DIIS - I3A
Universidad de Zaragoza
C/ María de Luna num. 1
E-50018 Zaragoza
Spain

Internal Report: 2004-V01

Automatic computation of the fundamental matrix from matched lines[1]

Pellejero O.A., Sagüés C., Guerrero J.J.

If you want to cite this report, please use the following reference instead:
Automatic computation of the fundamental matrix from matched lines, Pellejero O.A., Sagüés C., Guerrero J.J., *Current Topics in Artifical Intelligence*, LNCS- LNAI 3040, pages 197-206, 2004.

[1] This work was supported by projects DPI2000-1272 and DPI2003-07986.

Automatic computation of the fundamental matrix from matched lines

Pellejero O.A., Sagüés C., Guerrero J.J.

Dpto. de Informática e Ingeniería de Sistemas (U.Z.)
María de Luna 3, E-50018 ZARAGOZA (SPAIN)
Phone 34-976-761940, Fax 34-976-761914
email: {oscarph, csagues, jguerrer}@posta.unizar.es ⋆

Abstract. This paper addresses the recovery of epipolar geometry using homographies computed automatically from matched lines between two views. We use lines because they have some advantages with respect to points, particularly in man made environments. Although the fundamental matrix cannot be directly computed from lines, it can be deduced from homographies obtained from them. We match lines lying on a plane, estimating simultaneously a planar projective transformation with a robust method. A homography allows us to select and to grow previous matches that have been obtained combining geometric and brightness image parameters. Successive homographies can be computed depending on the number of planes in the scene. From two or more, the fundamental matrix can be obtained.

1 Introduction

The recovery of epipolar geometry has been more broadly treated using points [1]. The use of lines as image features has some advantages, mainly in man made environments. Straight lines can be accurately extracted in noisy images, they capture more information than points, and they may be used where partial occlusions occur. The epipolar geometry cannot be computed directly from lines, but it can be made through homographies [2].

Line matching, which has also been previously treated [3], is more difficult than point matching [4] because the end points of the extracted lines are not reliable. Besides that, there is no geometrical constraint, like the epipolar, for lines in two images. The putative matching of features based on image parameters has many drawbacks, resulting in non matched or wrong matched features.

Perspective images of planar scenes are usual in the perception of man made environments, and how to work with them is well known. Points or lines on the world plane in one image are mapped to points or lines in the other image by a plane to plane homography, also known as a plane projective transformation [5]. This is an exact transformation for planar scenes or for small baseline image pairs. We match lines between two images computing simultaneously a planar projective transformation.

⋆ This work was supported by projects DPI2000-1265, DPI2000-1272.

So, the first homography allows us to select and to grow previous matches which have been obtained combining geometric and brightness image parameters [6]. After that, successive homographies can be computed until no more of them could be obtained. As we use least median of squares to compute the homography and several homographies are supposed, we play with some percentile to select the inliers and outliers. From at least two homographies the fundamental matrix can be directly obtained.

After this introduction, we will present the process to obtain the initial matches which will be used to compute the homography (§2). The robust estimation of homographies from lines, and the process to obtain the final matches using geometrical constraints given by the homographies, is explained in §3. After that, we present in §4 how to compute fundamental matrix once we have information about two homographies of the scene. Experimental results with real images are presented in §5. Finally, §6 is devoted to exposing the conclusions.

2 Initial matches

Lines are extracted using our implementation of the method proposed by Burns [7]. This method computes spatial brightness gradients to detect the lines in the image. Pixels having the gradient magnitude larger than a threshold are grouped into regions of similar direction of brightness gradient. These groups are named line-support regions (LSR). A least-squares fitting into the LSR is used to obtain the line. For each one, we store four geometric parameters: midpoint coordinates (x_m, y_m), the line orientation (θ) and the length of the extracted line (l). We also use two brightness attributes: agl and c (average grey level and contrast).

