A New Algorithm for Target Tracking Using Fuzzy-Edge-based Feature Matching and Robust Statistic

Alireza Behrad, Seyed Ahmad Motamedi
Electrical Engineering Department
AMIRKABIR University of Technology
424 Hafez Ave. 15914 TEHRAN, IRAN
Email:*b7723958@aku.ac.ir, motamedi@aku.ac.ir*

Kurosh Madani
Intelligence in Instrumentation and Systems Lab. (I^2S)
SENART Institute of Technology - University PARIS XII - Avenue Pierre POINT,
F-77127 LIEUSAINT – FRANCE
Email: *madani@univ-paris12.fr*

Mehdi Esnaashari
Computer Engineering Department
AMIRKABIR University of Technology
424 Hafez Ave. 15914 TEHRAN, IRAN

ABSTRACT

In this paper we present a new algorithm for real-time tracking of moving targets in terrestrial scenes using a mobile camera. We used fuzzy edge of the target and modified LMedS statistic for robust tracking purpose. In this method, we first select proper feature points from the edge of the target. These feature points are then matched with points in the region of interest in the next frame using fuzzy-edge-based feature matching. Then using the modified LMedS statistic and affine transformation, a motion model is calculated for the target. Using this model the location of the target is identified in the next frame. In addition to robust statistic, the use of reflectance information for edge detection has made our tracking algorithm reliable against high illumination changes. The tracking system is also capable of target shape recovery and therefore it can successfully track targets with varying distance from camera or while the camera is zooming.

1. INTRODUCTION

Visual tracking is one of the most challenging issues in computer vision. Applications of the visual tracking are numerous and they span a wide range of applications including surveillance system, vehicle tracking and aerospace application, to name a few. There are many challenges to tracking including noise of the target, change of scene illumination, change in target geometry, real-time constraint and camera motion, to mention a few. A number of methods have been proposed for tracking of targets including motion-based segmentation, active contour methods, region matching and feature matching. Motion-based segmentation algorithms such as optical flow [1], [2] and difference-based methods [3], [4] detect moving objects in the scene to track them. These methods fail in the case of illumination change or when the target stops moving. Most of the optical flow methods also have high computation cost and are not proper for real-time applications. When the camera is mobile such as most tracking systems, conventional difference-based methods also fail to detect moving targets, because all of the objects in the image sequence have an apparent motion, which is related to the camera motion so another consideration and restriction is required. For example in [5] it is assumed that the target constitutes small portion of the scenery and based on this assumption the camera motion is estimated and compensated and in [6] it is assumed that the camera motion is only because of pan and tilt motion and using an active camera these motions are compensated.

Feature based methods [7], [8], [9] are another methods for target tracking, which can resolve many tracking problems such as camera motion, real-time constraint and geometry change of the target. In these methods a few discrete features such as points and lines are matched and tracked in consecutive frames. Existing techniques for tracking a set of discrete features generally fall into two categories: two-frame based and long-sequence based. In two-frame based methods the process of finding corresponding points over an image sequence is broken into successive, yet independent problems of two-view matching. Long-sequence based methods in other side employ smoothness constraints to exploit temporal information existing in the sequence.

In essence, when the motion of the camera and target is smooth such that smoothness constraints hold, long-sequence based methods are likely to outperform two-frame based methods. On the contrary, if the movement of

the target or camera results in nonsmooth image motion, two-frame based methods have good results.

As cited above, illumination change of scene is another challenge for target tracking. Hager and Belhumer [10] proposed a tracking algorithm based on region matching. They modeled the illumination changes into SSD formulation by using a low-dimensional linear subspace. The main disadvantage of this algorithm is the need of several images of the same target in different illumination condition. In [11] spatial illumination variations is modeled by low-order polynomial but the algorithm is not enough fast for real-time applications.

In this paper we propose a new method for real-time tracking of various targets using a mobile monocular camera. Our method is based on point features matching, which features are selected from the edge of the target and matched using fuzzy-edge-based matching algorithm. We considered noisy targets with change in geometry and illumination. We also tried to use the merits of both long-sequence and two-frame based methods.

