Computer Vision - FMAN85
Assignment 4 - Spring 2020

Model Fitting and RANSAC

Hicham Mohamad, hi8826mo-s
hsmo@kth.se

March 9, 2020

1 Introduction

In this assignment, we are going to exercise on the basic elements of **Model Fitting**. We will use the general and very successful robust estimator RANdom SAmpling Consensus **RANSAC** algorithm to fit various models such as planes, homographies and essential matrices.

The resulting plots and values in this report are obtained by running the implemented Matlab scripts `ass4_CE1.m` and `ass4_CE2.m`.

2 Plane Fitting

2.1 Exercise 1

2.1.1 Degrees of freedom

In RANSAC, the size of the sample set depends on the degrees of freedom of the model. If we want to fit a **3D plane** to a set of points, the model has 3 degrees of freedom. This can be concluded from the equation of the a 3D plane

$$ax + by + cz + d = 0 \tag{1}$$

where we suppose that (x_i, y_i, z_i) are the measurements of 3D-points belonging to the plane, i.e. 3 variables.

2.1.2 Number of sample sets

If the point set contains 10% outliers, the task is to find how many **sample sets** we need to draw to achieve a success rate of 98%. To solve this question it is necessary to compute the probability of drawing 3 inliers and the probability of failing to do so, i.e.

P(of drawing 3 inliers) $= 0.9^3$.

P(of failing to do so) $= 1 - 0.9^3$.

P(of failing n times) $= (1 - 0.9^3)^n < 0.02$, i.e. to achieve a success rate of 98%.

From this we can write using the logarithm calculus

$$\log[(1-0.9^3)^n] < \log(0.02) \implies n\log[(1-0.9^3)] < \log(0.02) \implies n > \frac{\log(0.02)}{\log[(1-0.9^3)} = 3 \tag{2}$$

Thus, we see as expected that we need to draw 3 sample sets, equivalent to the number of degrees of freedom, in order to achieve a success rate of 98%.

2.2 Exercise 2 - Total Least Squares problem

In this section, we are dealing with the formula for the solution of the total least squares problem which we need to use them in the computer exercise. Suppose that x_i, y_i, z_i for $i = 1, 2, \cdots, m$ are 3D points that we want to fit to a plane (a, b, c, d). Thus, we want to solve

$$\begin{cases} \min \quad \sum_{i=1}^{m}(ax_i + by_i + cz_i + d)^2 \\ \\ \text{such that} \quad a^2 + b^2 + c^2 = 1 \end{cases} \tag{3}$$

That means to minimize the sum of **squared distance** from the plane to the points.

2.2.1 The optimal d

It can be shown by taking the derivative with respect to d

$$2\sum_{i=1}^{m}(ax_i + by_i + cz_i + d) = 0 \implies \sum_{i=1}^{m} d = md = -\sum_{i=1}^{m} ax_i - \sum_{i=1}^{m} by_i - \sum_{i=1}^{m} cz_i \tag{4}$$

This gives that the **optimal** d must fulfill

$$d = -\frac{1}{m}\sum_{i=1}^{m} ax_i - \frac{1}{m}\sum_{i=1}^{m} by_i - \frac{1}{m}\sum_{i=1}^{m} cz_i = -(a\bar{x} + b\bar{y} + x\bar{z}) \tag{5}$$

given a, b and c. Where

$$(\bar{x}, \bar{y}, \bar{z}) = \frac{1}{m}\sum_{i=1}^{m}(x_i, y_i, z_i) \tag{6}$$

