Answers to questions in
Lab 3: Image Matching & 3D Reconstruction

Name: Timo Haubner   Program: DD2423

Instructions: Complete the lab according to the instructions in the notes and respond to the questions stated below. Keep the answers short and focus on what is essential. Illustrate with figures only when explicitly requested.

Good luck!

Question 1: To compute a homography, what are the computational steps involved? Why do you think we are interested in finding the smallest eigenvalue, but not in the other eigenvalues?

Answers:

1. Detect Features
In both images, we first extract distinctive keypoints (e.g. using SIFT).

2. Match Features
We then match the features between the two images, which gives me corresponding point pairs.

3. Construct the A Matrix
For every matched point pair, we generate two linear equations that describe how the homography should map one point to the other. Repeating this for all pairs gives a large matrix A with two rows per correspondence.

4*. Set Up the System Ah = 0*
With at least four point pairs, we can form a homogeneous linear system. Since the homography is only defined up to scale, one additional constraint is needed to avoid the trivial solution.

5. Formulate a Constrained Least-Squares Problem
The goal is to find the homography vector h that makes Ah as small as possible, subject to the normalization constraint. Therefore, the langrage-multiplier-method is used.

6. Solve the Eigenvalue Problem
The optimization leads to an eigenvalue equation involving the matrix derived from A. The vector h we want must be an eigenvector of this matrix.

7. Choose the Eigenvector of the Smallest Eigenvalue

The smallest eigenvalue corresponds to the direction in which the residual error is minimized. The associated eigenvector gives the homography because it produces the best solution to the system. All other eigenvalues represent directions with larger error, so they are not relevant.

Question 2: How much image noise can you typically have before the error becomes significant? Is it preferable to use a few or many feature points when creating A?

Answers:

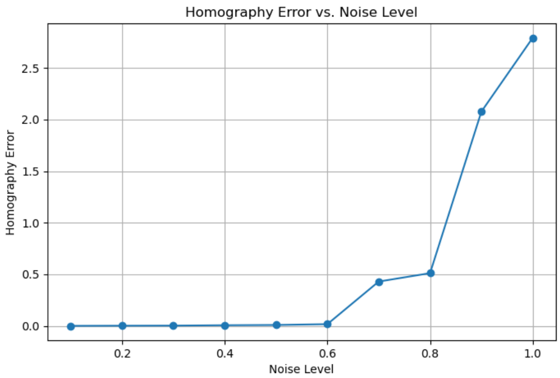

Figure 1

In Figure 1, it can be seen that the homography error remains relatively stable up to a noise level of about 0.5–0.6, after which it increases sharply. However, it should be noted that even small errors can have a significant impact on the accuracy of the estimated homographies. Since no outliers are present in this example, it is advantageous to use more points in A. With more point correspondences, additional equations can be included in A, resulting in a more robust estimation of the smallest eigenvector and making the method less sensitive to noise.

Question 3: How many outliers can you typically have before the errors start to increase too much? Given what you try to optimize, how can the sensitivity to outliers be explained?

Answers:

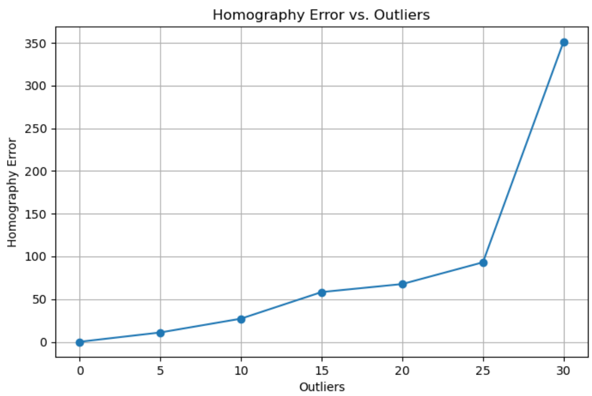

Figure 2

In Figure 2, it can be noted that the error increases much more sharply in the presence of outliers than it does with noise. With as few as five outliers among 100 data points, the result becomes virtually unusable. Outliers introduce a large error and therefore exert a strong influence on the minimization of the error of Ah, which in turn causes the estimated homography to become heavily distorted.