This paper is organized as follows. In next section the procedures for extracting feature and corresponding points and are described. Section 3 gives the motion model and the calculation algorithm employed in this work. Section 4 describes the method we applied for illumination change canceling. Experimental results are shown in section 5 and conclusion appears in section 6.

2. FINDING FEATURE AND CORRESPONDING POINTS

First stage of feature tracking is to find proper features. Features such as line, corner and edge can be used for this purpose. We used point features, which are extracted from edge of the target. To find precise location of corresponding points, points are selected as features, which have enough information for matching process. So we calculate edge count and edge variances in x and y directions in windows centered on the points. We select points that edge count is between two low and high thresholds. We also reject the points that the calculated variances are lower than a threshold. Then we give grades to remaining points based on the following equation.

$$grade = k_x v_x + k_y v_y + k_c (count - L)(H - count) \qquad (1)$$

where v_x and v_y are variances in x and y directions, *count* is the edge count, L and H are low and high thresholds for edge count, which are initialized based on window size and they are adaptively adjusted if enough features are not found for a typical target and k_x, k_y, k_c are fixed parameters. After the calculation of the grades, points with higher grades are selected as feature points. When the feature points are extracted, these features are matched with points in the region of interest in the next frame. In addition to noise and illumination change, geometry change of the target may deteriorate the matching process. Direct matching of edge is not good method to find corresponding points in the presence of change in target geometry, especially when the camera is zooming. To alleviate this problem, we used a fuzzy membership function to describe edge of the target as fuzzy values. The value of this membership function is 1 in edge points and a value between 0 and 1 for adjacent points. This value is determined based on selected membership function and different membership functions such as triangular and trapezoid may be used. When the fuzzy edge is obtained for the target and the region of interest in next frame, normalized correlation method is used to locate the corresponding points in the next frame. Normalized correlation equation is given by:

$$r = \frac{\sum_{x,y \in S} [f_1(x,y) - \bar{f}_1][f_2(x,y) - \bar{f}_2]}{\left\{ \sum_{x,y \in S} [f_1(x,y) - \bar{f}_1]^2 \sum_{x,y \in S} [f_2(x,y) - \bar{f}_2]^2 \right\}^{1/2}} \qquad (2)$$

here $\bar{f}_1$ and $\bar{f}_2$ are the average values of the fuzzy membership functions in the two regions being compared, and the summations are carried out over all fuzzy values with in small windows centered on the feature points.

3. MOTION MODEL ESTIMATION

When the feature points and their matches are found, a motion model, which describes the motion of the target is calculated. Several methods have been proposed to describe the motion of the target between two frames. Among these methods three transformation: Affine, projective and polynomial, are most commonly used in target tracking and image registration [7]. The affine transformation accounts for rotation, translation, scaling and skew between two images as follows:

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} a_5 \\ a_6 \end{bmatrix} \qquad (3)$$

where (x_i, y_i) are locations of points in the current frame and (X_i, Y_i) are locations of points in the next frame and a1-a6 are motion parameters. Although projective and polynomial consider more distortions, affine transformation is easy to use and proper for real-time applications. Furthermore, in this work the motion model is computed for windows in two consecutive frames so various distortions are expected to be small. Therefore affine transformation is selected. This transformation has six parameters; therefore, three matching pairs are required to fully recover the motion. We considered the noisy targets with change in illumination and geometry so it is necessary to select the three proper points to assure an accurate model for target motion. To calculate motion model we used modified LMedS method. This method,

which also considers previous motion models of the target, can be described in following five steps.

1. Select N random feature points and find their matches using the algorithm described in section 2.
2. Select M random sets of three feature points: (xi , yi , Xi , Yi) for i=1,2,3, from the N feature points obtained in step 1. (xi ,yi) are coordinates of the feature points , and (Xi , Yi) are their corresponds in next frame.
3. For each set calculate the affine transformation parameters.
4. Transform N feature points in step 1 using M affine transformations, obtained in step 3 and calculate the M medians of squared differences between corresponding points and transformed points. Then sort these errors and select the P first affine transformation, for which the median of squared difference are the minimum.
5. From P affine transformation parameters, obtained in step 4, select the affine parameters, which is the closest to previous motion model and use it to find the location and size of the target in the next frame.

