Multiple Object Tracking using Support Vector Machine

G Ramya[1] and Mrs Srilatha[2]

[1] vardhaman college of engineering

Received: 10 December 2013 Accepted: 31 December 2013 Published: 15 January 2014

Abstract

This paper presents an accurate and flexible method for robust recognition and tracking of multiple objects in video sequence. Object tracking is the process of separating the moving object from the video sequences. Tracking is essentially a matching problem in object tracking. In order to avoid this matching problem, object recognition is done on the tracked object. Background separation algorithm separate moving object from the background based on white and black pixels. Support Vector Machines classifier is used to recognize the tracked object. SVM classifier are supervised learning that associates with machine learning algorithm that analyse and recognize the data used for classification. SVM uses Kalman filter which makes the system more robust by tracking and reduce the noise introduced by inaccurate detections.

Index terms— object tracking, background subtraction, SVM, kalman filter, fuzzy.

1 Introduction

ormally all working environments need security. Security can be implemented [1] in many ways, sometimes audio, video or by any other means. Video surveillance systems are most common today. Intelligent video surveillance systems deal with the realtime monitoring of persistent and transient objects within a specific environment. This type of video surveillances can be applied not only to various security systems, but also for environmental surveillance. This surveillance can be used for many other purposes like event detection, [2] visual surveillance and robotics. A normal object detection algorithm can be applied for this purpose, but it may be difficult to detect unknown objects with significant changes in color, shape and texture. So most surveillance systems use static cameras which make the object detection much more easy. In such cases a background model is trained with data obtained from empty scenes and foreground regions are identified using the dissimilarity between the trained model and new observations. This method is normally used in all static cameras. Difficulties in tracking objects can arise due to abrupt object motion, [3] changing appearance patterns of both the object and the scene, non-rigid object structures, object-to-object and object-to-scene occlusions, and camera motion. Tracking is usually performed in the context of higher level applications that require the location and/or shape of the object in every frame.

The rest of this paper is organized as follows. Section II is a brief review of existing method. The proposed method is given in section III. Simulation results are given in section IV. Section V includes conclusion.

2 II.

3 EXISTING METHOD

A In the literature, many features are used for the detection of moving objects. In the case of color features, [4]some authors make the foreground detection in each dimension independently and then aggregate the corresponding foreground mask using the binary operator (OR). The disadvantage is that a false positive in one dimension generate false positive in the final result.

We propose thus to use a fuzzy operator i.e. the Choquet integral to aggregate the results obtained in each dimension to avoid crisp decision. fuzzy aggregation is the most promising, as it considers the correlation [5] existing among the features by fuzzy measure. Fuzzy Logic is a problem-solving control system methodology that lends itself to implementation in systems ranging from simple, small, embedded micro-controllers to large, networked, multi-channel PC or workstation-based data acquisition and control systems.

Fuzzy technique can be used to perform the accurate segmentation of the foreground object. fuzzy logic system, which fuses multiple sources of information together for decision making. In fuzzy technique the following are the drawbacks Using [6] more quantization levels increases the co-occurrence matrix size, thereby increasing computational complexity. Using less number of levels results in losing the information content in the frame. Thresholds were chosen empirically for all the competing [7] approaches so that their best results are used for comparison. High value of threshold results in more pixels getting falsely classified as foreground and vice versa.

4 III.

5 Proposed Method

The capability of multiple objects tracking [8] is essential for most surveillance systems. One major challenge of this requirement is to detect multiple objects with occlusion. In this paper, the tracking algorithm combines feature matching model and Kalman filter framework to resolve this problem. The real-time visual surveillance system detects and tracks multiple people and monitors their activities in an outdoor scene [9] using a single monocular grayscale or IR camera. It uses a combination of shape analysis and tracking to locate people and their parts, and to create models of their appearance such that they can be tracked through interactions as occlusions. The method uses the results of an iterative and dynamic probabilistic approach for object recognition. The object of interest is initialized by a user-specified bounding box, [10] but its category is not provided. Meanwhile, [11] video-based object recognition is applied on the tracked objects. When the target is recognized properly, the offline target model will be automatically incorporated to provide more information about the target. So target measurement at time t is) (t t t x I Z =

where It is the input image at time t. Xt is the target state. Frame splitting is the process of splitting the given video into number of frames. In one second 24 frames are generated. Frames are stored in a separate file. Alternate Frame Rendering (AFR) is a technique of graphics [12] rendering in personal computers which combines the work output of two or more graphics processing units (GPU) for a single monitor, in order to improve image quality, or to accelerate the rendering performance. Extracting the background image from sequences of frames is a very important task in order to help tracker detect motion. This task is repeated from time to time in order to incorporate any changes in the illumination of the tracking scene, Here we are using the first frame as background image.

