COMPARATIVE STUDY OF DIFFERENT APPROACHES FOR EFFICIENT RECTIFICATION UNDER GENERAL MOTION

Mr.V.SRINIVASA RAO[1] **Prof.A.SATYA KALYAN[2]**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
PRASAD V POTLURI SIDDHARTHA INSTITUTE OF TECHNOLOGY
KANURU, VIJAYAWADA

Abstract: This paper is concerned with the generating an efficient fundamental matrix for image rectification problem under motion. Computation of fundamental matrix is the information source to estimate the accurate optic flow in computer vision area. This includes extraction of images from real world, using 8-point algorithm, and calibration for predicting optic flow. We are performing a comparative study over different fundamental matrix, and implementation techniques to propose a method which reduces the complexity in analysis of an image rectification and improves computational speed for vehicle head lamp tracking. Fundamental matrix estimation is difficult since it is often based correspondences that are spoilt by noise and outliers. Outliers must be thrown out via robust statistics, and noise gives uncertainty. In this article we provide a closed-form formula for the uncertainty of the so-called 8 point algorithm, which is a basic tool for fundamental matrix estimation via robust methods. As an application, we modify a well established robust algorithm accordingly, leading to a new criterion to recover point correspondences under epipolar constraint, balanced by the uncertainty of the estimation.

Keywords: Image rectification; Image segmentation; vehicle velocity; Optical flow.

1. Introduction

The stereo matching problem can be solved much more efficiently if images are rectified. This step consists of transforming the images so that the epipolar lines are aligned horizontally. In this case stereo matching algorithms can easily take advantage of the epipolar constraint and reduce the search space to one dimension (i.e. corresponding rows of the rectified images).The ceaseless advances in camera technology and areas such as computer vision have lead to the development of automatic methods used in traffic surveillance. To overcome this difficulty, one may create and employ virtual images that preserve the necessary characteristics that are required to estimate the velocity. These images, also known as rectified images, represent a “top view” of the observed scenario. In this paper, a technique is presented to automatically obtain these images, estimate vehicle velocity and count the number of vehicles per lane. It is known that the fundamental matrix can be
Computed from pixel co-ordinates of corresponding points in uncalibrated images

Until now, many methods have been proposed for estimating fundamental matrix. These methods can be divided into three classes: linear methods, iteration methods and robust methods. Linear methods can largely reduce the computing time. Iteration method s and robust methods are more accurate but they are time-consuming. Linear methods can supply good initial values for iteration methods. Robust methods are just repeated applications of linear methods in which weightings of the residuals of corresponding points are introduced so that the false matches between two images can be removed. Thus, linear methods are foundation of various methods of estimating the fundamental matrix. Linear methods are mainly based on various least–squares methods, and the method using Eigen analysis can obtain the best results because an orthogonal least squares minimization is more realistic than the classic one.

In this paper, we propose a new linear approach to estimating the fundamental matrix. Instead of using the eigenvector corresponding to the smallest Eigen value from the orthogonal least-squares method to from the fundamental matrix, we make use of two eigenvectors corresponding to the two smallest Eigen value to construct a 3*3 generalized eigenvalue problem. This paper is organized in four main sections. The first, named Image rectification focuses on the method that originates rectified images. The second, on the procedure that leads to the determination of the necessary scale factor, while the third, to the process that estimates each vehicle’s velocity.

RANSAC-like methods make it possible to detect outliers, but the inaccuracy in the image point location is rarely modeled. Seminal contributions w.r.t this are MLEASAC and MAPSAC where point location is endowed with a Gaussian model. Such methods require to first estimate not only the variance of the point image location, but also outlier distribution and prior probabilities. Besides, they focus on model fitting and do not take into account the uncertainty of the fundamental matrix estimation from the minimal set although it is also spoilt by inaccuracy.

2. Image Rectification

Image rectification is the process of re-aligning corresponding epipolar lines to become collinear and parallel with the X axis. For a stereo sensor, mechanical adjustments of this caliber are difficult to achieve. However, given a description of the projective geometry between the two views, projective transformations can be applied resulting in rectified images. We choose the projective transformations uniquely to minimize distortions and maintain as accurately as possible the structure of the original images. This helps during subsequent stages, such as matching, ensuring local areas are not unnecessarily warped. Rectification can be described by a transformation that sends the Epipoles to infinity; hence the epipolar lines become parallel with each other. Additionally, we ensure that corresponding points have the same y coordinate by mapping the Epipoles in the direction $e=(1,0,0)^T$ or equivalently $e=(e_u,0,0)^T$.

