Computer Vision Report UE 2 - Group 12:

Tomas Musil (1167504)
Andreas Wittmann (1225854)
Hanna Huber (0925230)

Assignment 4: Image Stitching

A. SIFT Interest Point Detection

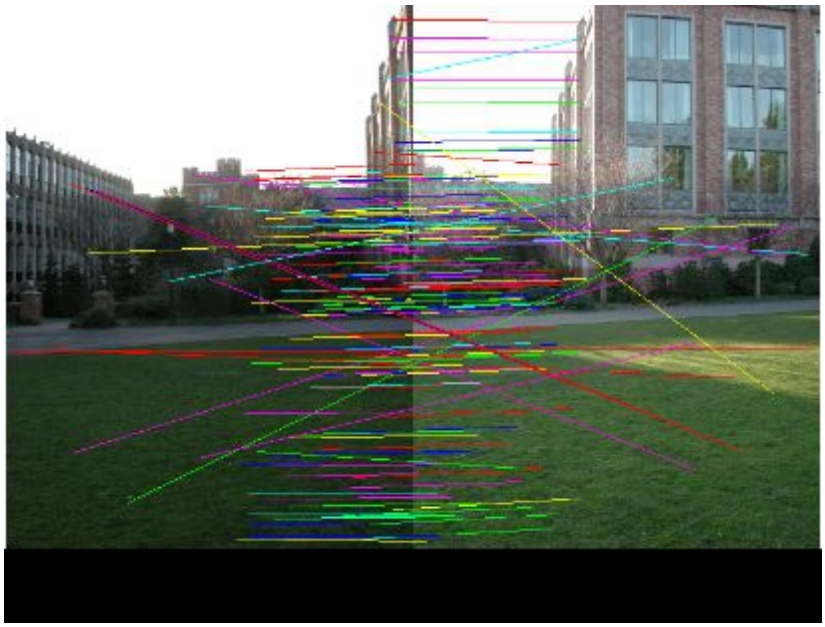
Figure 1 shows SIFT features detected on campus4.jpg:

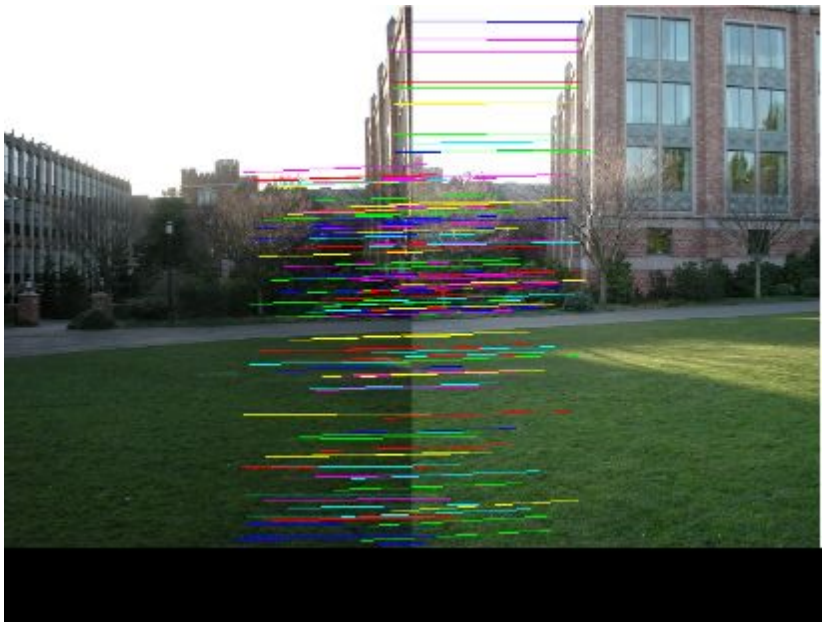
- Circle center → location of the keypoint
- Circle radius → scale where the keypoint was detected
- Lines → orientation of the keypoint