We determine correspondences between lines in two images without knowledge about motion or scene structure. The initial matching is made using the weighted nearest neighbor. Naming $\mathbf{r_g}$ the difference of geometric parameters between both images $(1, 2)$, $\mathbf{r_g} = [x_{m1} - x_{m2},\ y_{m1} - y_{m2},\ \theta_1 - \theta_2,\ l_1 - l_2]^T$, and $\mathbf{r_b}$ the variation of the brightness parameters between both images, $\mathbf{r_b} = [agl_1 - agl_2,\ c_1 - c_2]^T$, we can compute two Mahalanobis distances, one for geometric parameters, $\mathbf{d_g} = \mathbf{r_g}^T \mathbf{S}^{-1} \mathbf{r_g}$, and the other for brightness parameters, $\mathbf{d_b} = \mathbf{r_b}^T \mathbf{B}^{-1} \mathbf{r_b}$. To establish the matches, we test the geometric and the brightness compatibility. A line in the first image can have more than one compatible line in the second image. From the compatible lines, the line having the smallest $\mathbf{d_g}$ is selected as putative match. Details are in [6].

3 From lines to homographies

The representation of a line in the projective plane is obtained from the analytic representation of a plane through the origin: $n_1 x_1 + n_2 x_2 + n_3 x_3 = 0$. The equation coefficients $\mathbf{n} = (n_1, n_2, n_3)^T$ correspond to the homogeneous coordinates of the projective line. All the lines written as $\lambda \mathbf{n}$ are the same as $\mathbf{n}$. As cameras have a limited field of view, observed lines have usually n_3 close to 0. Similarly, an image point $\mathbf{p} = (x_1, x_2, 1)^T$ is also an element of the projective plane. A projective transformation between two projective planes (1 and 2) can be represented by a linear transformation $\mathbf{H}_{21}$, in such a way that $\mathbf{p}_2 = \mathbf{H}_{21}\mathbf{p}_1$. Considering the above equations for lines in both images, we have $\mathbf{n}_2 = \left[\mathbf{H}_{21}^{-1}\right]^T \mathbf{n}_1$. A homography requires eight parameters to be completely defined, because there is an overall scale factor. A corresponding point or line gives two linear equations in terms of the elements of the homography. Thus, four corresponding lines assure a unique solution for $\mathbf{H}_{21}$, unless three of them are parallel or intersect in the same point. To have an accurate solution it is interesting to have the lines as separate as possible in the image.

3.1 Computing homographies from corresponding lines

Here, we will obtain the projective transformation of points ($\mathbf{p}_2 = \mathbf{H}_{21}\mathbf{p}_1$), but using matched lines. To deduce it, we suppose the start (s) and end (e) tips of a matched line segment to be $\mathbf{p}_{s1}, \mathbf{p}_{e1}, \mathbf{p}_{s2}, \mathbf{p}_{e2}$, which will not usually be corresponding points. The line in the second image can be computed as the cross product of two of its points (in particular the observed tips) as

$$\mathbf{n}_2 = \mathbf{p}_{s2} \times \mathbf{p}_{e2} = [\mathbf{p}_{s2}]_\times \mathbf{p}_{e2}, \tag{1}$$

where $[\mathbf{p}_{s2}]_\times$ is the skew-symmetric matrix obtained from vector $\mathbf{p}_{s2}$.

As the tips belong to the line we have, $\mathbf{p}_{s2}^T\mathbf{n}_2 = 0$; $\mathbf{p}_{e2}^T\mathbf{n}_2 = 0$. As the tips of the line in the first image once transformed also belong to the corresponding line in the second image, we can write, $\mathbf{p}_{s1}^T\mathbf{H}_{21}^T\mathbf{n}_2 = 0$; $\mathbf{p}_{e1}^T\mathbf{H}_{21}^T\mathbf{n}_2 = 0$. Combining with equation (1) we have,

$$\mathbf{p}_{s1}^T\mathbf{H}_{21}^T[\mathbf{p}_{s2}]_\times\mathbf{p}_{e2} = 0\ ;\ \mathbf{p}_{e1}^T\mathbf{H}_{21}^T[\mathbf{p}_{s2}]_\times\mathbf{p}_{e2} = 0. \tag{2}$$

Therefore each couple of corresponding lines gives two homogeneous equations to compute the projective transformation, which can be determined up to a non-zero scale factor. Developing them according to the elements of the projective transformation, we have

$$\begin{pmatrix} Ax_{s1} & Ay_{s1} & A & Bx_{s1} & By_{s1} & B & Cx_{s1} & Cy_{s1} & C \\ Ax_{e1} & Ay_{e1} & A & Bx_{e1} & By_{e1} & B & Cx_{e1} & Cy_{e1} & C \end{pmatrix} \mathbf{h} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

where $\mathbf{h} = (h_{11}\, h_{12}\, h_{13}\, h_{21}\, h_{22}\, h_{23}\, h_{31}\, h_{32}\, h_{33})^T$ is a vector with the elements of $\mathbf{H}_{21}$, and $A = y_{s2} - y_{e2}$, $B = x_{e2} - x_{s2}$ and $C = x_{s2}y_{e2} - x_{e2}y_{s2}$, being (x_{s1}, y_{s1}) the coordinates of the start tip $\mathbf{p}_{s1}$.

