# A Robust Vision-based Moving Target Detection and Tracking System

Alireza Behrad, Ali Shahrokni, Seyed Ahmad Motamedi

Electrical Engineering Department
AMIRKABIR University of Technology
424 Hafez Ave. 15914 TEHRAN, IRAN
Email: b7723958@aku.ac.ir, ashahrokni@yahoo.com, 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

## Abstract

In this paper we present a new algorithm for real-time detection and tracking of moving targets in terrestrial scenes using a mobile camera. Our algorithm consists of two modes: detection and tracking. In the detection mode, background motion is estimated and compensated using an affine transformation. The resultant motion-rectified image is used for detection of the target location using split and merge algorithm. We also checked other features for precise detection of the target location. When the target is identified, algorithm switches to the tracking mode. Modified Moravec operator is applied to the target to identify feature points. The feature points are matched with points in the region of interest in the current frame. The corresponding points are further refined using disparity vectors. The tracking system is capable of t*a*rget 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. The refined points define the new position of the target in the current frame. Experimental results have shown that the algorithm is reliable and can successfully detect and track targets in most cases.

**Key words**: real time moving target tracking and detection, feature matching, affine transformation, vehicle tracking, mobile camera image.

## 1 Introduction

Visual detection and tracking is one of the most challenging issues in computer vision. Application of the visual detection and tracking are numerous and they span a wide range of applications including surveillance system, vehicle tracking and aerospace application, to name a few. Detection and tracking of abstract targets (e.g. vehicles in general) is a very complex problem and demands sophisticated solutions using conventional pattern recognition and motion estimation methods. Motion-based segmentation is one of the powerful tools for detection and tracking of moving targets. It is simple to detect moving objects in image sequences obtained by stationary camera [1], [2], the conventional difference-based methods fail to detect moving targets when the camera is also moving. In the case of mobile camera all of the objects in the image sequence have an apparent motion, which is related to the camera motion. A number of methods have been proposed for detection of the moving targets in mobile camera including direct camera motion parameters estimation [3], optical flow [4], [5], and geometric transformation [6], [7]. Direct measurement of camera motion parameters is the best method for cancellation of the apparent background motion but in some application it is not possible to measure these parameters directly. Geometric transformation methods have low computation cost and are suitable for real-time purpose. In these methods, a uniform background motion is assumed. An affine motion model could be used to model this motion. When the apparent motion of the background is estimated, it can be exploited to locate moving objects.

In this paper we propose a new method for detection and tracking of moving targets using a mobile monocular camera. Our algorithm has two modes: detection and tracking. This paper is organized as follows. In Section 2, the detection procedure is discussed. Section 3 describes the tracking method. Experimental results are shown in Section 4 and conclusion appears in Section 5.

# 2 Target detection

In the detection mode we used affine transformation and LMedS (Least median squared) method for robust estimation of the apparent background motion. After the compensation of the background motion, we apply split and merge algorithm to the difference of current frame and the transformed previous frame to obtain an estimation of the target positions. If no target is found, then it means either there is no moving target in the scene or, the relative motion of the target is too small to be detected. In the latter case, it is possible to detect the target by adjusting the frame rate of the camera. The algorithm accomplishes this automatically by analyzing the proceeding frames until a major difference is detected.

We designed a voting method to verify the targets based on a-priori knowledge of the targets. For the case of vehicle detection we used vertical and horizontal gradients to locate interesting features as well as constraint on area of the target as discussed in this section.

## 2.1 Background motion estimation

Affine transformation [8] has been used to model motion of the camera. This model includes rotation, scaling and translation. 2-D affine transformation is described as follow:

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} a_5 \\ a_6 \end{bmatrix} \quad (1)$$

where (xi , yi ) are locations of points in the previous frame and (Xi , Yi ) are locations of points in the current frame and a1-a6 are motion parameters. This transformation has six parameters; therefore, three matching pairs are required to fully recover the motion. It is necessary to select the three points from the stationary background to assure an accurate model for camera motion. We used Moravec operator [9] to find distinguished feature points to ensure precise match. Moravec operator selects pixels with the maximum directional gradient in the min-max sense.

If the moving targets constitute a small area (i.e. less than 50%) of the image, then LMedS algorithm can be applied to determine the affine transformation parameters of the apparent background motion between two consecutive frames according to the following procedure.

1. Select N random feature point from previous frame, and use the standard normalized cross correlation method to locate the corresponding points in the current frame. Normalized correlation equation is given by:

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

here $\bar{f}_1$ and $\bar{f}_2$ are the average intensities of the pixels in the two regions being compared, and the summations are carried out over all pixels with in small windows centered on the feature points. The value *r* in the above equation measures the similarity between two regions and is between 1 and -1. Since it is assumed that moving objects are less than 50% of the whole image, therefore most of the N points will belong to the stationary background.