4. ILLUMINATION CHANGE CANCELING

When there is no noise and illumination change in the consecutive frames, brightness constancy assumption holds as:

$$I_0(x+u, y+v)= I_1(x, y) \quad (6)$$

where I_0 and I_1 are image intensity function in two consecutive frames and u and v are motion vectors. In the presence of noise and illumination change, the above equation can be written in its generalized form as:

$$I_0(x+u, y+v) = \alpha(x, y)I_1(x, y) + \beta(x, y) \quad (7)$$

where $\alpha(x,y)$ is a low varying function showing illumination change and $\beta(x,y)$ is the sum of image noise and a low varying part showing illumination change. Low varying part of $\beta(x,y)$ can not effect our algorithm because we used edge of image for matching . To reduce the effect of noise and $\alpha(x,y)$ we used the algorithm illustrated in figure 1. In this algorithm we first filter the image to remove noise then we calculate logarithm of filtered image to convert multiplication to addition. Now using averaging filter the low varying part of signal can be removed. Exponential function is then applied to recover $I_1(x, y)$. This algorithm also removes the low varying part of $I_1(x, y)$ but this part of signal is not important for edge detection.

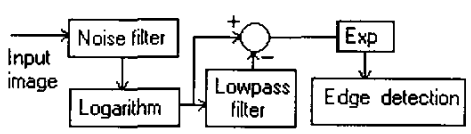

Figure 1: Illumination change removing algorithm

5. RESULTS

We first tested various edge detection and filtering algorithms to obtain a robust edge detection algorithms against noises. We tested sobel, canny, prewitt and LOG edge detection algorithms together with various filters including edge-preserving filters [12], [13]. To test these algorithms we applied various edge detection algorithms to both noiseless and filtered noise-added images then using the normalized correlation the similarity of fuzzy edge for these two images is calculated. Experimental result showed that the sobel edge detection algorithm with gaussian filter and nonmaxima suppression thinning have better results. The results for cameraman image and gaussian filter are shown in Table 1. The triangular membership function with the base of 7 pixels was used for construction of fuzzy edge in this experiment.

Table 1: Similarity measurement for different noise levels and edge detection algorithms

Noise	Canny	LOG	Prewitt	Sobel
5dB	0.317	0.551	0.918	0.935
7.5dB	0.412	0.710	0.960	0.965
10dB	0.481	0.822	0.964	0.968

The final algorithm has been implemented on a Pentium III 500Mhz using a Visual C++ program. We used triangular membership function with the base of 7 pixels. We have tested the algorithm with both simulated and actual sequences of images for various targets in different landscapes. We achieved the frame rate of 16 frames/second for tracking of 90*80 pixels target without illumination canceling algorithm and 11 frames/second in the case of illumination canceling algorithm in 352*288 pixels video frames. Proper implementation helped us to achieve such speed. Using of look up table for logarithm function, good implementation of average filter (because of its equal weights) and proper implementation of normalized correlation (considering most of edge data are zero) were some of tricks we used to have fast algorithm. We tested the proposed algorithm with wide variety of image sequence. Our algorithm can successfully track targets with gaussian noise up to 6dB and change in illumination and geometry. For simulation of illumination changes we used second order illumination model i.e. $ax^2+ by^2+cxy+dx+ey+f$, for $\alpha(x,y)$ and $\beta(x,y)$. For each frame we used random numbers for a,b,c,d,e,f and random origin which is one of the corner of the target. A typical target before and after noise and illumination addition is shown in figure 2. Figure 3 shows some tracking results for two different objects, a vehicle and a human face. As it is shown the algorithm has successfully tracked the targets. Experimental results have shown that the algorithm is reliable and can successfully track targets in most cases such as targets with noise and changes in geometry and

illumination. Comparison of results generated by the proposed method with those of other methods showed that more reliable results could be obtained with the aid of the proposed method in real-time.

Figure 2: Typical target image before and after noise and illumination addition.

