3D recovery of urban scenes. Session 3

Guillem Capellera[1], Johnny Núñez[1], Anna Oliveras[1], and Razvan Apatean[1]

1. Universitat Pompeu Fabra, Barcelona, Spain

1 Introduction

The objective of this session is to estimate the fundamental matrix (F) which relates two different viewpoints of the same scene. In this case, it is calculated with the normalized 8-point algorithm, so a minimum of eight keypoint correspondences will be used. In order to make the algorithm more robust, RANSAC is used. In addition, the fundamental matrix estimation will be verified using the epipolar lines. Finally, these concepts will be used to solve the Photo-sequencing problem. Note that this is a simplification of the method proposed in Dekel, Moses, and Avidan, 2013.

2 Estimation of the fundamental matrix

2.1 DLT algorithm

Given a set of point correspondences between a pair of images, the goal is to estimate the fundamental matrix using the normalized 8-point algorithm.

First we created an auxiliary function, similarly to last week, to normalize the points. This function is called $NormalizePoints$ and its pseudo-code is shown in Algorithm 1. Then to estimate the fundamental matrix, the normalized version of the 8-point algorithm is used. The pseudo-code for this function is shown in Algorithm 2.

In order to verify that the function to estimate the fundamental matrix works correctly, we used an example where the ground truth image was known. In this example we have the following camera matrices: P = [I|0] and P' = [R|t], where I is the identity matrix, R is the rotation matrix and t is the translation vector. We know that by definition, F = $K'^{-T}EK^{-1}$, where E is the essential matrix, and in our case, K is the identity matrix, so F = E. Note that this matrix is obtained as the cross product of t and R, which is the same as converting t to a skew-symmetric matrix and multiplying it by R, so, F = $[T'_x]R$.

So, we computed the Estimated fundamental matrix with Algorithm 2 and then we compared the result with the Ground truth (GT) fundamental matrix. Then we computed the error between them, which was almost 0 (error $9.74 \cdot 10^{-15}$), meaning that we can consider both matrices equal. Thus, we can state that the algorithm correctly estimates the fundamental matrix, that in this case is $F = \begin{pmatrix} 0.098 & 0.365 & -0.189 \\ -0.365 & 0.098 & 0.567 \\ 0.036 & -0.597 & 0 \end{pmatrix}$.

Algorithm 1 NormalizePoints

Given a set of points correspondences between two images, compute the transformation matrices T and T' so that the new origin is the centroid of the points and scale it so that the mean square distance from the origin to the points is 2 pixels.

- **Inputs:** **Points**: Set of point correspondences between a pair of images
- **Outputs:** **PointsNorm**: Set of normalized point correspondences
  T: Transformation matrix

1. Convert the points to homogeneous coordinates.
2. Compute the centroid of the points
3. Compute the mean distance of the points to the origin
4. Compute the scaling factor
5. Compute the translation
6. Compute the transformation matrix
7. Apply the transformation to the points so that they become normalized

Algorithm 2 FundamentalMatix

Given a set of points correspondences p_i and p'_i between two images, estimate the fundamental matrix using the 8-points algorithm.

- **Inputs:** **Points1**: Set of points (p_i) from image 1
  Points2: Set of point (p'_i) from image 2
- **Outputs:** **PointsNorm**: Set of normalized point correspondences
  F: Estimated fundamental matrix

1. Normalize the points and obtain the transformation matrices H, H' (using Algorithm 1).
2. Create the matrix W from the p_i and p'_i point correspondences
3. Compute the Singular Value Decomposition of W : W $= UDV^T$
4. Compute the fundamental matrix F (Rank 3): F is the result of reshaping the last row of V^T into a 3x3 matrix.
5. Compute the Singular Value Decomposition of F $= UDV^T$ and remove the last singular value of D so that F has rank 2.
6. Recompute F using the modified D
7. Denormalize F using H and H'

2.2 Robust estimation of the fundamental matrix

As normally the image correspondences contain outliers, in this section we propose a robust version of the previous Algorithm 2 to estimate the fundamental matrix using RANSAC. As an example, Figure 1 shows the obtained matches using the ORB keypoint detector. Note that there are many outliers that would have a negative impact on the estimation of the fundamental matrix.