Using four corresponding lines, we can construct a 8×9 matrix $\mathbf{M}$. In order to have a reliable transformation, more than the minimum number of matches and an estimation method may be considered. Thus from n matches a $2n \times 9$ matrix $\mathbf{M}$ can be built, and the solution $\mathbf{h}$ can be obtained from SVD decomposition of this matrix [5]. In this case the relevance of each line depends on its observed length, because the cross product of the segment tips is related to the segment length.

It is known that a previous normalization of data avoids problems of numerical computation. As our formulation only uses image coordinates of observed tips, data normalization proposed for points [8] has been used.

3.2 Robust estimation

The least squares method assumes that all the measures can be interpreted with the same model, which makes it very sensitive to out of norm data. Robust estimation tries to avoid the outliers in the computation of the estimate. From the existing robust estimation methods [9], we have chosen the least median of squares method. This method makes a search in the space of solutions obtained from subsets of minimum number of matches. The algorithm to obtain an estimate with this method can be summarized as follows:

1. A Monte-Carlo technique is used to randomly select m subsets of 4 features.
2. For each subset S, we compute a solution in closed form $\mathbf{H}_S$.
3. For each solution $\mathbf{H}_S$, the median or other percentile M_S of the squares of the residue with respect to all the matches is computed.
4. We store the solution $\mathbf{H}_S$ which gives the minimum percentile M_S.

A selection of m subsets is good if at least in one subset the 4 matches are good. Assuming a ratio ϵ of outliers, the probability of one of them being good can be obtained [10] as, $P = 1 - \left[1 - (1-\epsilon)^4\right]^m$.

Once the solution has been obtained, the outliers can be selected from those of maximum residue. As in [9] the threshold is fitted proportionally to the standard deviation of the residue, estimated as [10], $\hat{\sigma} = 1.4826\,[1 + 5/(n-4)]\,\sqrt{M_S}$. Assuming that the measurement error is Gaussian with zero mean and standard deviation σ, then the square of the residues follows a χ^2_2 distribution with 2 degrees of freedom. Taking, for example, that a 95% probability is established for the line to fit in the homography (inlier) then the threshold will be fixed to $5.99\,\hat{\sigma}^2$.

3.3 Growing matches from homography

From here on, we introduce the geometrical constraint introduced by the estimated homography to get a bigger set of matches. Thus final matches consist of two sets. The first one is obtained from the initial set of matches selected after the robust computation of the homography that passes an overlapping test compatible with the transformation of the segment tips additionally. The second set of matches is obtained using all the segments not matched initially and those previously rejected. With this set of lines a matching process similar to the basic matching is carried out. However, now the matching is made with the nearest neighbor segment transformed with the homography. The transformation is applied to the end tips of the image segments using the homography $\mathbf{H}_{21}$ to find, not only compatible lines but also compatible segments in the same line. In the first stage of the matching process, there was no previous knowing of camera motion. However, in this second step the computed homography provides information about an expected disparity and therefore the uncertainty of geometric variations can be reduced.

3.4 Several homographies

Previously, we have explained how to determine a homography from lines, assuming some of the lines used are outliers. The first homography can be computed in this way, assuming that a certain percentage of matched lines between images are good matches of lines in the plane to be extracted, and the others are outliers. We have not got a priori knowledge about which plane of the scene is going to be extracted first, but the probability of being chosen increases with the number of lines on it. If we execute the process two times, the same plane will be extracted with high probability, unless the number of lines in two main planes of the scene are similar. So, the only reasonable way to extract a second plane of the scene consists of eliminating lines used to determine first homography, expecting the second plane to have the largest number of lines then. Here we are assuming that lines belonging to the first plane do not belong to the second plane. That is true except for the intersection line between both planes.

So, once we have computed a first homography we should eliminate all lines belonging to that plane, that is to say, the lines which verify the homography. Probably it is better to eliminate all lines belonging to the region where the plane has been extracted, but at the moment it is not easy to determine the limits of the region, and other planes could exist inside.