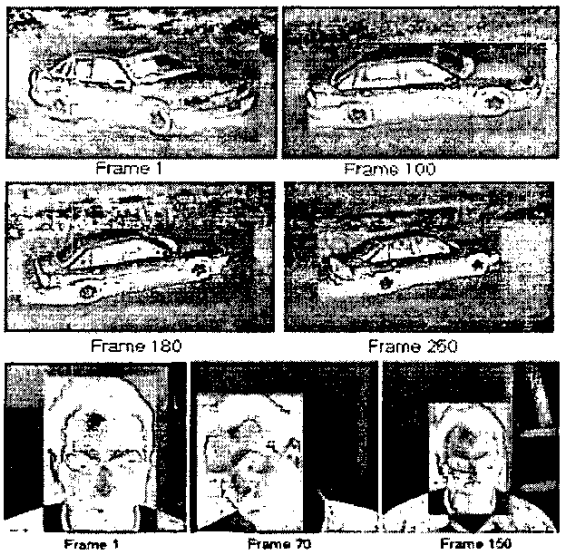

Figure 3: Tracking results for a vehicle and a face for gaussian noise of 10dB and illumination change (targets are shown without noise and illumination change).

6. CONCLUSION

In this paper we proposed a new method for tracking of the objects using mobile camera. We assumed noisy targets with change in geometry and illumination. We used edge of the target to select feature points, these feature points are matched with points in next frame using fuzzy-edge-based feature matching. Then using modified LMedS statistic an affine transformation is calculated to estimate location and size of the target in the next frame. To make our algorithm robust against illumination change we used an algorithm to extract reflectance information of image for edge detection. The proposed algorithm successfully tracks moving vehicles and objects in arbitrary scenes obtained from a mobile video camera. The tracking system is also capable of target shape recovery and therefore it can successfully track targets with varying distance from camera or while the camera is zooming. Local and regional computations have made the algorithm suitable for real-time applications. Experimental results have shown that the algorithm is reliable and can successfully track targets in most cases.

7. REFERENCES

[1] J.L. Barron, D.J. Fleet, S.S. Beauchemin and T.A. Burkitt, "Performance of optical flow techniques," *Proceeding of the computer vision and pattern recognition*, pp. 236-242, 1992.

[2] H.H. Nagel and M. Haag, " Bias-corrected optical flow estimation for road vehicle tracking", *Sixth International Conference on Computer Vision (ICCV'98)*, January 1998.

[3] A.J. Lipton, H.F. and R.S. Patil, "Moving target classification and tracking from real-time video", *Proc. IEEE Workshop Application of Computer Vision*, pp. 8-14, 1998.

[4] J. Hwang, Y. Ooi and S. Ozawa, "Visual feedback control system for tracking and zooming a target," *Proc. of the International Conference on Power Electronics and Motion Control*, IEEE 1992.

[5] S. Araki, T. Matsuoaka, N. Yokoya and H. Takemura, "Real-time tracking of multiple moving object contours in a moving camera image sequence," *IEICE Trans. Inf. & Syst.*, Vol. E83-D, No. 7, July 2000.

[6] D. Murray and A. Basu, "Motion tracking with an active camera," *IEEE Trans. Pattern Anal. & Mach. Intell.*, vol. 16, no.5, pp. 449-459.

[7] Y.S. Yao and R. Chellappa, "Tracking a dynamic set of feature points," *IEEE Trans. on Image Processing*, Vol. 4, No. 10, 1995.

[8] __, "A Robust Vision-based Moving Target Detection and Tracking System", Proceeding of Image and Vision Computing conference (IVCNZ2001), University of Otago, Dunedin, New Zealand 26 th –28 th November 2001.

[9] A. Fusiello, E. Tommasini and V. Roberto, " Improving feature tracking with robust statistics," *Pattern Analysis & Applications*, Vol. 2, pp. 312-320, 1999.

[10] G.D. Hager and P.N. Belhumeur, "Efficient region tracking with parametric Models of Geometry and illumination," *IEEE Trans. Pattern Anal. & Mach. Intell.*, Vol. 20, No 10, 1998.

[11] S.H. Lai and M. Fang, "Robust and efficient image alignment with spatially varying illumination model," *Proceedings of IEEE Conference on Computer Vision & Pattern Recognition*, Vol. 2, pp. 167-172, June 23-25, 1999.

[12] X. wang, "Optimal Edge-preserving hybrid filter," *IEEE Trans. On Image processing*, Vol. 3, No. 6, 1994.

[13] X. wang, "On gradient inverse weighted filter," *IEEE Trans. on Image processing*, Vol. 40, No. 2, 1992.