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 in the previous frame, and (Xi , Yi ) are their corresponds in current 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 select the affine parameters for which the median of squared difference is the minimum.

According to the above procedure, the probability p that at least one data set in the background and their correct corresponding points are obtained is derived from the following equation [7]:

$$p(\varepsilon, q, M) = 1 - (1 - ((1 - \varepsilon)q)^3)^M \quad (3)$$

where ε(<0.5) is the ratio of the moving object regions to whole image and q is the probability that corresponding points are correctly find.

In [7] it has been shown that the above method will give an accurate and reliable model.

## 2.2 Moving target detection using background motion compensated frames

When affine parameters are estimated, they can be used for cancellation of the apparent background motion, by transformation of previous frame. Now difference of the current frame and transformed previous frame reveals true moving targets. Then we apply a threshold to produce a binary image. The results of the transformation and segmentation are shown is figure 1-a and 1-b. Some parts are segmented as moving targets due to noise.

Connected component property can be applied to reduce errors due to noise. We use split and merge algorithm to find target bounding-boxes. If no target is found, then it means either there is no moving target in the scene or, the relative motion of the target is too small to be detected. In the latter case, it is possible to detect the target by adjusting the frame rate of the camera. The algorithm accomplishes this automatically by analyzing the proceeding frames until a target is detected. Our special interest is detection and tracking of the moving vehicles so we used aspect ratio and horizontal and vertical line as constraints to verify vehicles. Our experiments show that comparison of the length of horizontal and vertical lines in the target area with the perimeter of the target will give a good clue about the nature of the target.

# 3 Target tracking

After a target is verified, the algorithm switches into the tracking mode. Modified Moravec operator is applied to the target to identify feature points. These feature points are matched with points in the region of interest in the current frame. Disparity vectors are computed for the matched pairs of points. We used disparity vectors to refine the matched points. The refined points define the new position of the target in the current frame. The algorithm switches to the detection mode whenever the target is missed. Although the detection algorithm described above can be used for tracking too but the tracking algorithm, we describe in this section has very low computation cost in contrast with the detection algorithm described above. On the other hand when the target is detected it is not restricted to keep moving in tracking mode. The target can also be larger than 50% of the scenery in the tracking mode and this means camera can zoom to have a larger view of the target while tracking.

**Figure 1: two consecutive frames and difference of them after background motion compensation, the calculated affine parameters are: a1=0.9973, a2=-0.004, a3=0.008, a4=1.0022, a5=1.23, a6=-2.51**

When the size of the target is fixed the normalized cross-correlation or SSD (Sum of Squared Differences) method can be directly applied for target tracking. But we don't restrict the target to have a fixed size. Our tracking algorithm is capable of updating the target shape and size. To achieve this goal, the algorithm is based on dynamic feature points tracking. We select feature points from the target area and we track them in the next frame. Horizontal and vertical lines are important features for vehicle tracking. So we used optimized Moravec operator, which selects feature points, considering only horizontal and vertical gradients. This improves selection of interesting feature points located on the geometrically well-defined structures such as vehicles. This feature is very useful when dealing with occlusion. Our tracking algorithm consists of four steps, described as follow.

1. Apply the modified Moravec algorithm to select feature points in the target area in the previous frame.
2. Find corresponds of the feature points in the ROI (region of interest) of the current frame using normalized cross-correlation.
3. Calculate the disparity vectors and based on these vectors refine feature points. The refining is defined as omitting features with inconsistent vectors. This helps removal of non-target feature points.
4. Based on the location of refined corresponding points and the previous size of the target determine the location and size of the target in the current frame.

To refine the feature points described in the step 3, we calculated the mean and variance of the disparity vectors and based on these values we removed points, which have disparity vectors far from the mean disparity. In figure 2 the refined disparity vectors are shown for a target.

When there is no feature point for tracking or correspond points are not found properly, we assume that the target is lost so the algorithm switches to detection mode to find a target.

# 4 Results

The algorithm has been implemented on a Pentium III 500Mhz using a Visual C++ program. We have tested the algorithm with both simulated and actual sequences of images of vehicles in different landscapes. The system can detect and track targets in real-time. We achieved the frame rate of 4 frames/second for detection and 15 frames/second for tracking of 100*90 pixels target in 352*288 pixels video frames. We tested the proposed algorithm with wide variety of image sequence. Figure 3 shows some results for detection of various objects in arbitrary background. As it is shown the algorithm has successfully detected the targets. In figure 4 the tracking results for a vehicle are shown. In this example, the tracked vehicle turns over the road and its shape and size change. As it shown in the pictures both the size and the shape of the vehicle varies but our tracking algorithm can successfully track it.