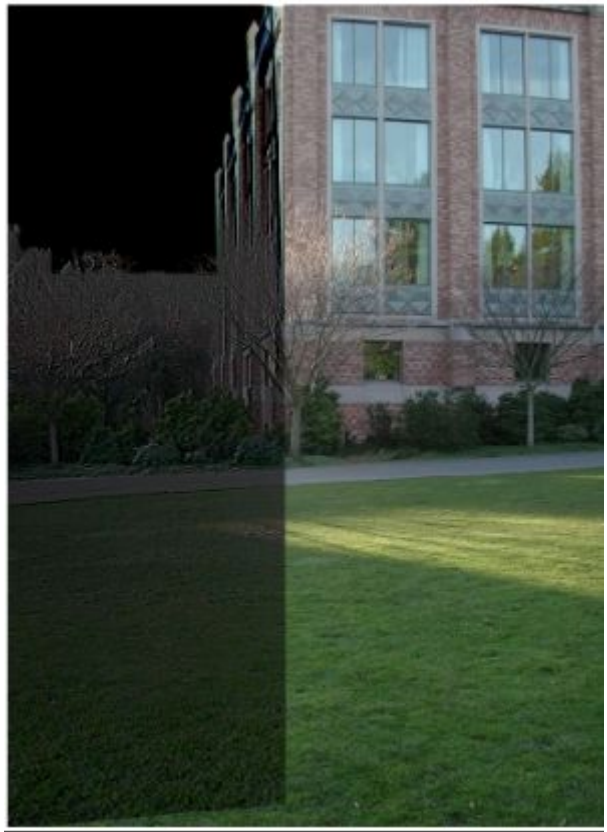
B. Interest Point Matching and Image Registration

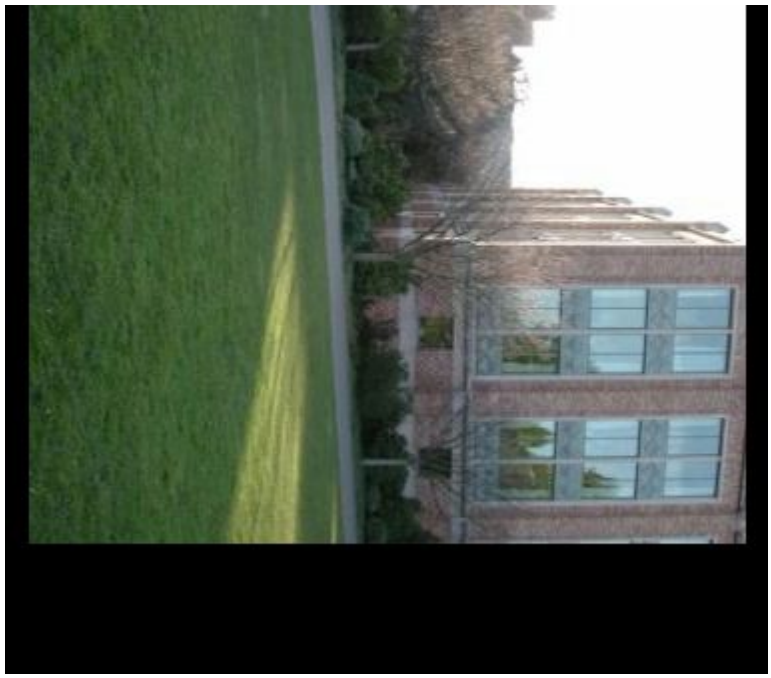
- Images ‘campus1.jpg’ and ‘campus2.jpg’ are aligned.
- The function vl_ubcmatch takes two sets of SIFT descriptors and returns matches. To find corresponding descriptors, the function compares every descriptor from the first set with every one from the second and computes their Euclidean distance as a measure of similarity. A descriptor d2 from the second set is considered a match for the descriptor d1 from the first set, if it is much more similar to it than the second best choice d2’. This is guaranteed by multiplying the Euclidean distance d(d1,d2) with a given threshold and checking if it is still smaller than d(d1,d2’).

