International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:11, No:9, 2017

An Efficient Fundamental Matrix Estimation for Moving Object Detection

Yeongyu Choi, Ju H. Park, S. M. Lee, Ho-Youl Jung

***Abstract*—**In this paper, an improved method for estimating fundamental matrix is proposed. The method is applied effectively to monocular camera based moving object detection. The method consists of corner points detection, moving object's motion estimation and fundamental matrix calculation. The corner points are obtained by using Harris corner detector, motions of moving objects is calculated from pyramidal Lucas-Kanade optical flow algorithm. Through epipolar geometry analysis using RANSAC, the fundamental matrix is calculated. In this method, we have improved the performances of moving object detection by using two threshold values that determine inlier or outlier. Through the simulations, we compare the performances with varying the two threshold values.

***Keywords*—**Corner detection, optical flow, epipolar geometry, RANSAC.

I. Introduction

RECENTLY, vehicles equipped with ADAS (Advanced Driver Assistance Systems) are on the market. ADAS is now considered essential because it provides not only convenience but also safety to drivers. It is regarded as an intermediate stage for fully-operated autonomous vehicles. Assistance systems and autonomous vehicles are not separate concepts. The level of ADAS will ultimately evolve into autonomous technology. On a typical road, drivers and pedestrians are always exposed to unexpected accident. In particular, collisions with pedestrians or other vehicles are very dangerous. In order to prevent these accidents, many researchers have studied on moving object detection. Yamaguchi proposed a method for moving object detection [1]. It analyzes epipolar geometry between consecutive two images in monocular camera. Next, corner points are detected by Harris corner detector [2], essential matrix is estimated by using 8-point algorithm [3] and RANSAC [4]. In this paper, we estimate fundamental matrix including camera intrinsic parameter, and propose to use two threshold values. One is value to form maximum consensus in RANSAC, another one is used for determining if feature points are inlier or outlier.

In the simulations, we test evaluation on video data recorded by front and rear camera mounted at a vehicle. For performance evaluation, we use recall, precision, and F-measure representing the result of integrating both measures.

This paper is organized as follows: In the next section, we explain the related works. In Section III, we present the method which uses two threshold values. In Section IV, the performances of simulated result are described. Finally, we conclude the paper in Section V.

Yeongyu Choi is with the Department of Information and Communications Engineering, Yeungnam University, Gyeongsan, Republic of Korea (e-mail: ygchoi1111@gmail.com).

Ju H. Park is with the Department of Electrical Engineering, Yeungnam University, Gyeongsan, Republic of Korea (e-mail: jessie@ynu.ac.kr).

S. M. Lee is with Department of Electronic Engineering, Kyungpook National University, Daegu, Republic of Korea (e-mail: moony@knu.ac.kr).

Ho–Youl Jung is with the Department of Information and Communications Engineering, Yeungnam University, Gyeongsan, Republic of Korea (e-mail: hoyoul@yu.ac.kr).

II. Related Works

To analyze the epipolar geometry from consecutive two images, the feature points are detected. These features need to be robust enough to match the same points in two images, such as corner points. At any t and $t-1$, corner feature points are detected by Harris corner detector and matched. And then, optical flows are estimated by pyramidal Lucas-Kanade algorithm. Optical flows describe not only motions, but also the corresponding points of feature points. These corresponding points at time t would almost match feature points at time $t-1$. The fundamental matrix can be computed by using an 8-point algorithm between two images with matching feature points. However, any feature point may be point of moving object. This optical flow would be a vector of summation of the two flows generated by the camera's motion and its own motion. Therefore, existing research estimates parameters using RANSAC [1].

A. Harris Corner Detection

Harris et al. [2] proposed the following method that detects traceable feature points. When a point $p(x,y)$, local window W, and the shift vector $(\Delta x, \Delta y)$ are indicated, the sum of squared differences S between these 2 local window is given by:

$$S(x,y) = \sum_W [I(x,+\Delta x, y+\Delta y) - I(x,y)]^2 \qquad (1)$$

Equation (1) is rewritten as (2) by a first-order taylor series approximation.

$$S(x,y) = [\Delta x \quad \Delta y] M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \qquad (2)$$

When two eigenvalues of M are λ_1, λ_2, the flat, edge, and corner are determined according to the scale and difference of the two eigenvalues. The result is given as:

$$R = \det(M) - k * trace(M)^2 \qquad (3)$$

where, the value k $(0.04 \le k \le 0.06)$ is determined as the experienced knowledge of the parameter. $(|R| = small)$, $(R < 0)$, $(R > 0)$ means flat, edge, and corner respectively.