4 Fundamental matrix from homographies

Fundamental matrix is a 3×3 matrix of rank 2 which encapsulates the epipolar geometry. It only depends on cameras' internal parameters and on relative motion.

As the images are obtained with the same camera whose projection matrix in a common reference system are $\mathbf{P}_1 = \mathbf{K}[\mathbf{I}|\mathbf{0}]$, $\mathbf{P}_2 = \mathbf{K}[\mathbf{R}|\mathbf{t}]$ (being $\mathbf{R}$ the camera rotation, $\mathbf{t}$ the translation and $\mathbf{K}$ the internal calibration matrix), then, the fundamental matrix can be expressed as $\mathbf{F}_{21} = \mathbf{K}^{-T}\ ([\mathbf{t}]_\times\,\mathbf{R})\ \mathbf{K}^{-1}$. It can be computed from corresponding points [1], in such a way that epipolar constraint for points in both images can be expressed as $\mathbf{p}_2^T\,\mathbf{F}_{21}\,\mathbf{p}_1 = 0$.

It can also be computed from homographies obtained through two or more planes. In this way when there are planar structure, corresponding lines in two images can be used. If at least two homographies $(\mathbf{H}_{21}^{\pi_1}, \mathbf{H}_{21}^{\pi_2})$ can be computed between both images corresponding to two planes (π_1, π_2), a homology $\mathbf{H} = \mathbf{H}_{21}^{\pi_1} \cdot (\mathbf{H}_{21}^{\pi_2})^{-1}$, that is a mapping from the second image into itself, exists. Under this mapping the epipole in second image $(\mathbf{e}_2)$ is a fixed point and therefore $\mathbf{e}_2 = \mathbf{H}\,\mathbf{e}_2$. So $(\mathbf{e}_2)$ may be determined by the eigenvector of $\mathbf{H}$ corresponding to unary eigenvalue [5]. From two planes, the fundamental matrix can be either computed as $\mathbf{F}_{21} = [\mathbf{e}_2]_\times\,\mathbf{H}_{21}^{\pi_1}$ or $\mathbf{F}_{21} = [\mathbf{e}_2]_\times\,\mathbf{H}_{21}^{\pi_2}$.

Our algorithm consists of an iteration of robust estimation of a homography and elimination of lines verifying it. We can iterate while planes are extracted. Then, if we have extracted two or more, we can compute the fundamental matrix as we have described previously. If we have estimated a first homography, any new homography will not be good for fundamental matrix estimation. A fundamental matrix computed from two close planes is inaccurate. This can be avoided checking the condition number of the homology $\mathbf{H} = \mathbf{H}_{21}^{\pi_1} \cdot (\mathbf{H}_{21}^{\pi_2})^{-1}$. If $\mathbf{H}_{21}^{\pi_1}$ and $\mathbf{H}_{21}^{\pi_2}$ are two homographies of the same plane, the condition number of the homology $\mathbf{H}$ would be close to 1. We demand a greater value than a threshold. However, a large value is not suitable, because large condition numbers indicate a nearly singular matrix. So the condition number should fit into a fixed range. In this way, if second homography leads us to a homology with a condition number out of the fixed range, it is discarded and lines belonging to that plane are also eliminated, in order to let us get a new homography with the rest.

Moreover, knowledge of one homography of the scene restricts others. Linear subspace constraints on homographies have been previously derived in [11]. They showed that the collection of homographies of multiple planes between a pair of views span a 4-dimensional linear subspace. This constraint, however, requires the number of planes in the scene to be greater than 4. Zelnik-Manor and Irani apply it [12], replacing the need for multiple planes with the need for multiple views. They only need two planes, and often a single one (under restricted assumptions on camera motion). We work only with two images, and more than four planes in a scene are not always available, but easier constraints are possible. In fact, the knowledge of first homography allows us to compute the second only from three matched lines, and knowledge of first and second let us compute the third from two.

This method may fail if only one plane in the scene exists, or in case of pure rotation, because epipolar geometry is not defined and only one homography can be computed, but an automatic detection of this situation is made through condition number of the homology.

5 Experimental Results

The proposed method works properly with indoor images and outdoor images provided that at least two planes, containing enough lines, exist in the scene.