Rectification can be the perspective transformation associated to image formation, distorts certain geometric Properties, such as length, angle and area ratios. Consequently, the employment of video or image sequences in traffic surveillance is challenging, in particular for the task of vehicle velocity estimation. How-ever, this problem can be solved by using rectified images that restore the lost geometric properties to the images of the monotorized scenario. In this paper, a method presented by D. Liebowitz is successfully employed in the rectification of images obtained by uncalibrated cameras. This technique requires the estimation of two vanishing points and the prior knowledge of two angles on the ground plane. Given the nature of a roadway structure, i.e. the large amount of parallel and perpendicular lines, these parameters can be easily obtained. In a general manner, this method estimates the projective transformation by establishing three matrices or trans-formations.

$$H = Hs.Ha.Hp \qquad (1)$$

Where

H_s represent the similarity transformation,
H_a the affine and H_p the pure projective transformation.

Each one of these transformations is responsible for the reinstatement of certain geometric and metric properties and can be achieved using known parameters on the image and ground planes. Namely, the pure projective transformation is responsible for restoring line parallelism and area ratios to the scenario. This transformation can be easily acquired by estimating the homogeneous representation of the vanishing line. Once known the location of two vanishing points, this representation is quite straightforward, as can be seen in the following equation:

$$l = l_1 \quad l_2 \quad l_3 \qquad T = v_1 \, x \, v_2 \qquad (2)$$

Where

l is the homogeneous representation of the vanishing line and
v_1 and v_2 the vanishing points that are represented on the upper left box in Figure 1.

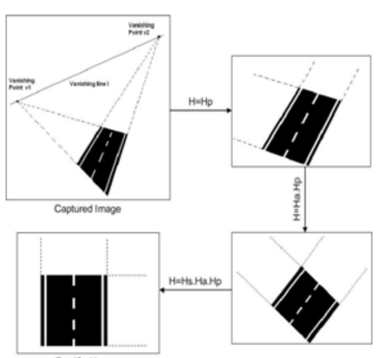

Fig. 1. Stages of the rectification process.

Therefore, the pure projective transformation can be resented by the following matrix:

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix} \qquad (3)$$

Hence, a correct estimation of this transformation relies on the accurateness of the location of the vanishing points. These are obtained by applying the Hough transform to edges extracted from the imaged highway lanes and to edges identified on the foreground image. An activity map is estimated due to the fact that edges detected from the background image may contain unnecessary edges. An activity map is an image that represents the regions that contain movement throughout the image sequence, i.e. the regions of interest. The required vanishing point is obtained applying the least squares method to the obtained lines.

To determine the second vanishing point it is necessary to identify lines on the image that, on the ground plane, are perpendicular to the highway lanes. Given a vehicle's orthogonal structure, these lines can be obtained using edges extracted from segmented vehicles.

Figure 2 illustrates edges detected from the foreground using a Sobel edge detector. The lines represented in Figure 2 where obtained from the edges rep-resented in Figure 2.a and are only a few of the used in the calculation of the second vanishing point. Given the fact that a great amount of lines obtained from the foreground image

(a) (b)

Fig. 2. (a) Edges detected by applying the Sobel Detector to the segmented vehicles. (b) Lines identified using the edges represented in (a) are outliers, the estimation of this vanishing point using the least squares method is ill-conditioned.

Thus, a RANSAC based algorithm is used in this estimation. In order to do so, it is necessary to adopt a 3D homogeneous representation for the extracted lines, seeing as this form of representation takes into account points at infinity. Each iteration of the RANSAC algorithm estimates a possible vanishing point using equation (4), where l_1 and l_2 represent the homogeneous coordinates of two lines, and the one with the largest number of inliers is taken as being correct.

$$p = l_1 \times l_2, \qquad (4)$$

On the other hand, the affine transformation reinstates angle and length ratios of non parallel lines, and can be obtained using two known angles on the ground plane as explained. This approach estimates two parameters α and β using constraints on the ground plane. These parameters represent the coordinates of the circular points on the affine plane. Liebowitz and Zisserman in propose three types of constraints:

- A known angle on the ground plane;
- equality of two unknown angles;
- a known length ratio.