2.2.2 Solution to the TLS is an eigenvector

By assuming $(\tilde{x}, \tilde{y}, \tilde{z}) = (x_i - \bar{x}, y_i - \bar{y}, z_i\bar{z})$ and substituting into 3, we get

$$\begin{cases} \min \quad \sum_{i=1}^{m}(a\tilde{x}_i + b\tilde{y}_i + c\tilde{z}_i)^2 \\ \\ \text{such that} \quad 1 - a^2 + b^2 + c^2 = 0 \end{cases} \tag{7}$$

this can be considered a **constrained optimization problem** of the type

$$\begin{cases} \min \quad f(t) \\ \text{such that} \quad g(t) = 0 \end{cases} \tag{8}$$

which can be solved by using **Lagrange multipliers** method and the solution is given by

$$\Delta f(t) + \lambda \Delta g(t) = 0 \tag{9}$$

We need to show that the solution to the optimization problem in 7 must be an **eigenvector** of the matrix

$$M = \sum_{i=1}^{m} \begin{pmatrix} \tilde{x}_i^2 & \tilde{x}_i\tilde{y}_i & \tilde{x}_i\tilde{z}_i \\ \tilde{y}_i\tilde{x}_i & \tilde{y}_i^2 & \tilde{y}_i\tilde{z}_i \\ \tilde{z}_i\tilde{x}_i & \tilde{z}_i\tilde{y}_i & \tilde{z}_i^2 \end{pmatrix} \tag{10}$$

corresponding to the smallest eigenvalue. In fact we can rewrite the objective function in the optimization problem in 7 in matrix form as

$$\begin{cases} \min \quad t^T M t \\ \text{such that} \quad 1 - t^T t = 0 \end{cases} \qquad \text{where } t = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{11}$$

Using Lagrange in 9, we can write

$$\Delta(t^T M t) + \lambda \Delta(1 - t^T t) = 0 \implies 2Mt - 2\lambda t = 0 \implies Mt = \lambda t \tag{12}$$

This result corresponds to the eigenvector definition, where t must be the eigenvector of the matrix M. Now looking at the objective function in 11, we can write

$$Mt = \lambda t \implies t^T M t = t^T \lambda t = \lambda t^T t = \lambda \tag{13}$$

which means we need to select the eigenvector with the **minimal eigenvalue**.

2.3 Computer Exercise 1

In this task, we consider two images of a house and a set of 3D points from the walls of the house. The goal of this exercise is to estimate the **location of the wall** with the most 3D points. The file `compEx1data.mat` contains cameras P, inner parameters K for both cameras, scene points X and some extra points x from image 1.

a) Solving the Total Least Squares - All the points

After computing the homogeneous 3D-points X, we solve the total least squares problem with all the points. Then, we compute the RMS distance between the 3D-points and the plane using the implementation of the distance formula

$$e_{RMS} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\frac{(ax_i + by_i + cz_i + d)^2}{a^2 + b^2 + c^2}} = 0.5168 \tag{14}$$

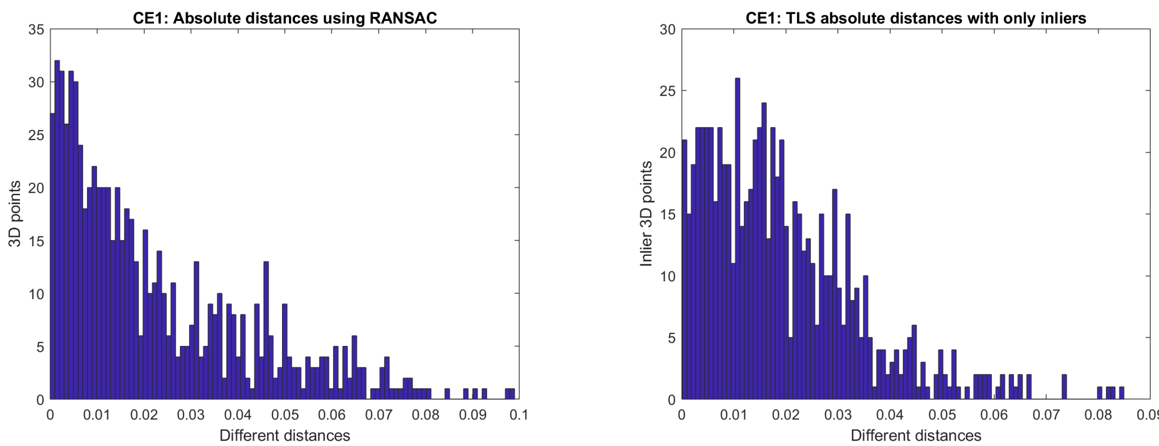
b) Fitting the plane using RANSAC

Random sample consensus (RANSAC) is a method for removing outliers. It is nice to show here he outline of the RANSAC algorithm as follows:

1. Randomly select a **small subset** of measurements and solve the problem using only these.
2. Evaluate the **error residuals** for the rest of the measurements under the solution from 1. The **Consensus Set** for this solution is the set of measurements with error residuals less than some predefined **threshold**.
3. **Repeat a number of times** and select the solution that gives the largest consensus set.