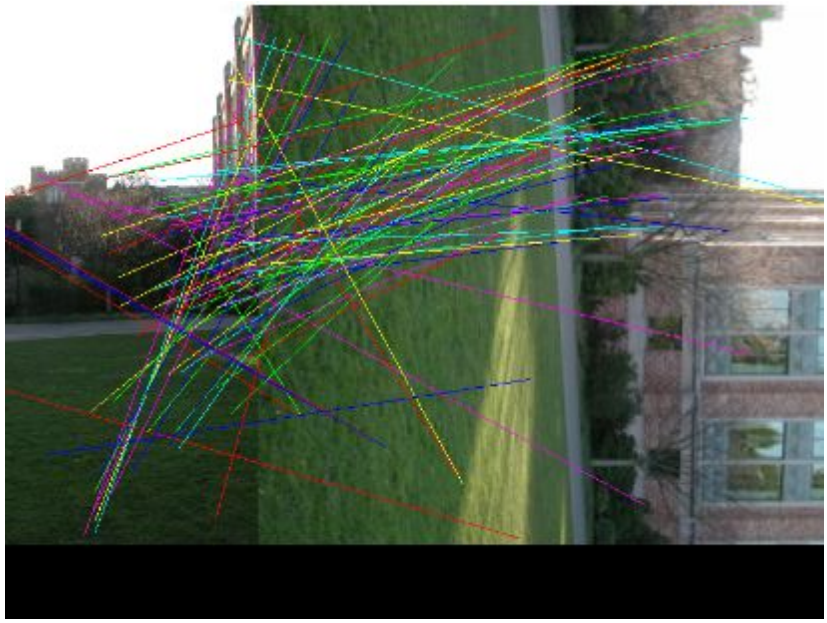
- The matches found by vl_ubcmatch are shown in Figure 2. Crossing lines imply that some pairs were matched incorrectly, because they correspond to a different image transformation compared to the majority of matches.
- The inliers corresponding to the best suited homography H are shown in Figure 3. Here, lines do not cross, because only those pairs are drawn where H produces (sufficiently) correct results, i.e. H*x1 is sufficiently close to x2.
- The absolute differences between the transformed first and the second image are shown in Figure 4. The contours show that they are well-aligned.
- Figures 6, 7 and 8 show the results for the alignment after rotating and resizing ‘campus2.jpg’ (see Figure 5). The absolute differences in Figure 8 show that the alignment procedure is scale and rotation invariant. The lines connecting the inliers in Figure 7 only cross according to the rotation, while the incorrect matches in Figure 6 correspond to completely different transformations.