Given the orthogonal structure of the highway lanes, we chose to employ the first constraint in this algorithm, i.e. a known angle on the ground plane. Each known angle on the ground planes defines a constraint circle. This fact is quite useful seeing as α and β lie within this circle represented on a complex space
defined by (α, β). Therefore, in order to obtain the required parameters one may estimate the intersection of two constraint circles obtained using two different known angles.

$$H_a = \begin{bmatrix} \frac{1}{\beta} & \frac{-\alpha}{\beta} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (5)$$

To conclude, the last transformation which is known as similarity transformation performs rotation, translation and Isotropic scaling of the resultant image.

$$H_s = \begin{bmatrix} s.\cos\theta & -s.\sin\theta & t_x \\ s.\sin\theta & s.\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \qquad (6)$$

3. Epipolar geometry

The epipolar geometry describes the relations that exist between two images. Every point in a plane that passes through both centers of projection will be projected in each image on the intersection of this plane with the corresponding image plane. Therefore these two intersection lines are said to be in epipolar correspondence. This geometry can easily be recovered from the images even when no calibration is available. In general 7 or more point matches are sufficient. Robust methods to obtain the fundamental matrix from images are described in [16, 18]. The epipolar geometry is described by the following equation:

$$M'^{T}F_{m}=0 \qquad (1)$$

Where and are homogenous representations of corresponding image points and is the fundamental matrix. This matrix has rank two, the right and left null space cor-responds to the epipoles and which are common to all epipolar lines. The epipolar line corresponding to a point is thus given by with meaning equality up to a non-zero scale factor (a strictly positive scale factor when oriented geometry is used, see further).

$$1' \sim H^{-T} 1 \text{ or } 1 \sim H^{T} 1' \qquad (2)$$

with a homography for an arbitrary plane. Valid homography can immediately be obtained from the fundamental matrix:

$$H= [e']_{X} F+e'a^{T} \qquad (3)$$

with a random vector for which det so that is invertible. If one disposes of camera projection matrices an alternative homography is easily obtained as:

$$H^{-T}=(P'^{T})^{\dagger} P^{T} \qquad (4)$$

Where indicates the Moore-Penrose pseudo inverse.

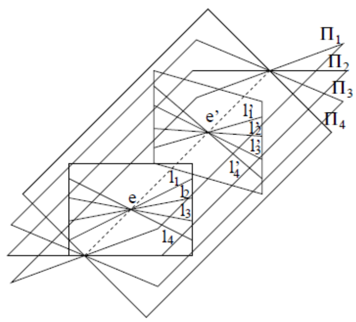

Figure 1. Epipolar geometry with the epipoles in the images. Note that the matching ambiguity is reduced to half epipolar lines.

Orienting epipolar lines, the epipolar lines can be oriented such that the matching ambiguity is reduced to half epipolar lines instead of full epipolar lines. This is important when the epipole is in the image.

Figure 1 illustrates this concept. Points located in the right halves of the epipolar planes will be projected on the right part of the image planes and depending on the orientation of the image in this plane this will correspond to the right or to the left part of the epipolar lines.

In practice this orientation can be obtained as follows. Besides the epipolar geometry one point match is needed (note that 7 or more matches were needed anyway to determine the epipolar geometry). An oriented epipolar line separates the image plane into a positive and a negative region.

Note that in this case the ambiguity on is restricted to strictly positive scale factor. For a pair of matching points both and should have the same sign. Since is obtained from through equation (2), this al-lows to determine the sign of. Once this sign has been determined the epipolar line transfer is oriented. We take the convention that the positive side of the epipolar line has the positive region of the image to its right. This is clarified in Figure 2.

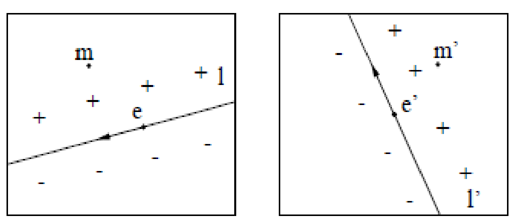

Figure 2. Orientation of the epipolar lines.

4. Finding Subsequent Homographies

Image pairs often contain more than one plane. When a plane is identified, the matches compatible with it, can be removed from the candidate set. Incorrect matches having at least one of their points lying in a part of an image where the detected plane is visible, can also be eliminated.