After implementing RANSAC algorithm in Matlab, we use it to robustly fit the plane to the 3D points X. We consider the **predefined threshold** such that a 3D point is accepted as an inlier if its distance to the plane is less than 0.1.

Our goal is that we want to fit a plane to the set of points. Thus, we randomly select 3 points and fit a plane to these in each RANSAC iteration. As we have seen in Exercise 1, the number of iterations required to find at least one set of only inliers with probability $p = 98\%$ is 3 iterations, but in practice it is a good idea to run more iterations than what is needed since, because of noise, not all inlier sets work equally well for estimating the solution. At the end, we get the largest **Consensus set** with number of inliers equal to 734.



(a) (b)

Figure 1: a) Plot of the absolute distances between the plane obtained with RANSAC and the points in a histogram with 100 bins. b) The absolute distances between the plane and the inliers 3D points after solving the Total Least Squares problem.

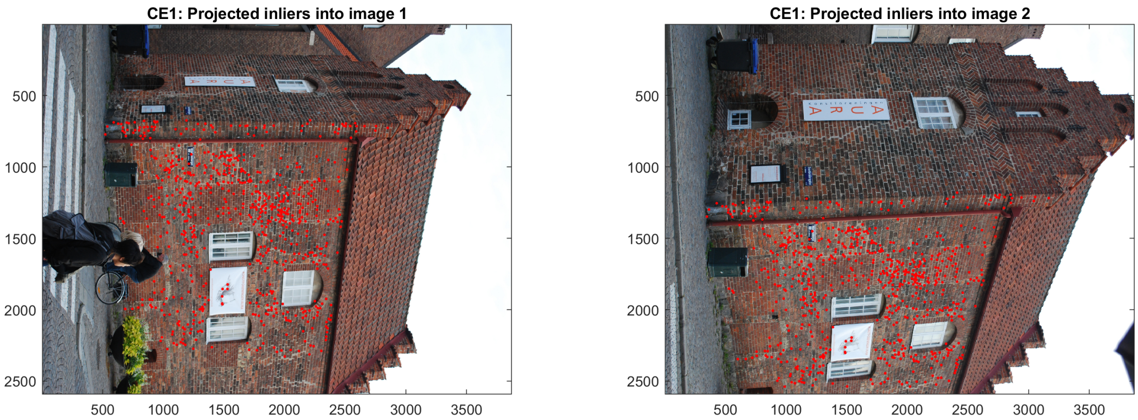
The resulting RMS distance between the plane obtained with RANSAC and the 3D points is 0.03009. It is clear that this error residual is much smaller than that one computed before without using RANSAC.

In Figure 1a it is illustrated the resulting plot of the **absolute distances** between the plane and the points in a histogram with 100 bins.

c) Solving the Total Least Squares - Only inliers

After solving the total least squares problem with only the inliers, we compute the RMS distance between the 3D-points and the new plane and obtain **0.024381**. The resulting plot of the absolute distances between the plane and the points in a histogram with 100 bins, is shown in Figure 1b. We can see that the estimation produced with RANSAC gives better result, with less residual errors.

In Figure 2 we can see the projection of the inliers into the images. These are located on the same wall of the house.



