

Presenting an Effective Algorithm for Tracking of Moving Object based on Support Vector Machine

Aref Altaf^{1*} and Arash Raeisi²

¹Information Technology Engineering (MIS), National Iranian South Oil Company (NISOC);
Aref.Altaf@gmail.com

²Information Technology Management, National Iranian South Oil Company (NISOC);
Ar_Raeisi@yahoo.com

Abstract

One of the most important and in progressing issue in image processing and machine vision is object tracking. In fact machine tracking is showing the position changing of the object and tracing it in a sequence of video image with a particular aim. Although the history of object tracking returned to military applications, but this object is used in various fields, so it considered by most researchers these days. In this paper an efficient algorithm for object tracking in video images using color and texture features and Support Vector Machine (SVM) is proposed. In the proposed algorithm, object's area is determined by user in the first frame at first, and then a same area is selected as field. If a part of the object is covered completely, Using SVM in the proposed algorithm cause that remains pixels be recognizable. Proposed algorithm is able to track the object yet there is partial covers and reach to suitable solutions.

Keywords: Image Processing, Machine Learning, Object Tracking, Video Images

1. Introduction

In recent years, machine vision one of the up-to-date branches of computer sciences has attracted lots of interests due to cheap and quality digital cameras and also increasing need for automatic video analysis. Through combination of methods related to image processing and machine learning tools, machine vision enables computer to intelligent comprehension of meaning and content of images. Object tracking, which is today under considerable levels of attention, is a basic operation for many high-level applications of machine vision such as recognition based on motion, automatic monitoring, indexing video files, human-computer mutual communications, traffic monitoring, and guidance of vehicles¹⁻⁴.

Many specialists in fields of electronics, telecommunications, computer, control and mathematics have presented various approaches for solving object tracking

problem. The most comprehensive research being conducted in ⁵, authors solved the point's correspondence problem with greedy approach and based on neighborhood and solid conditions. In ⁶, presented a greedy approach which has been bounded by steady estimation and the preliminary correspondences are obtained through evaluation of optical flow in the first two frames. In ⁷, the above works were improved by introduction of normal motions condition. In ⁸, authors formulated the point's correspondence problem as a graph theory problem. They defined matching of corresponding points in several successive frames as to be dependent on finding the best unique path for each point. Kalman filter has been frequently used for tracking in machine vision applications. In ⁹ Kalman filter has been used to track vehicles in a comparative framework. In ¹⁰, the developed Kalman filter was used to estimate the 3D motion path of object from motions of 2D image. One of the limitations

*Author for correspondence

of Kalman filter is the assumption that status variables have Gaussian distribution, so Kalman filter presents a weak estimation for status of those variables which do not follow Gaussian distribution. Generally, if the status is not assumed Gaussian, estimation of statuses can be done by particle filter¹¹. In ¹², particle filter has been used for tracking face and car. Template matching is one of the oldest and most common approaches of object tracking that has been used widely due to its relative simplicity and its date refers to Lucas_Kanade algorithm in 1981¹³. Other representations such as color histogram can be used instead of template using pixels inside rectangular or elliptical region and considered as object model.

In ¹⁴, authors made the object model by color mean and pixels inside rectangular region and searched the object in its 8 surrounding neighborhood to reduce computational complexity in next frame. In ¹⁵, to represent object, authors have used histogram of given weight being calculated from an elliptical region. They also used mean shift process^{16,17} for object locating instead of global search.

In ^{18,19}, authors considered object tracking problem as separation of object from its near background (the background which is located around object not the whole background). They calculated histograms of object region features and object's surrounding background using log likelihood ratio, tried to separate object being under tracking from its surrounding background in each frame, and then tracked the object by mean shift approach.

