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

Internal Report: 2001-V04

From Lines to Homographies between Uncalibrated Images[1]

Guerrero J.J., Sagüés C.

If you want to cite this report, please use the following reference instead:
From Lines to Homographies between Uncalibrated Images, Guerrero J.J., Sagüés C., *IX Spanish Symposium on Pattern Recognition and Image Analysis*, pages 233-240, 2001.

[1]This work was supported by projects CICYT TAP97-0992-C02-01 and DGA P29/98.

From Lines to Homographies between Uncalibrated Images

Guerrero J.J. and Sagüés C.
Dpto. de Informática e Ingeniería de Sistemas, Univ. de Zaragoza
María de Luna 3, E-50015 ZARAGOZA (SPAIN)
Phone 34-976-761940, Fax 34-976-761914
email {jguerrer, csagues}@posta.unizar.es †

Abstract

A robust method to compute homographies using lines in two uncalibrated images is presented and tested. Line features can usually be extracted and matched more accurately than points and they can be used in cases where there are partial occlusions. Lines and points are dual features to compute homographies, but some particular problems related to the representation and the normalization of data, avoiding singularities, must be considered in practice. Some experiments related to the robust selection of matched lines belonging to a plane have been used to test the presented method. The results show that the robust technique turns out stable and useful, and it opens interesting uncalibrated applications particularly in robot perception and navigation.

1 Introduction

In this paper a robust estimation method is used to compute homographies using corresponding lines in two uncalibrated images. Perspective images of planar scenes are usual in perception of man made environments, and the camera model to work with them is well known. Points or lines on the world plane or in one image of the world plane are mapped to points or lines in the other image by a plane to plane homography, also known as a plane projective transformation. A homography is represented by a 3x3 matrix, which allows computing geometrical measurement of the world, directly in the image [1].

Using lines instead of points has been considered by many researches. Straight lines can be extracted more accurately than points in noisy images, they are also easier to match than the latter, and may be used in cases where occlusions occur.

†This work was supported by projects CICYT TAP97-0992-C02-01 and DGA P29/98.

To compute homographies, points and lines are dual geometric entities, however line-based algorithms are generally more difficult than point-based ones [2]. Thus, some particular problems related to data representation and normalization must be considered in practice, because singularities or ill conditioning may occur. In this paper, we present a formulation to compute homographies from corresponding lines in two images that makes use of classical normalization of point data [3], and that avoids singularities.

On the other hand, robust estimators are actually unquestionable techniques to obtain results in real situation where outliers and spurious data are present, and many researches make use of them [4]. In this paper the least median of squares method [5] has been tested and used. It provides as output not only the solution in a robust way, but also a list of data that are in disagreement with it, which allows to reject bad matches.

Two main motivations underlie our work. Firstly we search a robust method to improve matching of lines [6]. Secondly, we try to segment planar regions of the scene using two uncalibrated images, which can be very useful for 3D perception and robot navigation in man made environments [7].

2 Points, lines and homographies

A point in the projective plane is represented by three Cartesian coordinates, $\mathbf{p} = (x_1, x_2, x_3)^T$, which represents a ray through the origin in the 3D space [8]. Only the direction of the ray is relevant, so all points written as $\lambda\mathbf{p} = (\lambda x_1, \lambda x_2, \lambda x_3)^T$ are equivalent. The classical Cartesian coordinates of the point (x, y) can be obtained intersecting the ray with a special plane perpendicular to x_3 axis and located at unit distance along x_3. This is equivalent to scale $\mathbf{p}$ as, $\mathbf{p} = (x, y, 1)^T$. Projected points in an image are represented in this way, considering this special plane as the image plane. As normal cameras have a limited field of view, the special case $x_3 = 0$, which corresponds with a ray parallel to the image plane, cannot be observed, and therefore observed points have always $x_3 \neq 0$.

