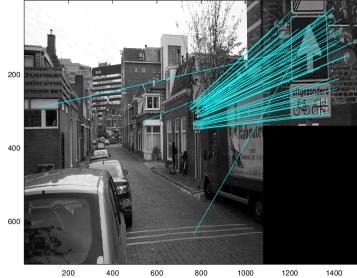


Figure 4: The SIFT matches of the street and large street scene with the detail 1 to 4 vertically images.

Normal street image



Large street image

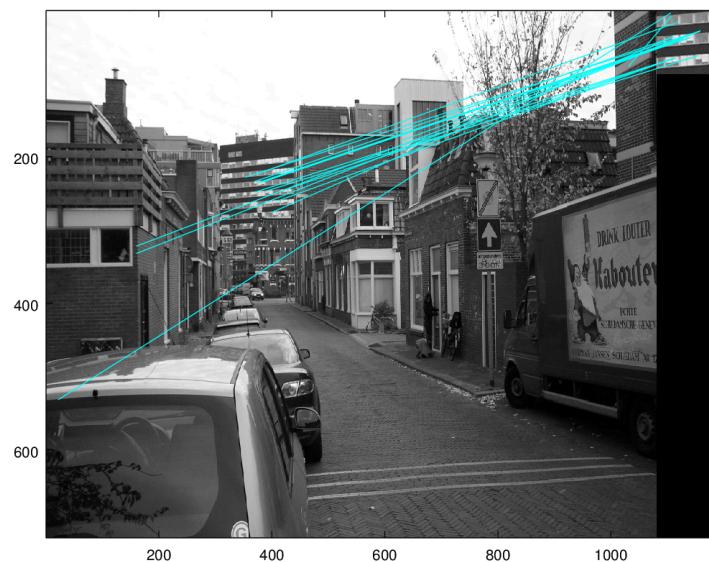


Figure 5: The SIFT matches of the street and large street scene with the detail 5 image.

Street	Large street	Detail 1	Detail 2	Detail 3	Detail 4	Detail 5
1452	5216	1592	46	588	368	126

Table 1: The number of keys per image.

	detail 1	detail 2	detail 3	detail 4	detail 5
Street	25	6	37	20	6
Large street	82	17	110	73	17

Table 2: The number of matches between the different street scenes.

## Exercise 2

## Exercise 3

Table 3 shows how the number of matches and the number of keys of the book changes with rotation. The number of matches is rather stable. This is good, because SIFT is supposed to be invariant to rotation. Of course this does not mean that it is completely insensitive to rotation. When sift samples the gradient it does so in a 4 by 4 grid with 8 different rotations thus generating 128 keys. The also is the cause of the insensitivity to rotation. The small fluctuations in the number of keys and matches is probably caused by how fitting the alignment is for detecting significant features.

rotation (degrees)	#matches	#keys
0	98	882
10	102	1020
20	95	1016
30	107	1031
40	101	1055
50	100	1055
60	102	1053
70	102	1019
80	107	1042
90	103	902
180	98	907
270	100	915

Table 3: The number of matches and keys of the book depending on the rotation.

## Exercise 4