To measure the goodness of the computed fundamental matrix we use the first order geometric error computed as the Sampson distance [5] for points extracted and matched manually,

$$\sum_i \frac{(\mathbf{p}_2^T\, \mathbf{F}_{21}\, \mathbf{p}_1)^2}{(\mathbf{F}_{21}\, \mathbf{p}_1)_f^2 \,+\, (\mathbf{F}_{21}\, \mathbf{p}_1)_s^2 \,+\, (\mathbf{F}_{21}^T\, \mathbf{p}_2)_f^2 \,+\, (\mathbf{F}_{21}^T\, \mathbf{p}_2)_s^2} \tag{3}$$

here $()_f$ and $()_s$ indicate the first and second components of the corresponding vectors.

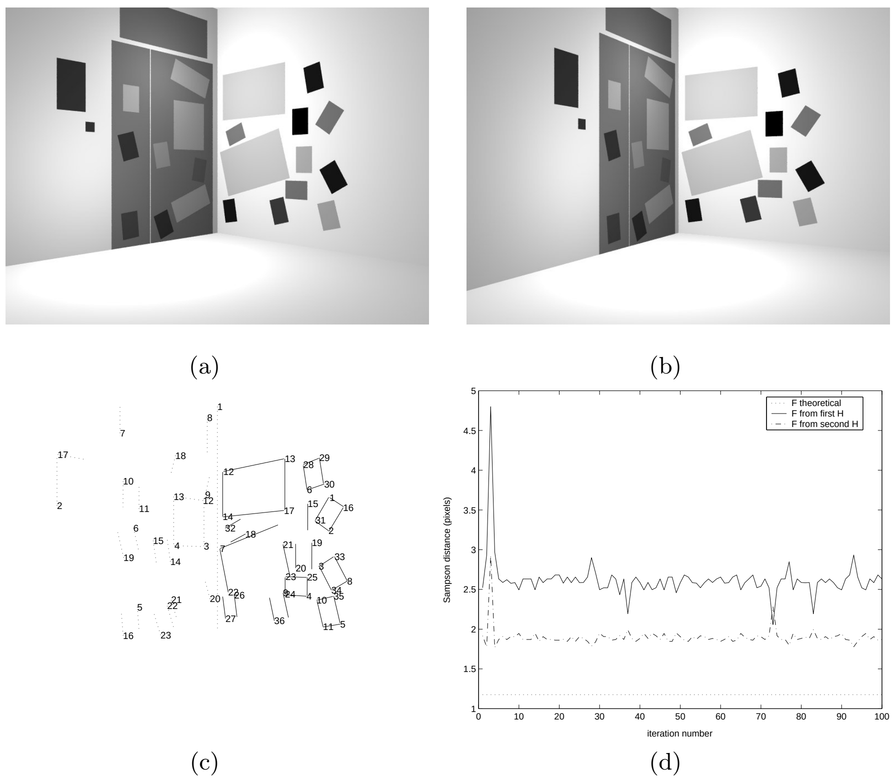

Fig. 1. (a) and (b) Synthetic images to compute the fundamental matrix. (c) Matches belonging to each plane in one execution. We show only one image because all matches are correct. (d) Sampson distance, with the two fundamental matrices computed from homology for 100 iterations and theoretical fundamental matrix. Sampson distance for theoretical matrix gives us a measure of quality of manually selected points. Notice that the average distance for fundamental matrices computed is only about a pixel over theoretical matrix.

The square root resulting on dividing such value between the number of points used, gives us a measure in pixels comparable with other pairs of images. One pixel of error could have been introduced in the manual selection of point matches. Since point matches are different for every image pair, measure noise is also different, and comparisons should be done with care.

The first experiment is made with synthetic images (Fig. 1). The motion of the camera includes translation and rotation. The first image camera centre is on an equidistant line to the planes showed, four meters from each one. The motion is a rotation of 15 degrees using as axis the intersection line of both planes. We can compute theoretically the epipole and the fundamental matrix knowing the camera motion. The theoretical epipole can be compared with the obtained epipole, computing the angle of both epipoles in relation to the camera

