Comparison of Fundamental Matrix Calculation Methods for Rectification of Stereo Images

[1]Vipil Vinod, [2]Jayalakshmi Surendran, [3]S Ashok

M-tech student, Department of electrical engineering, National Institute of Technology, Calicut,
Deputy Manager, Global R&D, Crompton Greaves ltd, Mumbai,
Professor, National Institute of Technology, Calicut.
Email: [1]vipilvinod26@gmail.com, [2]jayalaksmi.surendran@cgglobal.com, [3]ashoks@nitc.ac.in

Abstract. Image rectification is the transformation process to project both the images onto the same image plane. It simplifies the problem of matching correspondence points in the two images. For image rectification fundamental matrix is to be found out. There are mainly two methods to calculate fundamental matrix, in one method surf features and RANSAC is used to get fundamental matrix and in other case corresponding points found manually are used to find out fundamental matrix.Comparison of these methods is discussed in this paper. Along with this the main processing steps for calculating depth map using stereo vision is also discussed.

Keywords: Image rectification, Epipolar geometry, fundamental matrix, RANSAC.

I. INTRODUCTION

Stereo vision is a technique by which 3d information is obtained from a pair of images. In stereo vision different views of the same scene is obtained by stereo cameras. By comparing these two images, the relative depth information can be obtained, in the form of disparity maps, which are inversely proportional to the difference in distance between the objects.

Disparity map refers to the apparent pixel difference or motion between a pair of stereo images. To experience this, try closing one of your eyes rapidly while opening the other. Objects that are close to you will appear to jump a significant distance while objects further away will move very little, this apparent motion is the disparity. In case of a human visual system this disparity is due to two eyes which are translated with respect to each other.In case of obtaining stereo images using two different cameras it might be difficult to maintain perfect translation between two cameras due to issues with installation of cameras. So stereo images taken may be accompanied by some unintended rotation,and hence we need to do some preprocessing before calculating disparity map. The main step is the image rectification .Image rectification is the method by which images are projected back to a common plane to allow comparison of the image pairs. For image rectification, calculation of fundamental matrix is necessary. There are two methods to calculate fundamental matrix, in one method surf features and RANSAC is used to get fundamental matrix and in other case corresponding points found manually are used to find out fundamental matrix.

Stereo vision is highly important in fields such as roboticsto extract information about the relative position of 3D objects in the vicinity of autonomous systems, for example one chair in front of another, which the robot may otherwise not be able to distinguish as a separate object by any other criteria.[1][2]. In Video surveillance systems stereo vision might be very useful in increasing the robustness of various automated video analytics algorithms such as trip zone, object detection, motion detection, counting and etc. In these systems depth information can be utilized to reduce error and noise by filtering out depths which are irrelevant.

The organization of this paper is as follows. Section II discusses the depth estimationfrom stereo vision. In section III overview of the fundamental matrix calculation is discussed. In section IV experimental results is discussed and final section V is conclusion.

II. DEPTH ESTIMATION FROM STEREO VISION

Depth from stereo is based on the triangulation principle. Stereo cameras are used to obtain stereo imagessuch that their views overlap each other. The set up is as shown below.

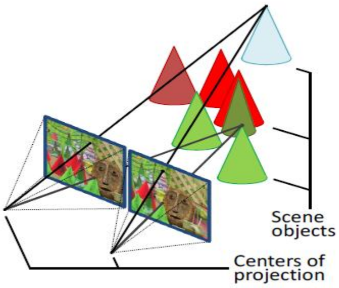
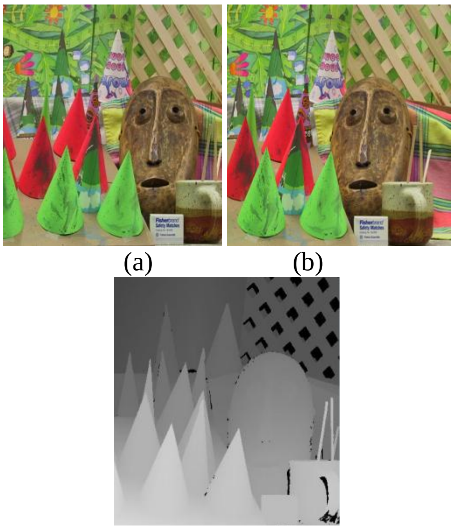