Results showed the accuracy of the method in detecting and tracking of moving objects. Comparison of results generated by the proposed method with those of other methods showed that more reliable results could be obtained using the proposed method in real-time.

(a) (b) (c)

**Figure 2: The refined disparity vector used for tracking. (a) Previous frame, (b) Current frame and (c) Current frame with disparity vectors**

**Figure 3: various targets detected using the detection algorithm.**

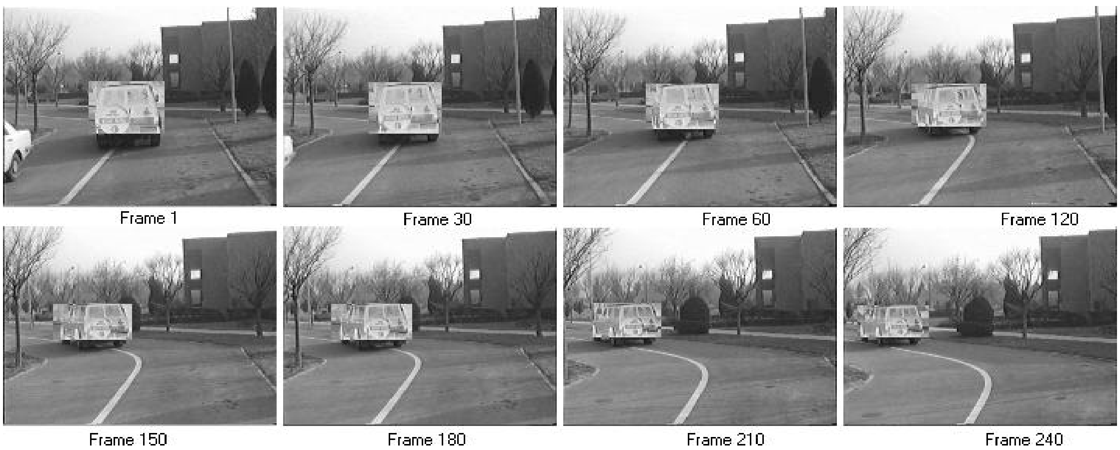


**Figure4: Tracking a vehicle while it turns over and changes its size and shape**

## 5 Conclusion

In this paper we proposed new method for detection and tracking of the moving objects using mobile camera. We used two different methods for detection and tracking. For detection, we used affine transform and LMedS method for estimation of the apparent background motion. When the apparent motion of the background is canceled, the difference of two consecutive frames is used for detection of the moving target. We also checked the target for some features.

We used modified Moravec operator and feature matching with disparity vectors refining for tracking. We tested our algorithms with wide variety of images sequence. The proposed methods successfully detect and track moving vehicles and objects in arbitrary scenes obtained from a mobile video camera. The tracking system is 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. Moreover, it is implemented in an understandable and structured way. Experimental results have shown that the algorithm is reliable and can successfully detect and track targets in most cases. To add more robustness to the tracking system it is possible to combine edge/texture based tracking methods with the current approach.

## 6 References


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

[2] Junghyun Hwang, Yoshiteru Ooi and Shinji Ozawa, "Visual feedback control system for tracking and zooming a target," *Proc. of the International Conference on Power Electronics and Motion Control* IEEE 1992.

[3] D. Murray and A. Basu, "Motion tracking with an active camera," *IEEE Trans. Pattern Anal. & Mach. Intell.*, vol. 16, no.5, 1995, pp. 449-459.

[4] 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,* 1992, pp. 236-242.

[5] Hans-Hellmut Nagel and Michael Haag, " Bias-corrected optical flow estimation for road vehicle tracking," *Sixth International Conference on Computer Vision (ICCV '98)*, Bombay/India, 4-7 January 1998.

[6] J. Odobez and P. Bouthemy, "Detection of multiple moving objects using multiscale MRF with camera motion compensation," *Proc. ICIP'94*, vol.2, 1994, pp. 257-261.

[7] Shoichi Araki, Takashi Matsuoaka, Naokazu Yokoya and Harua 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.

[8] E. Hayman, I.D. Reid and D.W. Murray, "Zooming while tracking using Affine transfer," *Proceedings of the British Machine Vision Conference* 1996.

[9] R. Jain, R. Kasturi and B.G. Schunck, *Machine vision*, McGraw-Hill Inc., 1995.

[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, October 1998.

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