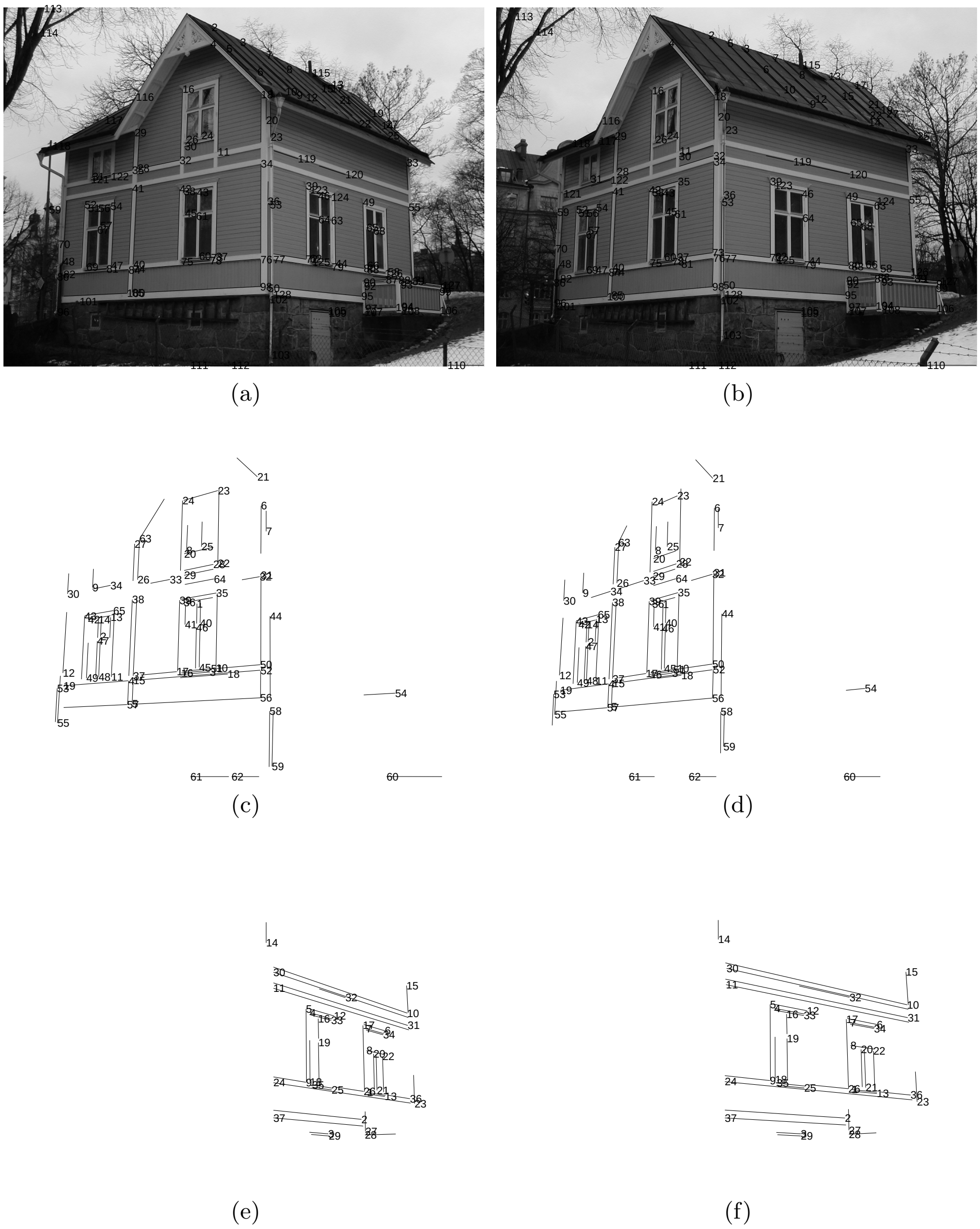

Fig. 2. Real images to compute the fundamental matrix. (a) and (b) Initial matching of straight lines. (c) and (d) Matches in the first homography. (e) and (f) Matches in the second homography. In both extracted planes we obtain new matches, whereas outliers are reduced significantly. (Images supplied by D. Tell, KTH Stockholm).

center. We have done 100 executions forcing the condition number over 1.4. The mean value of this error is 1.7628 degrees with a standard deviation of 0.6037. We want to emphasize that, in this case and for a standard execution, all final matches of lines between images are correct and are considered as belonging to the correct plane.