Author of paper²⁰ considered the object tracking problem as a classification problem of 2 classes and separated distributions of object and background features in each frame through constructing a time-varying discriminate function. Through training features of object region and its surrounding background, he formed a set of weak classifiers and then combined them to form object likelihood map. Furthermore, to find peak of likelihood map and thereby the object's new location, he used object mean shift approach. Weakness of this method is that various thresholdings have been used and also object changes have not been explicitly modeled so they must be known in advance to be used to determine thresholds. The biggest weakness of such approaches is that they are single-purpose which will limit their application in tracking a predetermined goal. The other weakness of this method is their need to an off-line learning process. In ²¹, correspondence of images has been conducted using

representation based on edge 1 Authors used Hausdroff distance to make a correlation level so that its minimum will be as object's new location. This method focuses on those parts of edges' plan which have not seriously modified by object motion so that the method improves tracking due to removal of edges having lots of motion.

Active contour is a shapeable curve which reveals a characteristic in the image, usually the smooth boarder of a region, through minimizing the defined energy for this contour. In 1987, this model was first presented in paper²². This method first faced many problems in terms of accuracy and speed but has been developed gradually.

Object revelation and tracking can be obtained by making a model of background, subtracting each entry frame from this model, and finding the object. Therefore, each great change in the image will be indicative of a moving object. This process is called background subtraction. Object region can also be obtained and tracked with the help of subtraction of those frames which are close to each other in terms of time²³. In order to learn gradual changes over time, paper²⁴ proposed that the color of each pixel of image $I(x,y)$ with a 3D Gaussian corresponding to colors of image (Y,U,V) become a statistical model. Further developments were done in this area by²⁵ using a mixture of Gaussians and updating their mean and covariance over time. In ²⁶, a mixture of illumination intensity and texture features has been also used for background subtraction.

2. The Proposed Method

In the proposed method, after selecting the desired object for tracking, a region as area as object and around it is considered as background. Then, color and texture features are extracted from these two regions, trained to Support Vector Machine (SVM) as features of object and background, and tested. The output of SVM will be an image in which the desired object has been distinguished from its surrounding background. Since the color and texture of image background will not remain the same over tracking, the trained SVM will not be able to distinguish object from background in all frames. To solve this problem, a concept called background development was defined by which all possible backgrounds are trained to SVM. For locating object, center of gravity of those pixels having been recognized as goal by SVM has been used along with mean shift approach.

2.1 Hypotheses of the Proposed Method

In the proposed method, at first, the desired object (which can be any specific object such as human being, car and etc.) is selected by user for tracking in the first frame using a rectangle. After that, it is expected that algorithm will track object as long as the object exists in the scene. In the presented method, 2D tracking of objects in a colorful video sequence is desired. A colorful video series is defined as a set of colorful images which are semantically dependent on and temporally adjacent to each other. The proposed method only needs the current frame and previous frames for tracking objects; therefore, theoretically has the capability of timely implementation. Besides, in this method, the focus is on presenting a general method for object tracking and simultaneously tracking of multiple objects has not been studied.

2.2 Steps of the Proposed Algorithm

2.2.1 Selecting Object and Background Regions

In the proposed method, at first, the desired object (which can be any specific object such as human being, car and etc.) is selected by user for tracking in the first frame using a rectangle. Second, a region approximately as area as object and around it is considered as background. Then, color and texture features are extracted from these two regions and trained to Support Vector Machine (SVM) as features of object and background. The reason why a region as area as object is selected around it is having almost equal number of object and background pixels to enhance performance of classifier.

2.2.2 Extracting Feature

Extraction of image features is one of the basic stages in image processing and pattern recognition. In the proposed method, we will use color and texture features for distinguishing object from its surrounding background. In the presented method, after preprocessing, color and texture features of object region and its surrounding background being selected by user are extracted and then these features will be taught to SVM as object and background features. SVM output will be a probability map of color and texture of object region and its surrounding background so that higher values of output indicate more probability that the pixel belongs to the object.

3. Simulation Results

To simulate the proposed algorithm, a computer having core i3 processor, 3GB memory and Microsoft Visual C#2010 programming software has been used. Furthermore, to evaluate the proposed method, the movement path obtained from proposed method has been compared to methods of frame correspondence (NCC), Mean Shift (MS¹⁵), and the real path having been extracted by expert.