Figure 1 stereo images

When we have right and left images taken from two cameras we need to find the point correspondence between the images. This can be done by using astereo matching algorithm. But the question arises; do we have to search in the entire image for point correspondence? The answer is no. Epipolar geometry [7] reduces the search to one dimension, i.e. for each pixel in the left image we can now search for the pixel on a single line in the right image; this line in the right image is known as the epipolar line. But since a single pixel value is not discriminative enough to reliably find the correspondence, one usually tries to find a match for a small window around each pixel against all possible windows in the right image along a line.In addition we can restrict the search to a region of interest instead of searching along the complete line. If good correspondence match is found then we can associate the left image pixel with corresponding right image pixel. The association is stored as the disparity map in the form of offset between the pixel postion.The stereo matching algorithms like SAD; SSD can be used to find the point correspondence. [2]


(a) (b)

(c)

Figure 2 a) left image b) Right image c) Disparity map. [Courtesy: www.middleburry.edu]

A. Image rectification

In cases when the relative displacement between cameras is pure translation (say horizontal) epipolar line for an object in the left image is along the same row in the right image. But if the relative motion includes rotation also then computation of epipolar line becomes very difficult. Disparity map computation may produce wrong results in such scenarios. The proof of this is shown in the results. To overcome this difficulty we employ a method called image rectification.

Image rectification is a method in which both the images are projected onto a common plane. So the scan line will be along the same row. For the image rectification we need to find the fundamental matrix. [7]

The fundamental matrix is a 3x3 matrix which relates the corresponding points in the stereo images .It contains the intrinsic and extrinsic parameters. The intrinsic parameters include the focal length, image centre, parameters of lenses distortion, etc.The extrinsic parameters are the relative rotation and translation between the cameras.

If the cameras themselves are calibrated for internal parameters, an essential matrixprovides the relationship between the cameras. The more general case (without camera calibration) is represented by the fundamental matrix. The essential matrix can be seen as a precursor to the fundamental matrix.

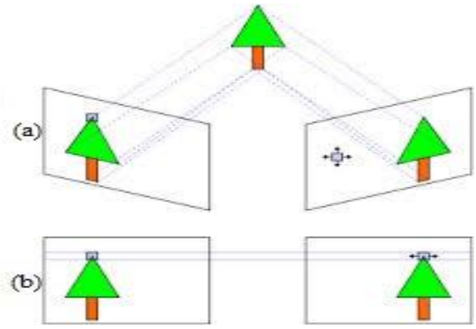


Figure 2.Search space before (a) and after (b) rectification

From the above figure it is seen that after image rectification the correspondence search is reduced to one dimension.

B. Epipolar geometry and fundamental matrix

The epipolar geometry is the projective geometry which gives the relation between two images taken by the camera.It is independent of scene structure and only depends on internal parameters and relative pose.

The fundamnetal matrix F contains the intrinsic geometry,it is a 3x3 matrix of rank 2.If a point in 3-D space is imaged as x in the first view andx' in the second,then the image points satisfies relation $x'^T F x=0$.

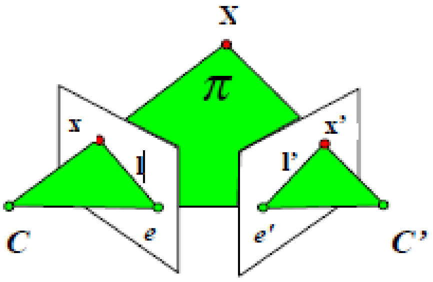


Figure 3.epipolar geometry

Given a point x in one image we have corresponding epipolar line I' in the other image.Any point x' corresponding to x must lie on the epipolar line I' .Epipolar line is the projection in the second image of the ray from the point throught the camera centre C of the first camera .There is a map

$$x \mapsto l'$$

from a point in one image to its corresponding epipolar line in the other image[3][4][5][6]

III.OVERVIEW OF FUNDAMENTAL MATRIX CALCULATION

The fundamental matrix is defined by the equation, $x'^{T} F x = 0$,for any pair of corresponding points$x$andx'in the 2 images**.**Each pair of correspondence gives 1 constraint. To recover fundamental matrix up to a scale we need at least 8 constraints i.e. 8 pairs of correspondences. The equation for a pair of points (x_1, y_1, 1) and (x_2, y_2, 1) is:

$$x_1x_2f_{11} + x_2y_1f_{12} + x_2 f_{13} + y_2x_1 f_{21} + y_2y_1 f_{22} + y_2f_{23} + x_1f_{31} + y_1f_{32} + f_{33} = 0$$

We have$AF=0$for n point matches, where A is formed by stacking the data vectors. The solution is the eigen vector corresponding to the smallest eigen values of $A^{T}A$.We call it “8 point algorithm’’ because we use 8 points to get the fundamental matrix.