Background subtraction, [13] also known as Foreground Detection, is a technique in the fields of image processing and computer vision wherein an image's foreground is extracted for further Processing. A classification task usually involves with training and testing data which consist of some data instances. Each instance in the training set contains one "target value" (class labels) and several "attributes" (features) . The goal of SVM [10] is to produce a model which predicts target value of data instances in the testing set which are given only the attributes. In this work, moment features of detected objects are calculated from the first twenty frames of the video sequences and are stored in a database . Then we use these feature data to train the SVM. In the subsequent video sequences, the corresponding features extracted from each detected object in each frame are sent to the trained SVM. The SVM can accurately predict the class label of each object; consequently the system can recognize each object correctly.

The moment features provide some robustness to occlusion and background clutter. Color Moments and Wavelet Moments, in object recognition system, typically a set of numerical features are extracted from an image.

The purpose of feature extraction technique in image processing is to represent the image in its compact and unique form of single values or vectors. When the Kalman filter tracking system has a reliable prediction on the target movement, feature matching can be carried out in the relatively small region thus[13] Kalman filter plays an extremely important role in improving processing speed and performance of the tracking system. With the aim of finding the corresponding relationship in this system, we propose a Kalman tracking method based on object recognition. In visual surveillance system, Kalman filter is widely used for target prediction and tracking. After setting various appropriate parameters, we initialize the Kalman filter by the object segmentation results of the first two frames.

IV.

6 Simulations Results

7 Comparison between Existing and Proposed Method

Th e performance of the prediction was evaluated in terms of recall, precision and accuracy.

8 Conclusion

In this paper we have proposed a robust and practical multiple objects recognition and tracking method. We integrate color moments and wavelet moments together for recognition and tracking. Moreover, we utilize a Kalman Filter framework to assist in tracking multiple objects. Our experimental results show that the proposed method can accurately recognize and robustly track multiple objects. Object recognition module improves the performance and accuracy of the Kalman filter tracking framework simultaneously the Kalman filter tracking framework can greatly improve the tracking speed. Furthermore, the detection of multiple moving objects with occlusion is successfully finished which is a problem for recognition and tracking based on feature matching. The combination of these two modules nicely compensates for the weaknesses of each individually. The major limitation of this method is that for each object there have to be twenty frames for SVM training. For future work, we plan to introduce multimodal features to represent objects so that the recognition and tracking performance can be further improved. [1]

1

Figure 1: Fig. 1 :