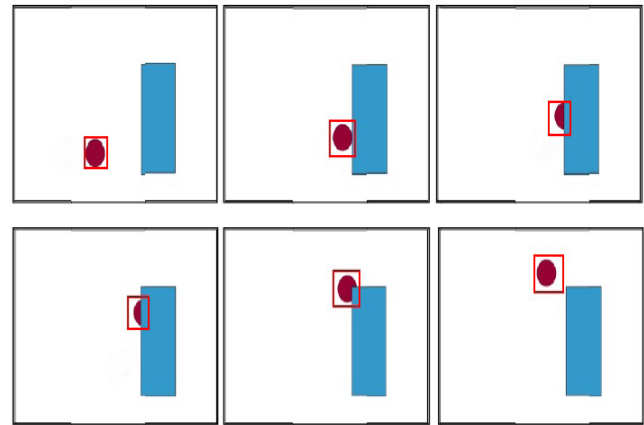


Figure 1. Results obtained for sequence NO.2 which indicates solid circular motion in red.

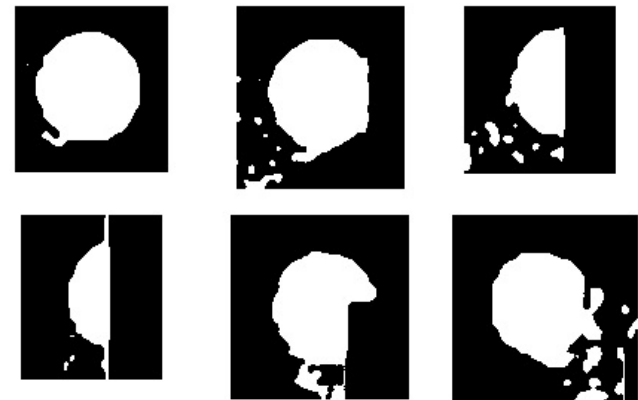


Figure 2. Confidence map of object frame and background regions obtained from proposed methods.

Photos from left to right and top to bottom respectively show tracking results for frames 10, 20, 25, 30, 35, and 40. In this sequence, the object is placed under slight

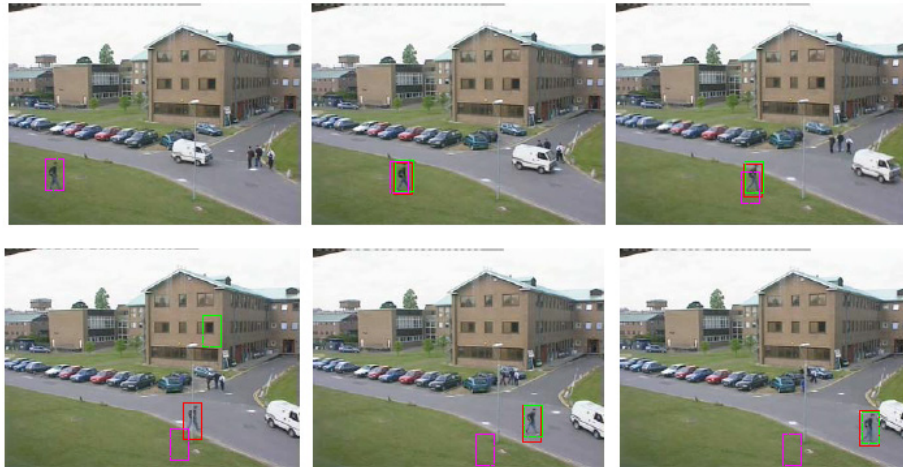


Figure 3. Results obtained for sequence NO. 3 that show a person is moving from left to right in the images.