Fundamental matrix can be calculated either by providing manually the corresponding points or by using Surf features to get the corresponding points .The method are discussed in the subsection.

A. Manually matching 8 points.

In this case fundamental matrix calculation is carried out by acquiring and processing stereo pairs of images. The corresponding points must be matched manually in this case. A point is marked in left image and the corresponding point is marked in the right image or vice versa. Minimum of 8 points are required to apply the 8 point algorithm.Manually matching the eight points might result in greater accuracy, but since it is offline it does not deal with unexpected change in extrinsic parameters due to rotation or translation of cameras.

B. Using surf features

In this case the corresponding points are calculated using the surf features and the fundamental matrix is found out using 8 point algorithm.SURF (Speeded up Robust Features) [8] is a robust local feature detector.There are three main steps

1. Feature detection

First, ‘interest points’ are selected at distinctive locations in the image, such as corners, blobs etc. The most valuable property of an interest point detector is its repeatability, i.e. whether it reliably finds the same interest points under different viewing conditions.

2. Feature extraction

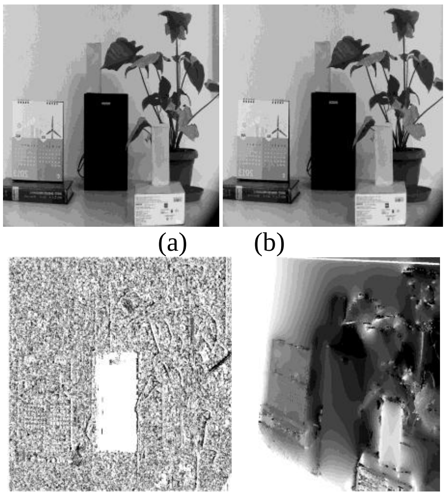
The neighbourhood of every interest point is represented by a feature vector. This descriptor has to be distinctive and, at the same time, robust to noise, detection errors, and geometric and photometric deformations

3. Feature matching

In the final step, the descriptor vectors are matched between different images. The matching is often based on a distance between the vectors. The dimension of the descriptor has a direct impact on the time it takes for computation, and a lower number of dimensions are therefore desirable. Mainly the sum of absolute difference algorithm is used. After getting the correspondence points RANSAC [1] is used to estimate the fundamental matrix.

IV. EXPERIMENTAL RESULTS

At first Stereo image of a scene is taken. After that depth map is calculated before and after rectifying the images. The results are as shown.


(a) (b)

(c) (d)

Figure4 a) Right image b) left image c) Disparity map before rectification d) Disparity map after rectification

As you can see from the above figure that disparity map obtained after rectification provides better results compared to that without rectification.

A. Fundamental matrix calculation

1. Manually matching points

In this case 8 points are matched manually as shown in below figure.

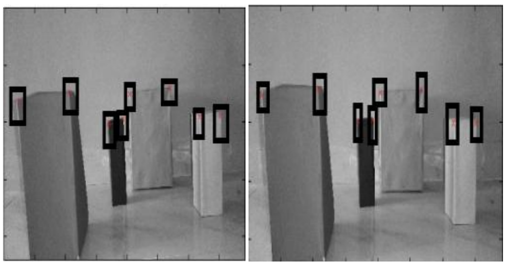

(a) (b)

Figure 5 corresponding points matched in (a) and (b)

Point are marked in figure (a) corresponding to that a points are then marked in figure (b).these 8 points are used to calculate the fundamental matrix.

2. Using surf features

In this method the surf feature matching method is used to match points. At first feature is detected and then feature vector is extracted and finally the features are matched. Surf features detected are shown in figure 6(green mark).