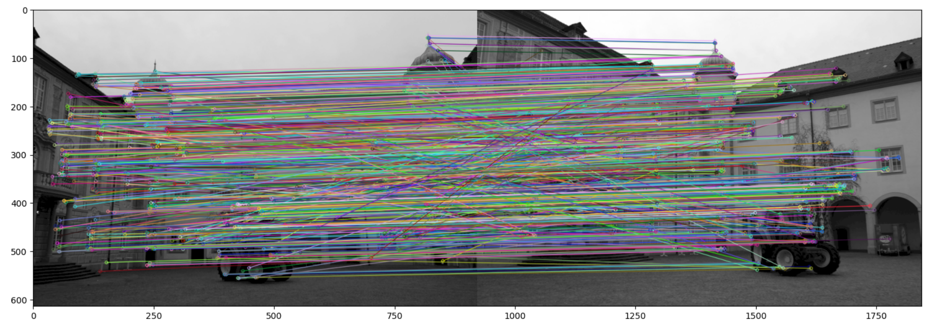


Fig. 1: Keypoint matches with outliers

Similarly as last week, we first developed an auxiliary function to select the inlier matches. However, in this case, the first order approximation of the geometric distance, called the Sampson error, which is calculated as:

$$\frac{\left(x_i'^T F x_i\right)^2}{(Fx_i)_1^2 + (Fx_i)_2^2 + (F^T x_i')_1^2 + (F^T x_i')_2^2} \tag{1}$$

The pseudo-code to obtain the inliers is shown in Algorithm 3.

Algorithm 3 Inliers

Given a set of correspondences x_i and x_i' determine which are inliers.

- **Inputs:** **F**: estimated fundamental matrix
  Points1: Set of points (x_i) from image 1
  Points2: Set of point (x_i') from image 2
  Th: threshold to determine the inliers
- **Outputs:** **Inliers**: indexes of the inliers

1. Transform x_i and x_i' with the estimated fundamental matrix F, which are needed for the divisor of the Sampson error expression.
2. Compute the Sampson error
3. Obtain the inliers: correspondences with a Sampson error $<$ threshold th.

To obtain the robust version of the fundamental matrix we use Algorithm 4, which is a modification of the RANSAC algorithm we used last week.

Algorithm 4 RansacFundamentalMatix

Robust version of the Algorithm 2 to estimate the fundamental matrix.

- **Inputs:** **Points1**: Set of points (x_i) from image 1
  Points2: Set of point (x'_i) from image 2
  Th: threshold to determine the inliers
  maxIt0: initial number of maximum iterations
- **Outputs:** **F**: Robustly estimated fundamental matrix
  Inliers: Indexes of the inliers

Initialization: Set the iterations $it = 0$
while it $<$ maxIt0 **do**

1. Select a random set of 8 correspondences
2. Estimate the fundamental matrix with the previous points using Algorithm 2
3. Get the inliers matches with Algorithm 3
4. Test if it is the best model so far (comparing the number of inliers)
5. Update the estimate of iterations $maxIt$ to ensure we pick, with probability p, an initial data set with no outliers.
6. Avoid infinite loops by setting the $maxIt0$ to the minimum between its value and the $maxIt$.
7. Increase by one the number of iterations it

end
Compute the fundamental matrix only using the best inliers found.

Figure 3 shows the inlier matches after applying RANSAC. We see that most of the mismatches where discarded, compared to Fig.1.