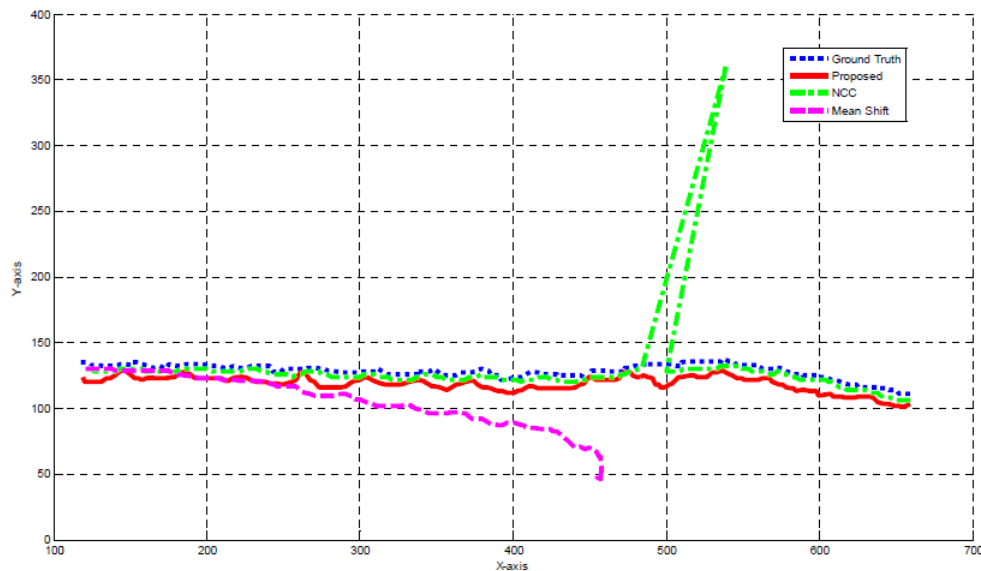


Figure 4. Comparing the real path of object motion (dotted line in blue) with the recognized path by the proposed method (solid line in red), frame correspondence method (NCC) (line-point in green), and Mean Shift method (MS) (dash line in pink) for second sequence.

latency in frames 20 to 37. The performance of proposed method has been shown by solid rectangle in red (Figure 1).

Photos from left to right and top to bottom and corresponding to Figure 1 show results obtained for frames 10, 20, 25, 30, 35 and 40 (Figure 2).

This sequence includes 200 frames (frames 860 to 1060 of the first scenario of test data in database PETS 2001) that the size of each video frame is 758×576.

Photos from left to right and top to bottom respectively show tracking results for frames 860, 900, 940, 990,

1030 and 1060 (Figure 3). In this sequence, object's surrounding background change and object is placed under slight latency. The slight latency of object is visible in frame 900. The proposed algorithm has been shown by solid rectangle in red, frame correspondence method (NCC) has been drawn by a green rectangle as line-point, and Mean Shift method (MS) has been illustrated by pink rectangle as dash line (Figure 4).

Mean shift method at first and gradually moves away from main path of object motion and. showing the tracking accuracy of proposed method has been shown