To decide which matches should be eliminated, the region of the image that agrees with the homography must be extracted. This is done by considering the result of applying the homography to one of the images in the pair. This trans-formed image should correspond almost exactly to the other one, in the area of the detected plane, while disagreeing in other regions. Thus, the absolute difference of the two images is computed, and threshold. The obtained binary image is then cleaned up, using morphological operators. A closing using a small mask to eliminate the isolated errors in the plane,

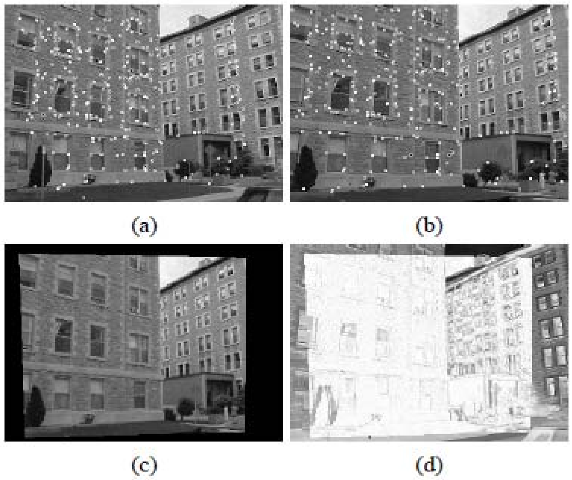

Figure3. Finding a second homography (a) Difference between the left image and Fig. 2 d. (b) Regions in Fig. 2 c that agreed with the homography. (c) an (d) The point matches that agree with a second homography found after eliminating matches in the regions of (b) from the candidate matches set.

and threshold. The obtained binary image is then cleaned up, using morphological operators. A closing using a small mask to eliminate the isolated errors in the plane region followed by an opening with a larger mask, to eliminate the isolated small areas falsely attached to the plane. All matches with a point in the identified region are discarded before attempting to find another homography. Fig. 3 illustrates the process of finding a second homography.

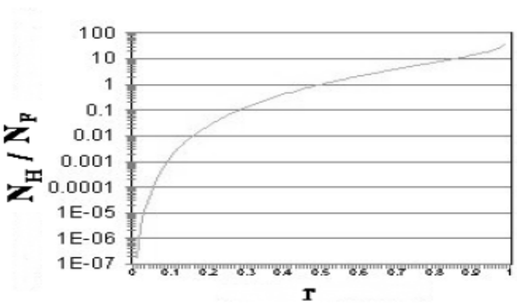

Figure4. The number of trials required to find a planar homography in a scene compared to the number of trials required for weak calibration, both using RANSAC, as a function of the reliability of the match set used.

Where p is the number of matches needed to compute a solution (4 for **H**, 8 for **F**), $^{2}{}_{out}$ is the proportion of the matches that are not compatible with the homography or fundamental matrix, and N the number of iterations performed. In the case of homography detection, $^{2}{}_{out}$ is larger since not all good matches lie on a single plane. We will make the reasonable assumption that half of the good matches are compatible with some homography. The ratio between the number of iterations that would be necessary to perform the two tasks with the same probability of success can be computed from (8):

$$\frac{N_{\mathbf{H}}}{N_{\mathbf{F}}} = \frac{\log\left(1 - r^8\right)}{\log\left(1 - \left(\frac{1+r}{2}\right)^4\right)}$$

Where N_H / N_F is the ratio of the number of trials needed to compute **H** to the number of trials needed to compute **F**, and r is the proportion of valid matches in the data set.

5. Simulation Results

Fig. 5 shows a homography that was identified in an image pair where the perspective difference between the images is significant. The four points were selected so that pairs of points lied on two different automatically dected lines. The algorithm succeeded in finding a homography despite the wide baseline.

(a) (b)

Figure 5. An image pair, and the matches agreeing with a detected planar homography.

5. Conclusion

In this paper we have proposed a new rectification algorithm. Although the procedure is relatively simple it can deal with all possible camera geometries. In addition it guarantees a minimal image size. Advantage is taken of the fact that the matching ambiguity is reduced to half epipolar lines. The method was implemented and used in the context of depth estimation. The possibilities of this new approach were illustrated with some results obtained from real images.

7. References

1. B. Bose and E. Grimson, “Ground plane rectification by tracking Moving objects,” in Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2003.
2. General motion estimation and segmentation from image sequences by M.Bober 1994.
3. Robust motion analysis, by M.Bober and J.kitter June 20-24, 1994, Pages 947-952-.IEEE Computer society press, 1994.
4. M.Bober.General Motion Estimation and Segmentation from Image Sequences. PhD thesis, VSSP Group, University of Survey, United Kingdom, 1994.
5. R.I Hartely.Estimation of relative camera positions for uncalibrated Cameras. In European Conference on Computer Vision, pages 579-587, 1993.
6. R.Hartely, “Cheiralityinvariants”, Proc. D.A.R.P.A. mage Understanding Workshop, pp.743-753, 1993.