(a) (b)

Figure 2: Projection of the inliers into image1 and image2 .

d) Homography from camera 1 and camera 2

After normalizing the cameras, we compute a homography H from camera 1 to camera 2 using the method in assignment 1, (Exercise 6,). If $\mathbf{x}_1 \sim P_1\mathbf{X}$ and $\mathbf{x}_2 \sim P_2\mathbf{X}$ and assuming that $\mathbf{X}$ belongs to the plane

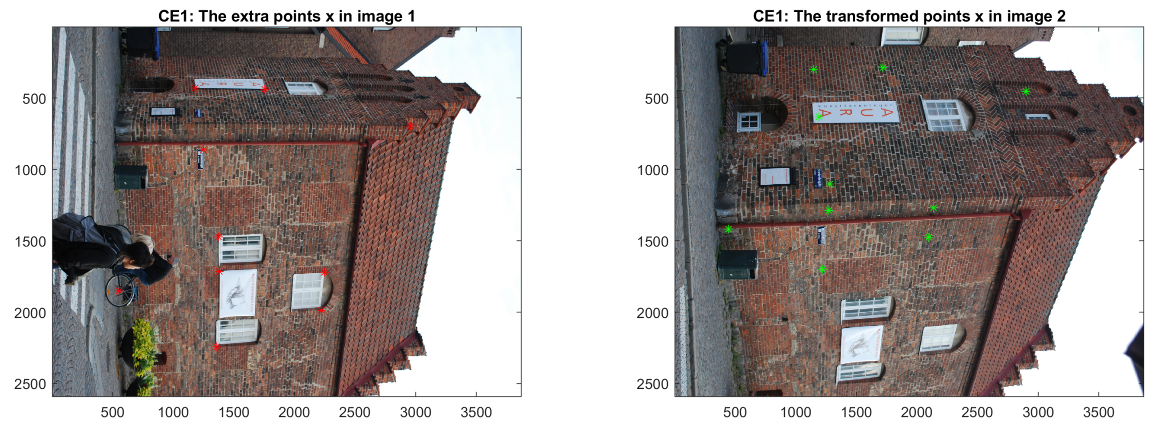
$$\Pi = \begin{pmatrix} \pi \\ 1 \end{pmatrix} \qquad \pi \in \mathbb{R}^3 \tag{15}$$

we can write

$$H = (R - t\pi^T) = \begin{pmatrix} 0.9939 & -0.0573 & 0.0185 \\ -0.0294 & 0.7758 & 0.0880 \\ 0.0422 & 0.3423 & 0.8436 \end{pmatrix} \quad \text{where } P_2 = \begin{bmatrix} R & t \end{bmatrix} \tag{16}$$

In Figure 3a we plot the points x in image 1. Then we transform the points using the obtained homography H and plot them in image 2 as shown in Figure 3b.

Which ones seem to be correct, and why? ???



(a) (b)

Figure 3: a) Plot of the points x in image 1. b) The transformed points x in image 2 after using the computed homography H.

3 Robust Homography Estimation and Stitching

3.1 Exercise 3

In this task, we need to show that if two cameras $P_1 = [A_1 \quad t_1]$ and $P_2 = [A_2 \quad t_2]$ have the same camera center then there is a homography H that transforms the first image into the second one.

By assuming that A_1 and A_2 are invertible, we can write these two cameras with the same center as

$$\begin{cases} P_1 C = \begin{bmatrix} A_1 & t_1 \end{bmatrix} \begin{bmatrix} C \\ 1 \end{bmatrix} = A_1 C + t_1 = 0 \\ \\ P_2 C = \begin{bmatrix} A_2 & t_2 \end{bmatrix} \begin{bmatrix} C \\ 1 \end{bmatrix} = A_2 C + t_2 = 0 \end{cases} \tag{17}$$

From this we can conclude that

$$C = -A_1^{-1} t_1 = -A_2^{-1} t_2 \tag{18}$$

Then we can write the projected points in in P_1 and P_2

$$\begin{cases} \mathbf{x}_1 \sim P_1 \mathbf{X} = \begin{bmatrix} A_1 & t_1 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix} = A_1 X + t_1 = 0 \\ \\ \mathbf{x}_2 \sim P_2 \mathbf{X} = \begin{bmatrix} A_2 & t_2 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix} = A_2 X + t_2 = 0 \end{cases} \tag{19}$$