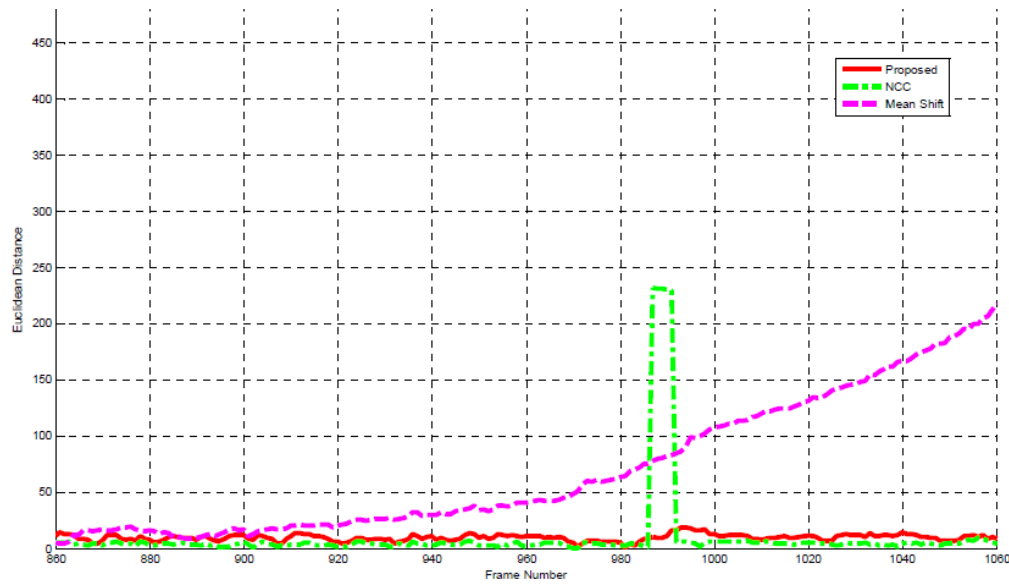


Figure 5. Show of tracking accuracy of the proposed method (solid line in red) in comparison with frame correspondence (NCC) (line-point in green), and Mean Shift (MS) (dash line in pink) for second sequence.

Table 1. Error mean obtained for the proposed method, frame correspondence method and mean shift method for the second sequence

Method name	Amount of Error mean during tracking (in pixel)
Proposed method	10.01
Frame correspondence (NCC)	10.06
Mean Shift (MS)	70.31

in Figure 5 in comparison with frame correspondence (NCC) and Mean Shift (MS) methods. The proposed method has a better performance in comparison with other methods. This is also observable in Table 1.

Table 1 shows error mean obtained for the proposed method, frame correspondence method, and mean shift for the second sequence. This error has been obtained through calculating Euclidean distance mean of real path by the proposed method, frame correspondence method, and mean shift (in pixel) along frames.

4. Studying and Comparing Algorithm Speed

Studying and comparing the tracking algorithms in terms of time depends on the way of implementing algorithm. In this section, it is tried to study correctly the time required for processing in each frame by the proposed algorithm. The main and time-consuming parts of proposed method have been shown in Table 2 for a selected object with dimensions 25×40 . Execution time of proposed method is different for the first frame and next ones. The first frame includes parts such as feature extraction, training and testing SVM, development of background, and retraining of SVM. Time-consuming parts for next frame include feature extraction and testing of SVM.

With the assumption of 4 repeats for calculating mean shift (to find object location in next frames), the time required for next frames will be 4 times more than time being displayed in Table 2 – namely 0.52 s.

Besides, time required for processing in each frame of proposed method and compared methods are available in Table 3. This time is related to a selected object with dimensions 25×40 and size of each frame is 480×640 . In the propose and mean shift methods, only the size of object window is effective but in the frame correspondence method, both the size of object window and size of

Table 2. The time (in second) obtained for calculations of first frame and next frames for a selected object with dimensions 25×40 in the proposed method

Algorithm/frame parts	First frame (second)	Next frames (second)
Feature extraction	0.12	0.12
Training and testing of SVM	0.72	-
development of background and retraining of SVM	0.77	-
Testing of SVM	-	0.01

Table 3. Processing time obtained for each frame in the proposed, frame correspondence, and mean shift methods

Method Name	Processing Time of Each Frame (in second)
Proposed Method	0.52
Frame Correspondence (NCC)	0.56
Mean Shift (MS)	0.63

image frame are effective. Times are expressed in second. For the proposed method, time required for execution of next frames (instead of first frame) has been considered as the time of processing each frame.

5. Conclusion

In this paper, an effective algorithm was presented for object tracking in video images using color and texture features and with the help of support vector machine. As it was expected, studying the results of experiments show that when untrained background pixels of next frames are tested by SVM, it properly distinguishes object from background. Besides, the proposed method tracks object carefully in all tested sequences up to the end. One advantage of using SVM and a pixel-based approach is that if a

part of object is under latency, pixels of remaining parts of object will still be recognizable. Another advantage of the proposed method is that the region occupied by object can often be extracted approximately in each frame and can be used in higher order processing such as recognition, interpretation and explanation of the type of object behavior and etc.

The possibility of using multiple features in the framework of proposed method can be named as another benefit, so that this method does not have weakness of histogram-based methods that is their inability to use lots of features.