Question 4: If you have num = 100 points in total and noutliers = 50 outliers, how many iterations do you typically have to do until you find a solution with a small enough homography error?

Answers:

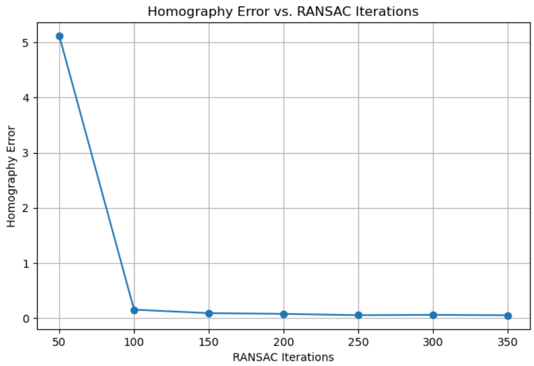

Figure 3

In Figure 3, it can be observed that RANSAC is able to significantly reduce the homography error when the number of iterations is sufficiently large. With 50 iterations, the error is still around 5.1, whereas with 100 iterations it drops to approximately 0.15 and is therefore comparatively low. To approximate the required number of RANSAC iterations, the method introduced in the lecture can be used, which has been implemented in the function task_2_outlier_test().

Question 5: Why does RANSAC make the homography estimation so much more robust compared to the original implementation? Try to come up with a short, but easily understood, explanation.

Answers:

When using RANSAC, not all points are used for estimating the homography. Instead, the algorithm repeatedly computes homographies from small random subsets of point correspondences and selects the best option. This prevents outliers from influencing the final homography, as only models that are consistent with the majority of inlier points are considered (meaning outliers get ignored).

Question 6: What do you observe in the results? Do you get an overlap similar to what is shown in Fig. 2? Can you somehow exploit the results even in regions for which there is no overlap?

Answers:

Applying the method to the images books1 and books2 resulted in significantly less accurate outcomes compared to flat images such as img1. In large parts of the scene, it is clearly visible that the overlaid images diverge substantially. This is because the depth structure of the scene cannot be represented by a single homography.

However, in certain regions the homography still performs reasonably well—for example, on the book Digital Image Processing. The surface of this book is approximately planar, and the viewpoints in both images differ only slightly, which makes the homography a good approximation in this local area. For such individual planar regions, the homography result could still be used meaningfully.

Question 7: What are the similarities and differences between the code you wrote for estimating a fundamental matrix and a homography?

Answers:

The estimation of the fundamental matrix and the approximation of homographies are two distinct mathematical concepts. Nevertheless, they share several methodological similarities. In both cases, matched point pairs are used to extract the required information. Each point pair contributes one equation for the fundamental matrix and two equations for homographies; these equations are then assembled into a linear system.

The overall procedure for solving the resulting systems is essentially the same: first, a system of equations is formulated in matrix form; then the square of this matrix is computed, and finally, the eigenvector corresponding to the smallest eigenvalue is determined. As a result, large parts of the implementation can be reused for both tasks.

Question 8: If you have 100 generated feature matches, how many outliers can you typically have before you get significant errors in the estimated fundamental matrices? Can you see a difference with normalization versus without?

Answers:

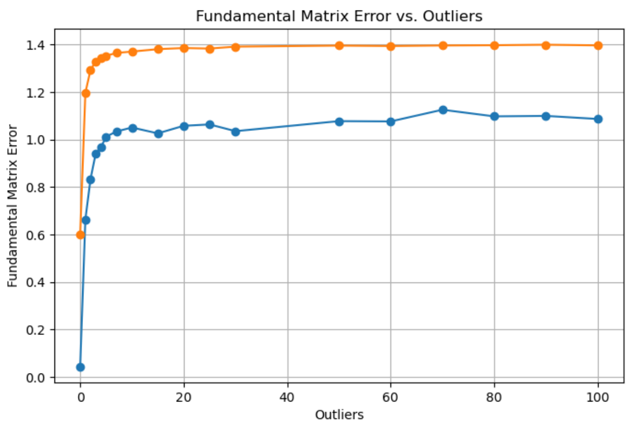

Figure 4

Figure 4 shows the fundamental matrix error as a function of the number of outliers, once with normalization (blue curve) and once without normalization (orange curve). For small numbers of outliers, the error increases rapidly and then remains approximately constant from around 10 outliers onward. From this, it can be concluded that this plateau at roughly 1.0 already renders the fundamental matrix unusable, as the error does not increase any further even at 100% outliers. The effect of normalization is clearly visible: the error with normalization is consistently lower than the error without normalization.

Question 9: If you have num = 100 points in total and noutliers = 50 outliers, how many iterations do you need until you find a solution with a small enough error in the estimated fundamental matrix? What is the difference compared to when you tried to find homographies?

Answers:

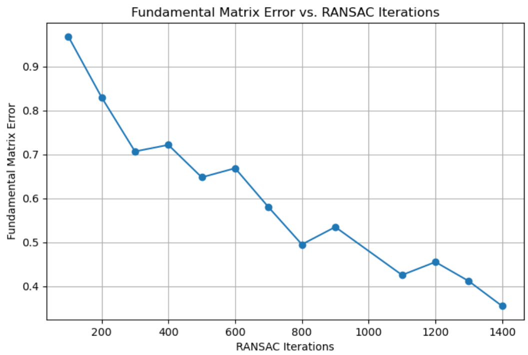

Figure 5

Figure 5 shows that the error decreases fairly consistently for up to 1400 iterations. The number of required iterations is significantly higher than for the homographies, where no substantial improvement could be observed after far fewer iterations. This is because 8 point correspondences are used instead of 4. Since all selected points are expected to be inliers, the probability of drawing a sample containing only inliers is considerably lower when 8 points are required per sample.

Question 10: How large a fraction of all original feature pairs seems to be classified as inliers? Do you see any obvious mismatch being classified as inlier and thus not removed? How much do the results depend on which pair of images you tried?

Answers:

The proportion of inliers varied between 37% and 75%, indicating a strong dependence on the images used. For example, the quality of the image features detected by SIFT may influence the accuracy. After applying RANSAC, the point correspondences proved to be highly accurate; only very few mismatches could be identified.

Question 11: If we assume the angle between two books in books1.jpg is 90°, is it possible to tell approximately how large the real focal length is by varying the focal length used for reconstruction?

Answers:

Yes, it is possible to estimate the approximate real focal length. By testing different assumed focal lengths and examining the resulting 3D reconstruction, only the correct focal length produces an angle between the book surfaces that is close to 90°. Incorrect focal lengths distort the geometry significantly, causing the reconstructed angle to deviate strongly from the expected value. In practice, however, this approach proved to be challenging, as it is difficult to interpret the 3D reconstruction in a meaningful way and relate it to the original images.

Question 12: How sensitive is the reconstruction to the choice of images and the characteristics of the 3D scene? Can it handle scenes primarily consisting of a flat surface, such as the scenes for which we earlier used homographies?

Answers:

The reconstruction quality depends strongly on the choice of images and on the depth structure of the scenes. This became evident in the experiments on “img1” and “img2”, which depict an almost planar scene. Using homographies, the results were very good; however, applying triangulation did not yield any meaningful output. The reason is that epipolar geometry relies on depth variations to estimate the fundamental matrix. In (approximately) planar scenes, this procedure becomes degenerate or leads to large errors.

Question 13: What do you suggest one could do to further improve the reconstructions? Could we somehow build on what has been done so far in the lab to get considerably better results?

Answers:

The most obvious improvement is to calibrate the cameras used. This would, for example, eliminate the error introduced by the unknown focal length. Additionally, more features could be incorporated, or more than two cameras could be combined to obtain even more accurate results.