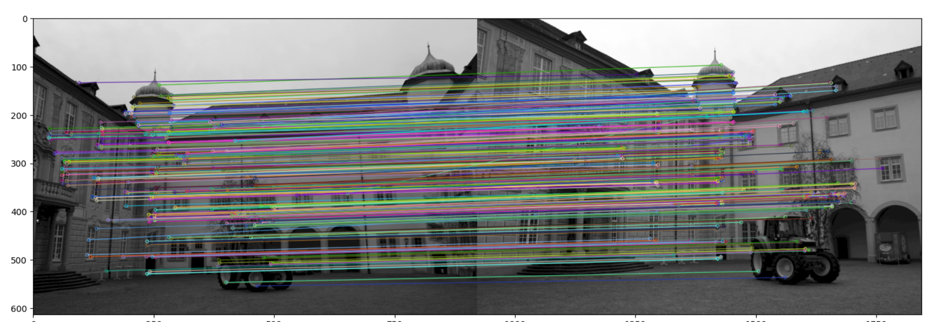
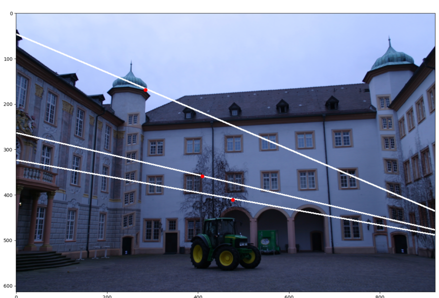


Fig. 2: Keypoint matches after applying RANSAC

2.3 Epipolar lines

In this section we calculate the correspondent epipolar lines of some given points using the robustly estimated fundamental matrix F. They are the intersection of the epipolar plane with the image plane. So, for each point in the first image, there exists a corresponding epiplolar line in the second image and vice versa. This mapping from points to lines is represented by the fundamental matrix. Thus, the epipolar lines in image 1 are calculated as $l = F^T p'_i$ while in image 2 they are calculated as $l' = Fp_i$, where the points p_i and p'_i are expressed in homogeneous coordinates.

As an example, Fig. 3 shows the epipolar lines of the points in red in images 1 and 2 respectively. Note that the lines don't exactly pass perfectly through the points. This may be due to the fact that the fundamental matrix is not perfect. Although we robustly estimated F, some outliers may still be present after RANSAC.



(a)

(b)

Fig. 3: Epipolar lines in a) Image 1 and b) Image 2

3 Application: Photo-sequencing

In this section, we implement the simplified algorithm on photo-sequencing(Dekel, Moses, and Avidan, 2013). In order to sequence a set of images we need to have an object that is dynamic, in our case, a van which we will assume that follows a straight trajectory. The three images we want to sequence are shown in Fig. 4. We will estimate the different locations of the van on the reference image using the intersection of the van's trajectory and the corresponding epipolar lines.

(a) Frame 0000

(b) Frame 0001

(c) Frame 0002

Fig. 4: Frames to photo-sequence

We first computed the keypoint matches between the reference image (0000) and the other two images, taken at different time instants. Note that the keypoint inliers were obtained using Algorithm.4 and the result is shown in Fig.5 and Fig.6.