As shown, the proposed method is able to track object even though it has very slight latency. But for tracking object having complete latency, the proposed method must be combined with a prediction method such as Kalman filter or particle filters that is suggested for future works.

6. References

1. Yilmaz A, Javed O, Shah M. Object tracking: A survey. *ACM Computing Surveys*. 2006 Dec; 38(4):1–45.
2. Baghyasree T, Janakiraman K, Parkavi D. Efficient remote video monitoring for industrial environment based on LPC2148 and GSM. *Indian Journal of Science and Technology*. 2014; 7(9):1333–41.
3. Kheirkhah E, Sadri Tabatabaie Z. A hybrid face detection approach in color images with complex background. *Indian Journal of Science and Technology*. 2015; 8(1):49–60.

4. Ghasemi Naraghi M, Koohi M, Sarbazy H. Object tracking based on estimation mesh method. *Indian Journal of Science and Technology*. 2012; 5(1):1851–4.
5. Sethi I, Jain R. Finding trajectories of feature points in a monocular image sequence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1987; 9(1):56–73.
6. Rangarajan K, Shah M. Establishing motion correspondence. *Proceeding of Conference Vision, Graphics Image Process Image Understanding (CVGIP)*. 1991; 54(1):103–8.
7. Veenman C, Reinders M, Backer E. Resolving motion correspondence for densely moving points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2001; 23(1):54–72.
8. Shafique K, Shah M. A non-iterative greedy algorithm for multi-frame point correspondence. *Proceeding of IEEE International Conference On Computer Vision (ICCV)*; 2003. p. 110–15.
9. Boykov Y, Huttenlocher D. Adaptive bayesian recognition in tracking rigid objects. *Proceeding of IEEE Conference Computer Vision and Pattern Recognition*; 2000. p. 697–704.
10. Rosales R, Sclaroff S. 3D trajectory recovery for tracking multiple objects and trajectory guided recognition of actions. *Proceeding of IEEE Conference Computer Vision and Pattern Recognition*; 1999. p. 117–23.
11. Kitagawa G. Non-Gaussian state-space modeling of non-stationary time series. *Journal of the American Statistical Association*. 1987; 8(400):1032–63.
12. Li B, Chellappa R. Simultaneous tracking and verification via sequential posterior estimation. *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*. 2000; 2:110–17.
13. Lucas B D, Kanade T. An iterative image registration technique with an application to stereo vision. *Proceeding of International Joint Conference on Artificial Intelligence*; 1981. p. 674–79.
14. Fieguth P, Terzopoulos D. Color-Based tracking of heads and other mobile objects at video frame rates. *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 1997. p. 21–7.
15. Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2003; 25(5):564–75.
16. Comaniciu D, Meer P. Mean shift analysis and applications. *Proceeding of IEEE International Conference on Computer Vision (ICCV)*. 1999; 2:1197–1203.
17. Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24(5):603–19.
18. Collins R, Liu Y. On-Line selection of discriminative tracking features. *Proceeding of ICCV*; 2003. p. 346–52.
19. Collins R, Liu Y, Leordeanu M. On-Line selection of discriminative tracking features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2005 Oct; 27(10):1631–43.
20. Avidan S. Ensemble tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2007; 29(2):261–71.
21. Huttenlocher D, Noh J, Rucklidge W. Tracking nonrigid objects in complex scenes. *Proceeding of IEEE International Conference on Computer Vision (ICCV)*; 1993. p. 93–101.
22. Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. *Proceeding of 1st International Conference on Computer Vision*. 1987; 1(4):321–31.
23. Haritaoglu I, Harwood D, Davis L W4: real time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000; 22(8):809–30.
24. Wren C, Azarbayejani A, Pentland A. Pfunder: Real-Time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1997; 19(7):780–85.
25. Stauffer C, Grimson W. Learning patterns of activity using real time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000; 22(8):747–67.
26. Liyuan L, Maylor L. Integrating intensity and texture differences for robust change detection. *IEEE Transactions on Image Processing*. 2002; 11(2):105–12.