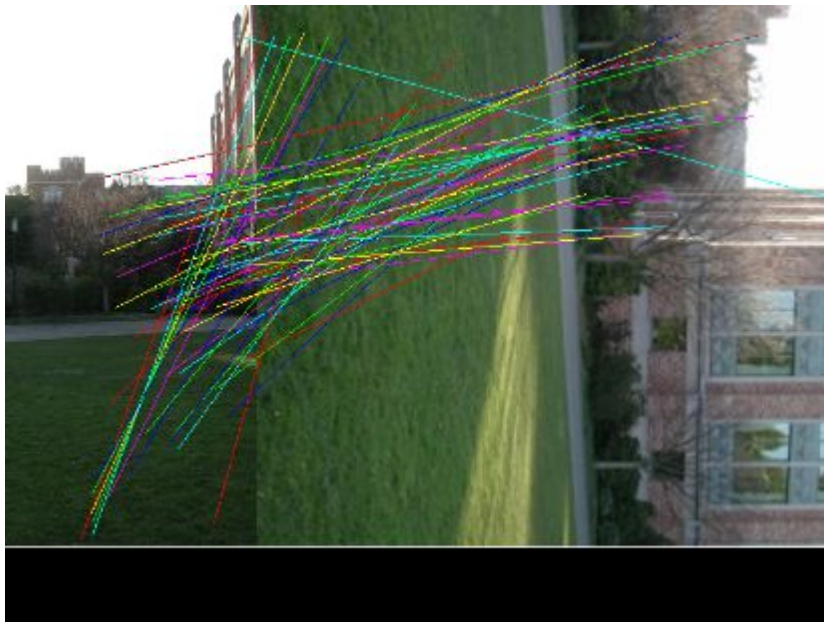


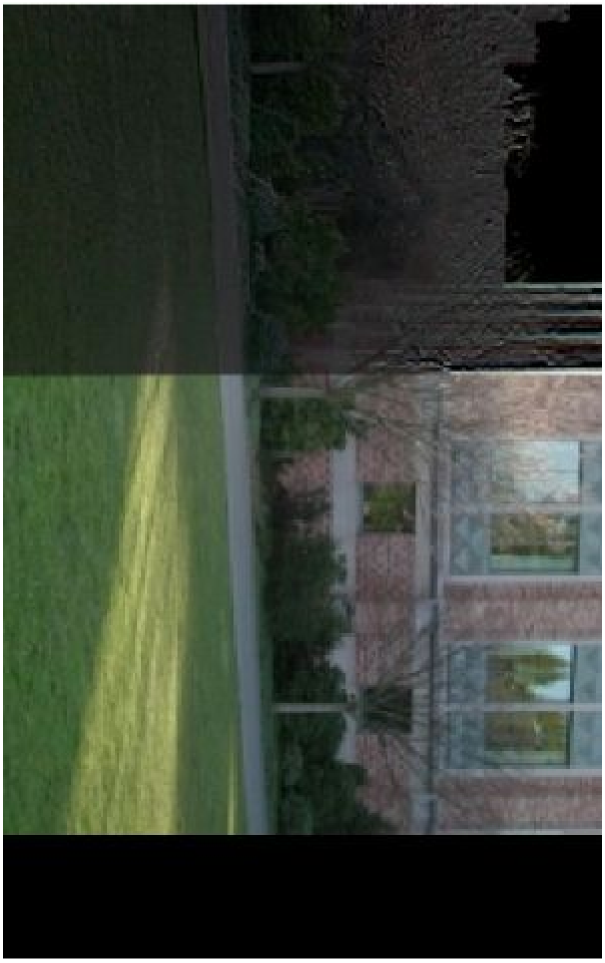












C. Image Stitching

Show and discuss the achieved results (with the two provided image sequences and your own sequence). The result might look quite realistic at a first glance but can you spot any errors by looking on details?

Campus:

At first glance this image looks very realistic. But there are two errors visible. The brightness is not everywhere constant. When the difference in brightness in images is big enough, it is possible to see the blending effect. The other error is visible by zooming in. There are areas, which occur twice (there are shifted only by few pixels) or are blurred. For example a building in two pictures can have different structures, because it is a difference, if the building is positioned in the middle, or on the side of the image. Putting this two images together can produce ghost or blurred images on overlapped areas. The best solution is to stitch images together, which are shifted only few pixels.

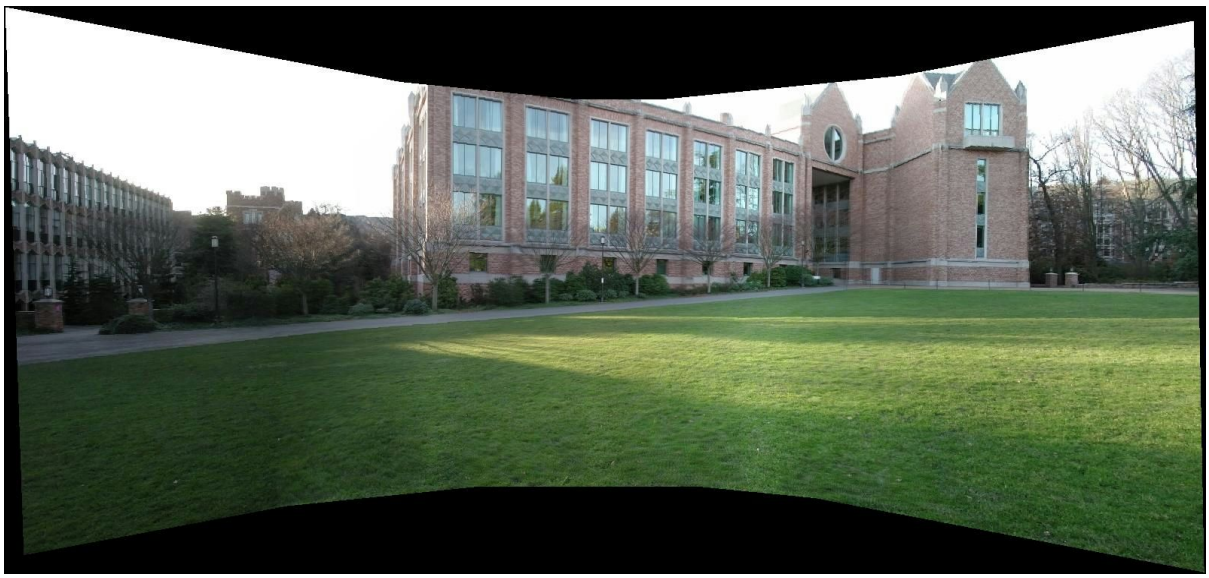
Office view:

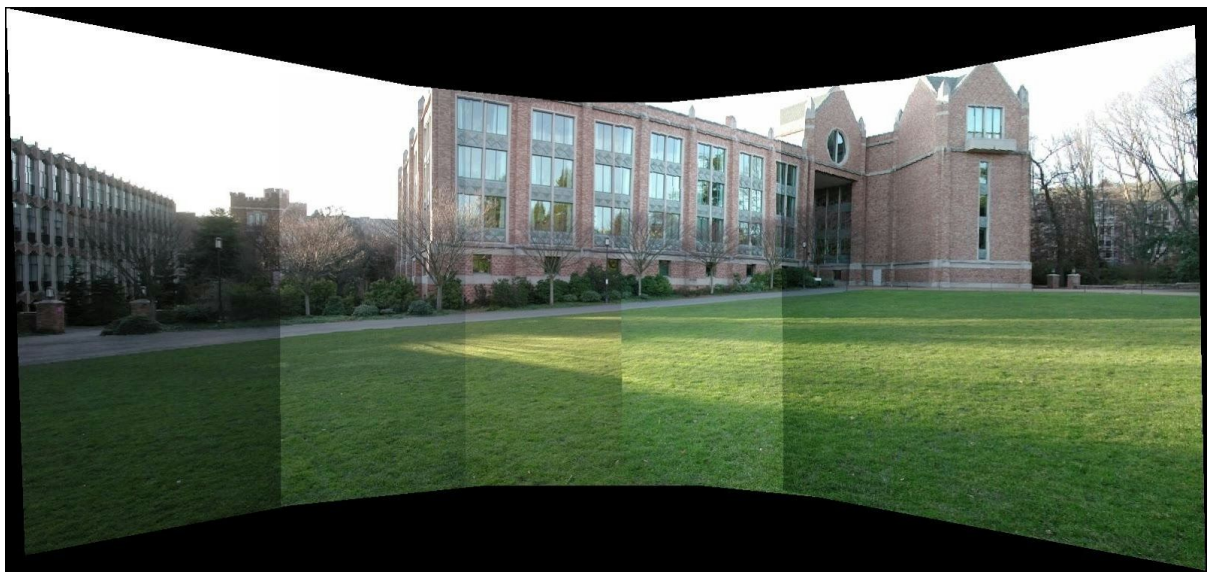
The result is very good. There are no errors visible, because the brightness in input images is constant and the shifted distance is very small.

Home view:

This result is also very realistic. Like in the first image there are any ghost images visible in overlapped areas.

Compare a result achieved with feathering to a result where no blending has been performed (i.e. the color values of only one image are taken for the stitched image). What is the difference of the two results?





In the non blended image there are hard edges visible, because of the difference of contrast from one image to the other. On the one hand the blended image looks more realistic, than the other one. On the other hand in the non blended image there cannot occur ghosted images, because there is no overlap of images.

Assignment 5: Scene Recognition with Bag of Visual Words

In general, this approach should achieve a classification rate of around 60%. Show and discuss your results in the report. Show the confusion matrix and describe if there are classes which can be more easily identified than others. Additionally, describe if there are class pairs which are confused with one another more often than other pairs.

conf_matrix:

	1	2	3	4	5	6	7	8
1	38	0	9	28	3	8	10	4
2	1	91	0	0	6	0	1	1
3	10	0	49	18	0	10	10	3
4	15	0	17	37	1	8	16	6
5	5	5	0	2	73	0	6	9
6	10	0	13	16	0	61	0	0
7	11	1	11	16	1	2	49	9
8	1	0	3	6	2	1	22	65

Our Bag of Visual Words classification algorithm has an average correct classification rate of ~60%. The classification rate varies due to randomly chosen SIFT features in the learning algorithm.
Images of the forest class could be classified the best with a classification rate of over 90%, while images of the bedroom class and images of the livingroom class will be classified correctly with less than 40%. Moreover, classes of natural images (forest and mountain) could be classified better than classes of man-made objects. This effect can not be generalised. If there would be an equal number of natural classes and man-made object classes, we would not expect better classification results of natural scenes.
Furthermore we can see a correlation between bedroom and livingroom images in the confusion matrix, which can be explained by similar furniture and materials in those scenes.

Take some own test images and investigate if they can be correctly classified by your simple scene recognizer. For instance, make a photo of your own kitchen or living room and classify it. You should at least test one own photo from three of the eight categories.

	1	2	3	4	5	6	7	8
1	0	0	0	0	0	0	0	0
2	0	2	0	0	0	0	0	0
3	1	1	0	0	0	0	0	0
4	0	1	0	0	0	0	1	0
5	0	0	0	0	0	0	0	0
6	0	1	0	0	0	0	0	1
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

We tested the algorithm with images of the class forest (class2), kitchen (class3), livingroom (class4) and office (class6). Every test class contained 2 images. Only the images of the forest class were classified correct. It surprises us, that one kitchen image and one office image is classified to the forest class.

Source of forest class images:
image_0001.jpg - https://unsplash.com/photos/ERbV0Rfn0PA, 19.01.2016
image_0002.jpg - https://unsplash.com/photos/jL1tJho10qc, 19.01.2016