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. Again $\lambda\mathbf{n}$ is the same line than $\mathbf{n}$. The case $n_3 = 0$ corresponds to a line through the origin of the virtual image plane. As cameras have a limited field of view, observed lines have usually n_3 close to 0.

The projective equation of a line can be represented in different ways because $\mathbf{n}$ and $\mathbf{p}$ may be interchanged without affecting the equation, which shows the duality of points and lines. Thus, $\mathbf{n} \cdot \mathbf{p} = \mathbf{n}^T \mathbf{p} = \mathbf{p}^T \mathbf{n} = 0$.

A projective transformation between two projective planes (1 and 2) can be represented by a linear transformation $\mathbf{p}_2 = \mathbf{T}_{21}\mathbf{p}_1$. If the transformation is represented in Cartesian coordinates it results non-linear.

Since points and lines are dual in the projective plane, the transformation of the line coordinates is also linear. Considering the equation previously presented, $\mathbf{n}_1^T \mathbf{p}_1 = 0$ and if a point $\mathbf{p}_1$, is transformed as $\mathbf{p}_2 = \mathbf{T}_{21}\mathbf{p}_1$ then, $\mathbf{n}_1^T \mathbf{T}_{21}^{-1} \mathbf{p}_2 = 0$.

Besides, as the transformed line equation is $\mathbf{n}_2^T \mathbf{p}_2 = 0$, the transformed line coordinates are, $\mathbf{n}_2 = \left[\mathbf{T}_{21}^{-1}\right]^T \mathbf{n}_1$. Therefore, lines and points are similarly transformed but the transformation matrix of lines is the transpose of the inverse of the matrix defining the point transformation [8].

A homography or a plane projective transformation 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 matrix. The most common and simplest method to compute an homography supposes an element of the matrix to be no zero, solving for the other eight elements. In principle four corresponding points or lines assure a unique solution for $\mathbf{T}_{12}$, if no three of the image points are collinear, or if no three of the image lines are parallel. To have an accurate solution it is interesting to have the features as separate in the image as possible.

3 Computing Homographies from Lines

Using corresponding points, two equations for each i match are usually written as, $(\lambda_i x_{2i}, \lambda_i y_{2i}, \lambda_i)^T = \mathbf{T}_{21}(x_{1i}, y_{1i}, 1)^T$ [8]. As points and lines are dual, this equation could be used to compute the projective transformation from lines. However, using lines it is not suitable to put $\mathbf{n}$ as $\mathbf{n} = (n_x, n_y, 1)$, because observed lines with real images are close to the singularity ($n_z = 0$).

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

$$\mathbf{n}_2 = \mathbf{p}_{s2} \times \mathbf{p}_{e2} = \tilde{\mathbf{p}}_{s2}\mathbf{p}_{e2} \tag{1}$$

where $\tilde{\mathbf{p}}_{s2}$ is the skew-symmetric matrix,

$$\tilde{\mathbf{p}}_{s2} = \begin{pmatrix} 0 & -1 & y_{s2} \\ 1 & 0 & -x_{s2} \\ -y_{s2} & x_{s2} & 0 \end{pmatrix}$$

As the tips belong to the line we have, $\mathbf{p}_{s2}^T \mathbf{n}_2 = 0$; $\mathbf{p}_{e2}^T \mathbf{n}_2 = 0$. And as the transformed tips also belong to the corresponding line, we can write, $\mathbf{p}_{s1}^T \mathbf{T}_{21}^T \mathbf{n}_2 = 0$; $\mathbf{p}_{e1}^T \mathbf{T}_{21}^T \mathbf{n}_2 = 0$. With the equation (1) we have,

$$\mathbf{p}_{s1}^T \mathbf{T}_{21}^T \tilde{\mathbf{p}}_{s2} \mathbf{p}_{e2} = 0 \; ; \; \mathbf{p}_{e1}^T \mathbf{T}_{21}^T \tilde{\mathbf{p}}_{s2} \mathbf{p}_{e2} = 0 \tag{2}$$

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

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

where $\mathbf{t} = (t_{11}\, t_{12}\, t_{13}\, t_{21}\, t_{22}\, t_{23}\, t_{31}\, t_{32}\, t_{33})^T$ is a vector with the elements of the homography matrix, and $A = y_{s2} - y_{e2}$, $B = x_{e2} - x_{s2}$ and $C = x_{s2} y_{e2} - x_{e2} y_{s2}$.