B. Pyramidal Lucas-Kanade Algorithm

Pyramidal Lucas-Kanade is a sparse optical flow and can find large motion by constructing pyramid by size from upper layer to lower layer of original image.

C. Epipolar Geometry and Fundamental Matrix

Essential matrix is a relative geometry represented algebraically in two normalized camera coordinates. Let p and p′ be respectively the normalized camera coordinates with z-coordinate value 1 in image at time $t-1$ and t [5], then,

$$\mathrm{p'^{T}Ep} = 0 \tag{4}$$

where, E is 3×3 matrix including camera's rotation and translation. In fact, there exists difference between the camera's image plane and normalized camera coordinates because of lens distortion and focal length. If the transformation between the two planes is the camera intrinsic parameter K, (4) can be rewritten as

$$m'^{T}\mathrm{K^{T}EK}m = 0 \tag{5}$$

where, m, m' are homogeneous image coordinates. Through (5) fundamental matrix is given as:

$$m'^{T}\mathrm{F}m = 0\ (\mathrm{F} = \mathrm{K^{T}EK}) \tag{6}$$

where, F is 3×3 matrix, which contains intrinsic parameter K in essential matrix E. Finally, (6) is rewritten in (7) by all N feature points m and corresponding points m′.

$$\mathrm{A}f = \begin{pmatrix} x_1x_1' & y_1x_1' & x_1' & x_1y_1' & y_1y_1' & y_1' & x_1 & y_1 & 1 \\ x_2x_2' & y_2x_2' & x_2' & x_2y_2' & y_2y_2' & y_2' & x_2 & y_2 & 1 \\ \vdots & & & & & & & & \\ x_Nx_N' & y_Nx_N' & x_N' & x_Ny_N' & y_Ny_N' & y_N' & x_N & y_N & 1 \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{pmatrix} \tag{7}$$

Fundamental parameter f is estimated by SVD (Singular Value Decomposition) [5]. RANSAC allows the matrix A to contain the feature points of the background except for the foreground, and the fundamental matrix is estimated more accurately than from all features.

III. Proposed Method

Fig. 1 shows system flowchart for moving object detection. Fig.. 2 shows feature points using Harris corner detector. As shown this figure, it is seen that only the corners of traceable feature are detected.

The first threshold value used in fundamental matrix estimation is smaller than the second for searching outlier points. In order to reduce the residual of the calculated fundamental matrix, it is important to select stationary features without movement and variation such as structure's. For example, wavering leaves and objects with a change in brightness would affect calculation of fundamental matrix.

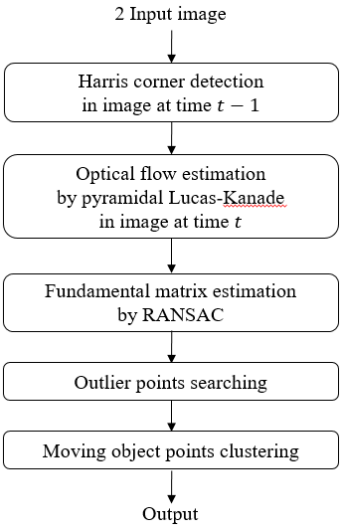


Fig. 1 Main flow chart of proposed system

(a) Original image

(b) Corner points

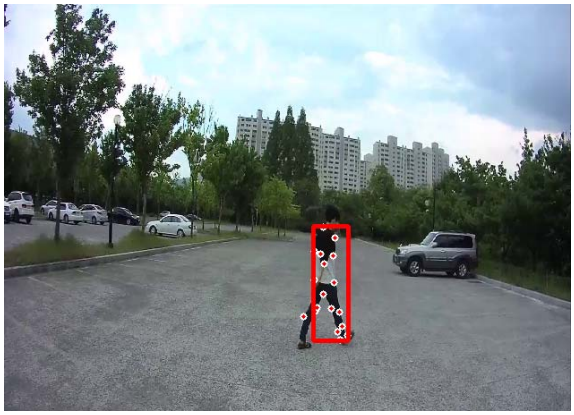
Fig. 2 An example of corner detection

If using smaller threshold value in RANSAC, the matrix A in (7) which consists of feature points is made up of stationary features with a high probability. It is because small inlier range does not permit objects with low noise value as well. And the larger threshold value for outlier search allows background points containing some noise.