We also show an experiment carried out with real outdoor images. Initial matches of lines and final matches obtained computing two homographies in one execution can be seen in Fig. 2. Outliers are common in initial matching, whereas in final matching they are rare. Moreover, lines not considered initial matches, appear after estimating the homography. However, lines not belonging to the region of the plane in the image also appear in this phase, because they are close to satisfying the constraint imposed by it, because they are nearly collinear with corresponding epipolar line. We think they do not spoil neither the computed homography, nor therefore the fundamental matrix estimation. We have executed the algorithm 200 times, ordering executions by condition number. In this experiment the condition number has not been limited in any way, in order to analyze its influence on the quality of fundamental matrix. In Fig. 3, we show the Sampson distance and the condition number obtained in each execution. We can notice that it exists an interval (from execution 125 to 185 approximately), where Sampson distance takes low values. That is to say, we can use condition number to determine when an execution is good. In this experiment, an execution with condition number between 1.53 and 1.64 can be considered good. Out of this range, fundamental matrix should be discarded, and a new execution done.

6 Conclusions and future work

We have presented a method to compute automatically the fundamental matrix between two images using lines.

The proposed method works properly with indoor images and outdoor images provided that at least two planes, containing enough lines, exist in the scene. The condition number of the obtained homology can be used to decide the goodness of the fundamental matrix.

However, matches obtained not always belong to the corresponding region of the plane extracted. Lines used for the computation of one homography are discarded for the following homographies, but we should check that regions containing such lines are not overlapping. Another improvement consists of constraints between homographies. Knowledge of one homography of the scene restricts others. The use of regions associated to planes and constraints between homographies will lead us to a faster and more accurate algorithm.

References

1. Zhang, Z.: Determining the epipolar geometry and its uncertainty: A review. International Journal of Computer Vision **27** (1998) 161–198

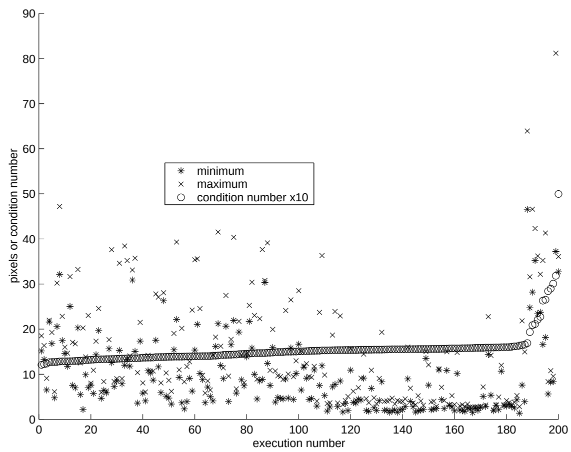

Fig. 3. Sampson distance per point for 200 executions. In each one two fundamental matrices are obtained. We show the minimum and the maximum of both Sampson distance, and the condition number of the homology obtained in the execution multiplied by 10. Executions are ordered by condition number. We can notice that it exists an interval of values from 1.53 to 1.64 (executions 125 to 185 approximately) where Sampson distance takes smaller values. These values can be used to decide the goodness of an execution.

2. Rother, C., Carlsson, S.: Linear multi view reconstruction and camera recovery using a reference plane. International Journal of Computer Vision **49** (2002) 214–223
3. Schmid, C., Zisserman, A.: Automatic line maching across views. In: IEEE Conference on CVPR. (1997) 666–671
4. Zhang, Z., Deriche, R., Faugeras, O., Luong, Q.: A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence (1995) 87–119
5. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge (2000)
6. Guerrero, J., Sagüés, C.: Robust line matching and estimate of homographies simultaneously. IbPRIA. Pattern Recognition and Image Analysis, LNCS 2652. (2003) 297–307
7. Burns, J., Hanson, A., Riseman, E.: Extracting straight lines. IEEE Trans. on Pattern Analysis and Machine Intelligence **8** (1986) 425–455
8. Hartley, R.: In defense of the eight-point algorithm. IEEE Trans. on Pattern Analysis and Machine Intelligence **19** (1997) 580–593
9. Zhang, Z.: Parameter estimation techniques: A tutorial with application to conic fitting. Rapport de recherche RR-2676, I.N.R.I.A., Sophia-Antipolis, France (1995)
10. Rousseeuw, P., Leroy, A.: Robust Regression and Outlier Detection. John Wiley, New York (1987)
11. Shashua, A., Avidan, S.: The rank 4 constraint in multiple (≥ 3) view geometry. In: Proc. European Conf. Computer Vision, St. Louis, Missouri (1996) 196–206
12. Zelnik-Manor, L., Irani, M.: Multiview constraints on homographies. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 214–223