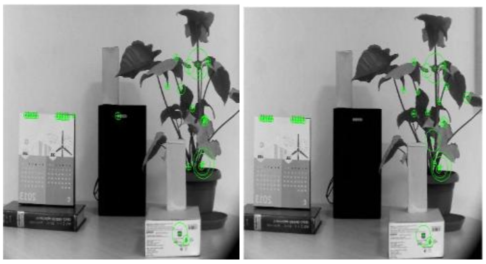
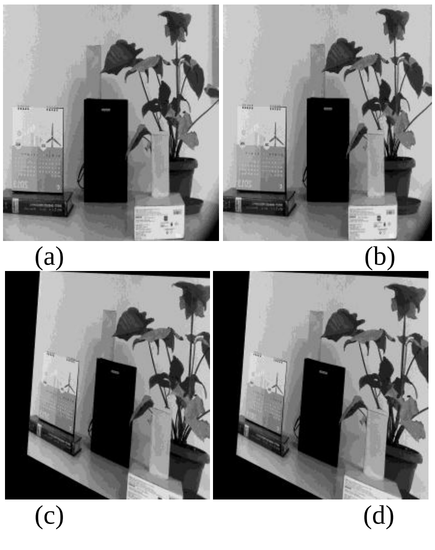

(a) (b)

Figure 6 (a) and (b) shows points detected by surf feature in left and right images

In this case many corresponding points are found out.RANSAC is used to find the fundamental matrix from these corresponding points.

B.Image rectification using Surf features and RANSAC

In this method the corresponding point are found out using the surf features [8].RANSAC is used to find the fundamental matrix. The result is as shown below. After finding the fundamental matrix rectification is carried out.

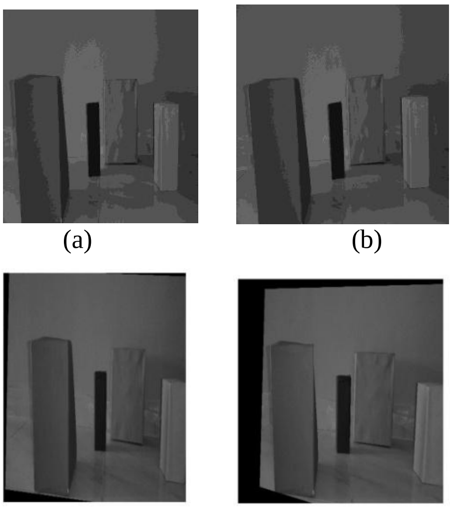

(a) (b)

(c) (d)

Figure5 a)Right image b)Left image c)Right rectified image d)Left rectified image

C.Image rectification by manually matching points

Corresponding points are manually matched in this case. Minimum of 8 points are required as mention in the above section. The fundamental matrix calculated is then used to rectify the images. The result is as shown below.


(a) (b)

(c) (d)

Figure6 a)Right image b)Left image c)Right rectified image d)Left rectified image

D. Comparison of results.

Advantage of method used to find the fundamental matrix depends on the application in which it is used. To find the image rectification for stereo vision surf features is advantageous because of the following reasons. It detects the corresponding points automatically .There is no manual intervention. Time taken for overall process is less in case of corresponding points.

V. CONCLUSION

Image rectification simplifies the problem of matching corresponding points. It produces better results as it is proved experimentally. For image rectification fundamental matrix is calculated. Fundamental matrix can be calculated by two methods, in one method corresponding points are matched using manually, in other method corresponding points are matched using surf features.Manually giving the eight points might result in greater accuracy, but since it is offline it does not deal with unexpected change in extrinsic parameters due to rotation or translation of cameras. With surf features the corresponding points are matched automatically providing better results.

REFERENCES

[1] Trucco, E Verri, A."Introductory Techniques for 3-D Computer Vision." Prentice Hall, 1998.

[2] Hartley, R; Zisserman, A. "Multiple View Geometry in Computer Vision." Cambridge University Press, 2003

.[3] Fusiello,Andrea, Trucco, Emanuele; Verri, Alessandro (2000-03-02). "A compact algorithm for rectification of stereo pairs". Machine Vision and Applications (Springer-Verlag)

[4] Pollefeys, Marc; Koch, Reinhard; Van Gool, Luc (1999). "A simple and efficient rectification method for general motion". Proc. International Conference on Computer Vision: 496–501.

[5] Lim, Ser-Nam; Mittal, Anurag; Davis, Larry; Paragios, Nikos. "Uncalibrated stereo rectification for automatic 3D surveillance".International Conference on Image Processing

[6] J. Gluckman and S. K. Nayar, "Rectifying Transformations that Minimize Resampling Effects", Proc. Int. Conf. on Computer Vision and Pattern Recognition, vol. 1, pages 111-117, 2001.

[7] Richard I. Hartley, "Theory and Practice of Projective Rectification" International Journal of Computer Vision, vol 35, no 2, pages 115-127, 1999.

[8] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, SURF: Speeded Up Robust Features, Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008

◈◈◈