To cluster the points, we considered Euclidean distance

between points, optical flow's magnitude and direction. Fig. 3 illustrates detected moving object's points and a result of clustering.

(a) Moving object's points



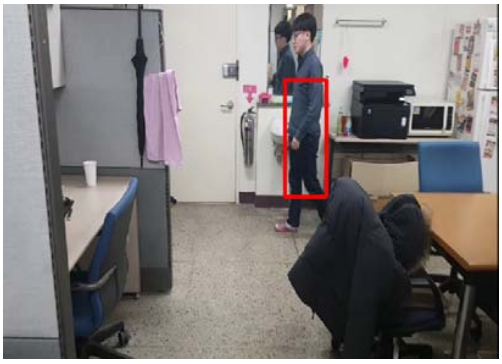
(b) A result of clustering

Fig. 3 An example of moving object detection

Fig. 4 shows result images in various environments. Fig. 4 (a) describes that a moving person is detected in a complex indoor environment, Figs. 4 (b) and (c) explain moving object detection from the front/rear camera mounted at vehicle, Fig. 4 (d) illustrates detection of 2 objects. In order to measure the similarity of each point in clustering, we considered 3 features. Therefore, as shown in Fig. 4 (d), the 2 objects could be detected separately.

IV. Experimental Results

For the evaluation, we test the method by varying the first threshold value thr_1 in various environments when the second thr_2 is fixed as 1. The experimental environment of videos consists of indoor and outdoor. The case of outdoor is recorded by cameras mounted at front/rear. The input image resolution was 640×480 pixels.



(a)

(b)

(c)

(d)

Fig. 4 Result images in various environments (a) one moving object indoor, (b) and (c) using front and rear camera outdoor, (d) two objects outdoor case

The performances are shown in Table I. the method is evaluated by recall, precision, and F-measure. As shown this table, it has better performances that the smaller first threshold value is applied at the step of fundamental matrix estimation and moving objects are determined as the larger second threshold value.

TABLE I
THE PERFORMANCES OF THE PROPOSED METHOD

	thr_1	thr_2	Recall	Precision	F-measure
	0.2	1	0.39	0.46	0.43
	0.4	**1**	**0.4**	**0.48**	**0.44**
Indoor	0.6	1	0.35	0.4	0.37
	0.8	1	0.31	0.36	0.33
	1	1	0.32	0.37	0.34
	0.2	**1**	**0.36**	**0.63**	**0.46**
	0.4	1	0.31	0.59	0.41
Outdoor (front)	0.6	1	0.3	0.54	0.39
	0.8	1	0.29	0.52	0.37
	1	1	0.25	0.41	0.36
	0.2	**1**	**0.41**	**0.58**	**0.48**
	0.4	1	0.36	0.54	0.44
Outdoor (rear)	0.6	1	0.34	0.51	0.41
	0.8	1	0.34	0.49	0.4
	1	1	0.24	0.3	0.27

V. CONCLUSION

In this paper, we present a method using two inlier range for moving object detection in dynamic scene. The performance of the method shows higher than using one inlier range. However, as shown table I, our method still has low detection rate. Therefore, we will research the way of improvement using tracking algorithm.

ACKNOWLEDGMENT

This work was supported in part by the Human Resource Training Program for Regional Innovation and Creativity through the Ministry of Education and National Research Foundation of Korea (NRF-2014H1C1A1073141).

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2017-2016-0-00313) supervised by the IITP (Institute for Information & communications Technology Promotion).

REFERENCES

[1] K. Yamaguchi, K. Takeo, and N. Yoshiki, “Vehicle Ego-motion Estimation and Moving Object Detection using a Monocular Camera," Proceedings of IEEE International Conference on Pattern Recognition, Vol. 4, pp. 610-613, 2006.

[2] C. Harris, and S. Mike, “A Combined Corner and Edge Detector," Proceedings of Alvey vision conference, Vol. 15, No. 50, 1988.

[3] R. Hartley. “In Defense of the Eight-point Algorithm," IEEE Transactions on pattern analysis and machine intelligence, Vol. 19, No. 6, pp. 580-593, 1997.

[4] M. Fischler, and R. Bolles, “Random Sample Consensus: a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Communications of the ACM, Vol. 24, No. 6, pp. 381-395, 1981.

[5] R. Hartley, and A. Zisserman, “Multiple View Geometry in computer vision,” 2000.