Using four corresponding lines, we can construct a 8x9 matrix $\mathbf{M}$. Then, the solution corresponds with the eigenvector associated to the least eigenvalue (in this case the null eigenvalue) of the matrix $\mathbf{M}^T\, \mathbf{M}$.

In order to have a reliable transformation, more than the minimum number of matches and an estimation method must be considered. The previous formulation is valid using as residue the algebraic distance corresponding to these equations. In this case the residue will be related to the distance from the tips of the observed segment to the corresponding transformed line, and 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 is suitable to avoid numerical computation problems. As our formulation only uses image coordinates of observed tips of lines, the data normalization proposed for points [3] has been directly used.

4 Robust estimation

The least squares method assumes that all the measures can be interpreted with the same model. This makes the method to be very sensitive to out of norm data. Robust estimation tries to avoid 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. If we need a minimum of 4 matches to compute the projective transformation, and there are a total of n matches, then the search space will be obtained from the combinations of n elements taken 4 by 4. As it is too big, several subsets of 4 matches are randomly chosen. 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 $\mathbf{T}_S$.
3. For each solution $\mathbf{T}_S$, the median M_S of the squares of the residue with respect to all the matches is computed.
4. We store the estimation of parameters $\mathbf{T}_S$ which gives the minimum M_S.

A selection of m subsets is good if at least one of them has no outliers. Assuming a ratio ϵ of outliers, the probability of one of them been good can be obtained as [5], $P = 1 - \left[1 - (1-\epsilon)^4\right]^m$. For example, if we want a probability $P = 0.999$ of one of them being good, having $\epsilon = 35\%$ of outliers, the number of subsets m should be 34.

Once the solution has been obtained, the outliers can be selected from those of maximum residue. Good matches can be selected between those of residue smaller than a threshold. In [9] the threshold is fitted proportional to the standard deviation of the residue, which is estimated as $\hat{\sigma} = 1.48\,[1 + 5/(n-4)]\,\sqrt{M_S}$.

When the bad matches have been rejected, a better solution can be obtained with the selected matches by the least squares method, as explained above.

5 Experimental Results

The method presented could be used to improve classical matching techniques [6], allowing also to compute a geometric relation between images. From it, many applications in robot navigation or scene reconstruction could be achieved.

In order to test the capacity of selection and accuracy of the robust method, we firstly have carried out experiments to select the lines belonging to a planar surface. These experiments have shown the influence of the arguments of the algorithm and have proved the goodness of the robust technique to eliminate outliers.

Figure 1: Results of the selection of lines belonging to a plane. In the figure we show both images and the matched lines selected by the robust method.

In the first experiment we show the results of the computation of the projective transformation from a set of lines and the segmentation of planes. To carry out it, we have used two images where the main object is the plane cover of a box (Fig. 1).

We have manually selected 35 good matches (27 of them (77%) correspond to the box), and some spurious matches which always have been automatically discarded by the robust method. The 27 matches belonging to the box have been perfectly selected, 6 have been rejected as not belonging to the plane. Besides that, 2 matches (those of the upper right corner) have been also selected although they do not belong to the box. This is because the epipolar plane is nearly the same as the projection plane of both lines and therefore any associated depth is compatible with the observations. In figure 1 we can see the results of the segmentation.