From this we can conclude that

$$X \sim A_1^{-1} \mathbf{x}_1 - A_1^{-1} t_1 \sim A_2^{-1} \mathbf{x}_2 - A_2^{-1} t_2 \tag{20}$$

By using the result obtained in 18 we can rewrite

$$X \sim A_1^{-1} \mathbf{x}_1 - A_1^{-1} t_1 \sim A_2^{-1} \mathbf{x}_2 - A_1^{-1} t_1 \implies X \sim A_1^{-1} \mathbf{x}_1 \sim A_2^{-1} \mathbf{x}_2 \tag{21}$$

From this we conclude that there is a homography H that transforms the first image into the second one,

$$\mathbf{x}_2 \sim A_2 A_1^{-1} \mathbf{x}_1 \implies H = A_2 A_1^{-1} \tag{22}$$

3.2 Exercise 4

a) Degrees of freedom

When we want to find a homography H that transforms one 2D point set into another, this homography has eight **degrees of freedom**. This because for one 2D-point $(x_1, y_i, 1)$ we have 3 elements and consequently the matrix H should have 9 elements, but the scale is arbitrary and we can assume that one of them is fixed. Therefore it only has 8 degrees of freedom.

b) Minimal number of point correspondences

In homography estimation we want to find the minimal number of **point correspondences** that we need to determine the homography. Given two sets of points y_i and x_i that are related by a homography H we want to solve

$$\lambda_i y_i = H x_i \quad i = 1, \cdots, n \tag{23}$$

For n points we therefore have $3n$ **equations** (3 for each point projection), but each new projection introduces one additional unknown λ_i and consequently this makes $8 + n$ **unknowns**. The problem can therefore be solved if

$$3n \geq 8 + n \implies n \geq 4 \tag{24}$$

Thus, we need at least 4 point correspondences to be able to determine the homography H.

c) Iterations of RANSAC

Assuming the number of incorrect correspondences is 10%, the task here is to find how many iterations of RANSAC we need to find an outlier free sample set with 98% probability.

To solve this question it is necessary to compute the probability of drawing 4 inliers and the probability of failing to do so, i.e.

P(of drawing 4 inliers) $= 0.9^4$.

P(of failing to do so) $= 1 - 0.9^4$.

P(of failing n times) $= (1 - 0.9^4)^n < 0.02$, i.e. to achieve a success rate of 98%.

From this we can write using the logarithm calculus

$$\log[(1 - 0.9^4)^n] < \log(0.02) \implies n \log[(1 - 0.9^4)] < \log(0.02) \implies n > \frac{\log(0.02)}{\log[(1 - 0.9^4)} = 3.665 \tag{25}$$

Thus, we see that we need at least **4 iterations** of RANSAC in order to find an inlier sample set with 98% probability.

3.3 Computer Exercise 2

In this exercise you will use RANSAC to estimate homographies for creating **panoramas**. Specifically we will need to use `VLfeat` as in assignment 2 to generate **potential matches**, and then determine inliers using RANSAC.

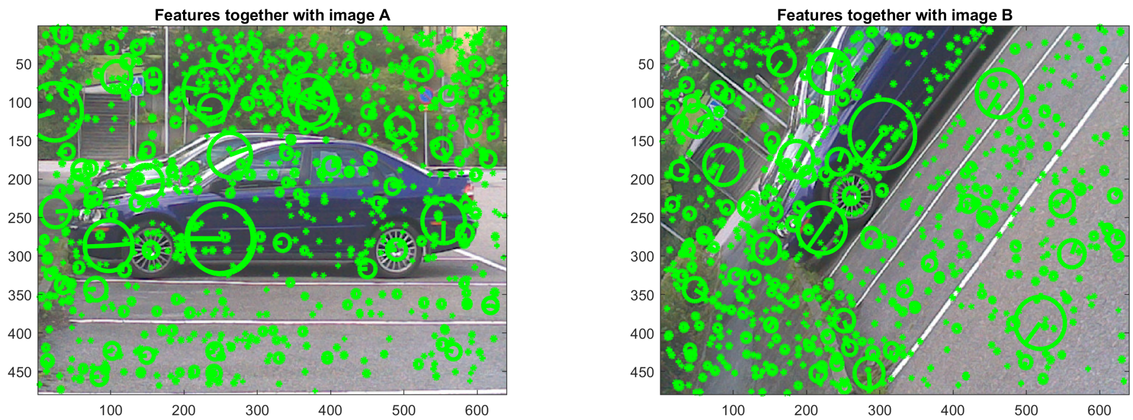
After loading the two images in Matlab and displaying them in Figure 4. We can notice that the images are partly overlapping. The goal is to place them on top of each other. To achieve this goal we need first to use VLFeat to compute `SIFT features` for both images and match them.