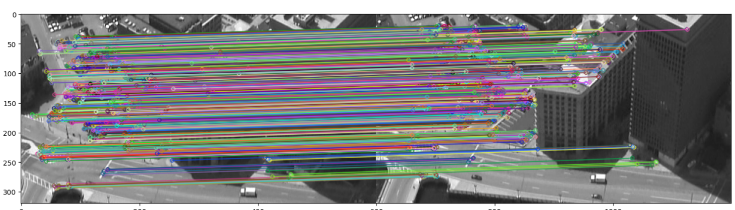


Fig. 5: Keypoints matching between 0000 and 0001 with RANSAC

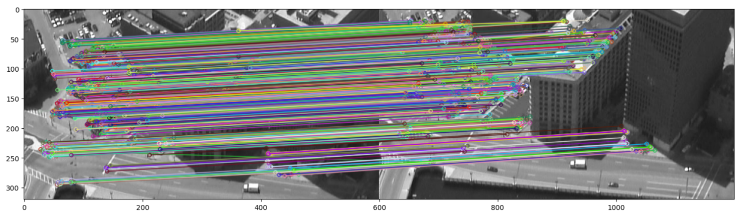


Fig. 6: Keypoints matching between 0000 and 0002 with RANSAC

Then we calculate the fundamental matrix for each pair of images using Algorithm.2. Then, selected a match of interest manually (from the initial matches, not after RANSAC), as can be seen in Fig. 7 and we calculated the trajectory of the van as the cross-product between the selected keypoint on the reference image and a manually selected point.

Fig. 7: Match of interest between image 0000 and 0001

After computing the fundamental matrix, we are interested in getting the epipolar lines, which are calculated using the formulas:

$$ep_1 = F_{12}^T * point_1$$

$$ep_2 = F_{13}^T * point_2$$

After that, the projection of the positions of the van can be calculated by finding the intersection of the found epipolar lines with the trajectory line using the cross-product. In the end, the order of the images is given by the order of the keypoints on the resulting line. 8 The cyan dot represents the position of the van in the second image (0001), while the blue point represents the position in the third image (0002).

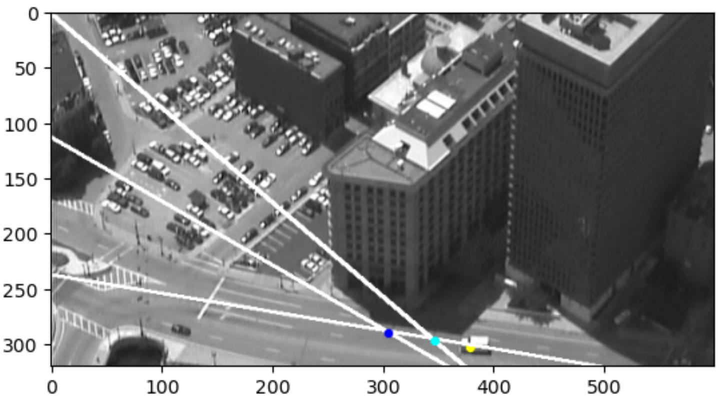


Fig. 8: Photo sequence results

4 Conclusions

In this laboratory, we have seen how to estimate the fundamental matrix and some of it's applications. We have seen that the 8-point algorithm can estimate correctly the fundamental matrix when there is no noise. However, in practice, it doesn't produce satisfactory results as it's very sensitive to outliers. In order to ameliorate this problem, we used RANSAC algorithm, which improved the accuracy of the estimations by removing most of the outliers. Nevertheless, using ORB keypoint detector we detected a lot of keypoints and sometimes was difficult to remove all the outliers with RANSAC. We also noticed the randomness of RANSAC, which sometimes removed the keypoints matches that we needed. Maybe using ORB keypoint detector would have produced better results than using ORB, as there would be fewer detections and thus, fewer outiers.

In the end, we have seen the application of photo-sequencing. We have verified that it is possible to determine the order in which where taken different pictures of a moving object (straight trajectory). Note that the pictures were taken approximately at the same position. One of the possible difficulties of this exercise occurs when calculating the key-point indices corresponding to the van for each image. Sometimes (rarely) ORB does not detect a match between these key-points of interest. In this case the key-point calculation must be rerun. Also, with respect to the photo sequence we have seen that the epipolar lines are very sensitive to small changes of the fundamental matrix. To obtain the final image of the photosequence we had to make more than one run to recalculate the fundamental matrices because the visual results were not entirely consistent.

References

Dekel, Tali, Yael Moses, and Shai Avidan (2013). "Space-Time Tradeoffs in Photo Sequencing". In: *2013 IEEE International Conference on Computer Vision*, pp. 977–984. DOI: 10.1109/ICCV.2013.125.