We have made also a Monte-Carlo experiment repeating for one hundred runs the computation of the homography. In figure 2 we show the mean absolute error for the 27 lines belonging to the box regardless of whether they have been selected or not. Taking 34 sets of 4 lines the error is about 1 pixel, but in 4 runs the error is higher (6 pixels) and then the homography is badly computed. However with 72 sets, which represents a probability of $99,9\%$ of taking a good set of 4 lines, when there exist 45% of outliers, the mean residue is in every run less than 1.1 pixels.

We present also a second experiment to show the influence of the threshold of outliers detection (it is function of the standard deviation of residue) when segmenting images (Fig. 3). In this case there is rotation and translation between images. In the experiments, 22 matches belong to the door and 11 to the wall. We have

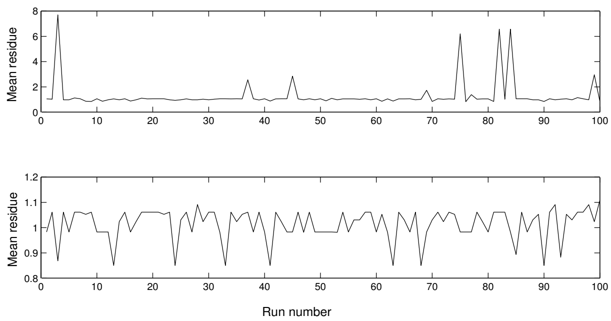

Figure 2: One hundred runs taking 34 (top) and 72 (bottom) initial sets of 4 lines respectively. We show the mean residue obtained with the lines belonging to the box, regardless of whether they have been selected or not.

a) b) c)

Figure 3: a)b) Images used with the initial good matches c) Selected lines in the plane of the door

chosen 72 sets of 4 lines to compute the homography.

The result is quite good (Fig. 3) because all the matches not belonging to the door are discarded except one that is nearly parallel to the epipolar line. We have also carried out one hundred runs to test the influence of the threshold from which the lines are selected to compute the homography, using $2\hat{\sigma}$ and $\hat{\sigma}/2$ as threshold.

The mean error for the lines in the plane of the door is lower when a smaller threshold is used, showing that the transformation obtained in second case ($\hat{\sigma}/2$) is better. Therefore we conclude that it is better to compute the projective transformation from few but good matches. However, in this case ($\hat{\sigma}/2$), as the threshold is more demanding, some short lines belonging to the door have been also rejected.

6 Conclusions

We have presented and tested a robust method to compute homographies using lines in two images. This provides a mapping between uncalibrated images, which is specially useful for close images or for images of planar surfaces. The experiments show that the robust technique turns out useful to select good matches and to compute image transformations, even in presence of outliers. We are now applying these ideas in robust matching and automatic correction of trajectories for robot navigation in man made environments.

References

[1] A. Criminisi, I. Reid, and A. Zisserman. A plane measuring device. In *Proc.1996 British Machine Vision Conference*, 1996.

[2] L. Quan and T. Kanade. Affine structure from line correspondences with uncalibrated affine cameras. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(8):834–845, 1997.

[3] R.I. Hartley. In defense of the eight-point algorithm. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(6):580–593, 1997.

[4] P.H.S. Torr and D. Murray. The development and comparison of robust methods for estimating the fundamental matrix. *International Journal of Computer Vision*, 24(3):271–300, 1997.

[5] P.J. Rousseeuw and A.M. Leroy. *Robust Regression and Outlier Detection.* John Wiley, New York, 1987.

[6] J.J. Guerrero and C. Sagüés. Tracking features with camera maneuvering for vision-based navigation. *Journal of Robotic Systems*, 15(4):191–206, 1998.

[7] C. Sagüés and J.J. Guerrero. Motion and structure for vision-based navigation. *Robotica*, 17(4):355–364, 1999.

[8] J.L.S. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision.* MIT Press, Boston, 1992.

[9] Z. Zhang. Parameter estimation techniques: A tutorial with application to conic fitting. Rapport de recherche RR-2676, I.N.R.I.A., Sophia-Antipolis, France, 1995.