a) Features and matches in the images

After computing the SIFT features for the two images as shown in Figure 5, we find 947 detected features for image A and 865 for image B. Then we find 204 matching points.

(a) (b)

Figure 4: Image a.jpg, b.jpg



(a) (b)

Figure 5: plot of the images A.jpg and B.jpg together with the features after computing SIFT using vlfeat.

b) Homography estimation and images stitching

Now we should find a homography describing the transformation between the two images. This goal can be achieved in two steps:

1. Because not all matches are correct, we need to use RANSAC to find a set of good correspondences (inliers).
2. To estimate the homography we use DLT with a minimal number of points needed to estimate the homography.

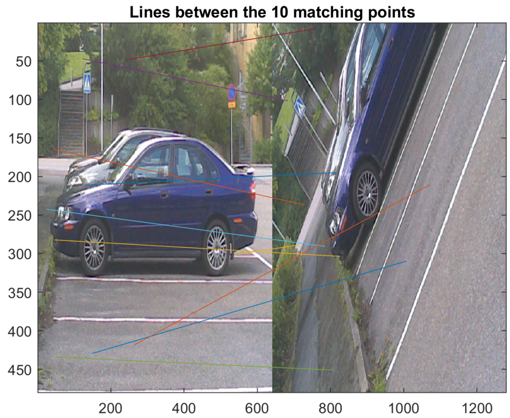
In this case, the least squares system will have an exact solution, so normalization does not make any difference. After implementing the 2 steps mentioned above and assuming a reasonable threshold for inliers as 5 pixels, we get the best homography H

$$H = \begin{pmatrix} -0.0026 & -0.0037 & 0.2122 \\ 0.0043 & -0.0027 & -0.9771 \\ 0.0000 & 0.0000 & -0.0056 \end{pmatrix} \tag{26}$$

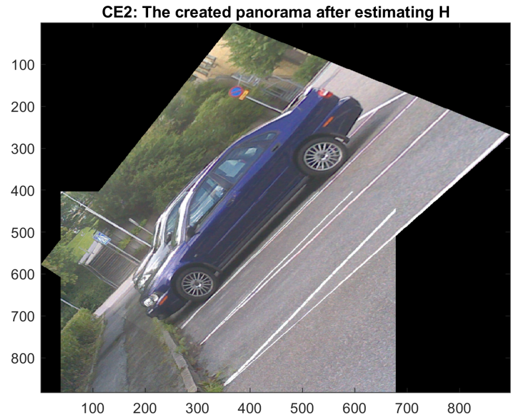
Using the estimated homography H, the couple images are transformed to a common coordinate system where the panorama is created by placing the images on the top of each other (image stitching), as shown in Figure 6b.

c) Consensus set

With 10 iterations of RANSAC, **119 inliers** are found in the largest consensus set.



(a)



(b)

Figure 6: a) Plot of the two images next to each other and the lines between the 10 matching points selected randomly. It seems that all of the matches appear to be correct. b) The created panorama by placing the images on top of each other.

References

[1] Carl Olsson, Computer Vision - FMAN85, Lectures notes: `https://canvas.education.lu.se/courses/3379`

[2] Hartley, Zisserman, Multiple View Geometry, 2004.

[3] Szeliski, Computer Vision - Algorithms and Applications, Springer.