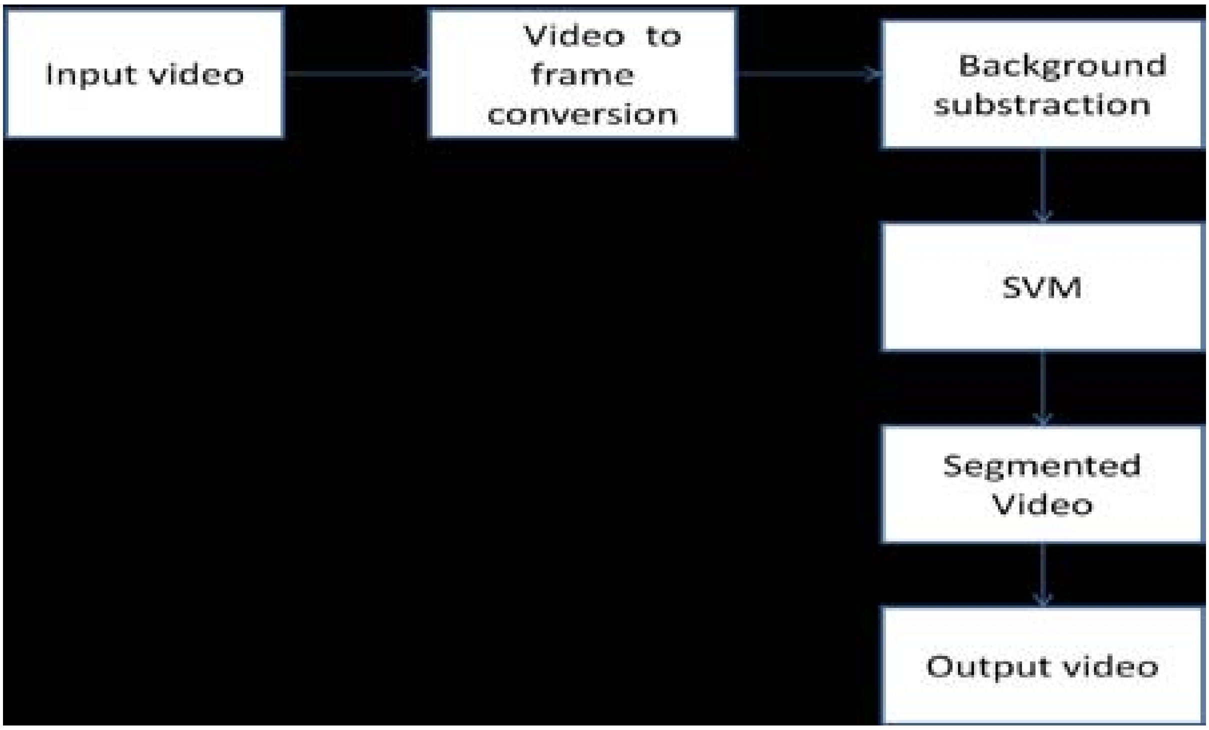

Figure 2:

24

Figure 3: Fig. 2 :Fig. 4 :

1

SVM using recall

Figure 4: Table 1 :

2

SVM using prescision		
Number of	Precision	
frames	Fuzzy	SVM
1	89	95
2	90	93
3	89	95
4	90	92
5	89	92
6	90	92
7	90	95
8	95	98
9	89	92
NUMBER OF	ACCURACY	
FRAMES	FUZZY	SVM
1	91	99
2	82	90
3	95	98
4	95	99
5	90	95
6	90	92
7	92	95
8	95	98
9	90	99

Figure 5: Table 2 :

3

Recall TP/TP+FP
Precision TP/TP+FN
Accuracy TP+TN/TP+TN+FP+FN

[Note: JVI.]

Figure 6: Table 3 :

[Baf et al. ()] 'A fuzzy approach for background subtraction'. F El Baf , T Bouwmans , B Vachon . *Proc. ICIP*, (ICIP) 2008. p. .

[Sun and Guo (2008)] 'A Robust Object Detecting and Tracking Method'. W Sun , B L Guo . *5th International Conference on Fuzzy Systems and Knowledge Discovery*, Oct. 2008. 4 p. .

[Jian et al. ()] 'Background subtraction based on a combination of texture, color and intensity'. X Jian , D Xiao-Qing , W Sheng-Jin , W Youshou . *Proc. IEEE 9th ICSP*, (IEEE 9th ICSP) 2008. p. .

[Cristani et al. ()] 'Background subtraction for automated multisensor surveillance: A comprehensive review'. M Cristani , M Farenzena , D Bloisi , V Murino . *EURASIP J. Adv. Signal Process* 2010. 2010 p. .

[Grabner et al. ()] 'Learning Adaptive Metric for Robust visual tracking'. H Grabner , M Grabner , H Bischof . *Proc. British Mach. Vis. Conf*, A D Jepson, D Fleet, T El-Maraghi (ed.) (British Mach. Vis. Conf) 2006. p. . (Real-time tracking via on-line boosting)

[Marcenaro et al. (2002)] 'Multiple objects tracking under heavy occlusions by using Kalman filters based on shape matching'. L Marcenaro , M Ferrari , L Marchesotti , C S Reqazzoni . *Proceedings of International Conference on Image Processing*, (International Conference on Image ProcessingRochester, NY, United States) June 2002. 3 p. .

[Chiranjeevi and Sengupta ()] 'New fuzzy texture features for robust detection of moving objects'. P Chiranjeevi , S Sengupta . *IEEE Signal Process. Lett* 2012. 19 (10) p. .

[Wren et al. (1997)] 'Pfinder: real-time tracking of the human body'. C R Wren , A J Azarbayejani , T J Darrell , A P Pentland . *IEEE Trans Pattern Anal Mach Intell* July. 1997. 19 p. .

[Boser et al. (1992)] 'Training algorithm for optimal margin classifiers'. B E Boser , I M Guyon , V N Vapnik . *Proceedings ofthe Fifth Annual ACM Workshop on Computational Learning Theory* July 1992. p. .

[Choi and Yang (2009)] 'Vehicle detection from aerial images usin local shape information,? Adv'. J Y Choi , Y K Yang . *Image Video Technol* Jan. 2009. 5414 p. .

[Zhou et al. (2004)] 'Visual tracking and recognition using appearance-adaptive models in particle filters'. S Zhou , R Chellappa , B Moghaddam . *IEEE Trans. Image Process* Nov. 2004. 13 (11) p. .

[What Are We Tracking: A Unified Approach of Tracking and Recognition Jialue Fan, Xiaohui Shen, Student Member (2003)] 'What Are We Tracking: A Unified Approach of Tracking and Recognition'. *Jialue Fan, Xiaohui Shen, Student Member* Oct. 2003. 11. IEEE. 25 (10) p. . (Mach. Intell.)