International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Multi Variant Feature Similarity Based Behavior Tracking in Video Surveillance using Ann

Shana L, C Seldev Christopher

Abstract: ***The problem of video surveillance has been well studied which has been adapted for several issues. The behavior of any human can be monitored through video surveillance. There are number of approaches available for the video surveillance and behavior analysis. The previous methods uses background models, object tracking for the problem of behavior analysis. The methods suffer with poor accuracy in behavior analysis. To improve the performance, a multi variant feature similarity model based behavior tracking in video surveillance is presented. The method involves in identifying interest points throughout the images of video. Second, the changing feature has been identified to measure the multi variant feature similarity by using multi variant feature model. Based on the MVFS, the object tracking is performed. The human tracking is performed in the same way and the multi variant features are trained with artificial neural network which has number of behavior classes. At the testing phase, the video has been removed with background features according to the multi feature model adapted. Once the object has been identified, then tracking and behavior analysis is performed by measuring MVFS with the features at different behavior classes. The artificial neural network has been used for the classification of behavior identified through video surveillance. The method would produce higher accuracy and improves the performance.***

Index Terms**: *Video Surveillance, Object Detection, Object Tracking, Behavior Tracking, ANN, MVFS.***

I. INTRODUCTION

The video surveillance is the process of monitoring the happenings in any indoor or outdoor location. Many organizations have engaged video surveillance systems for different purposes like security management, employee monitoring and so on. By adapting video surveillance systems, the administrator would be able to monitor the happenings and events in the location and can monitor the activities of any human. Similarly, the behavior of any human can be monitored and tracked using video surveillance systems. There are number of approaches available for video surveillance like object and template based models.

Manuscript published on 30 September 2019
* Correspondence Author
Shana L*, Computer Science and Engineering, St. Xavier's College of Engineering, Anna University, Nagercoil, India.
Prof Dr. C Seldev Christopher, Computer Science and Engineering, St. Xavier's College of Engineering, Anna University, Nagercoil, India

The template models maintain list of templates for different activities based on that the activities has been performed. Similarly, in object based models, the method uses list of objects and their features in identifying the activity or behavior of the user. Similar to that, number of scientific approaches available for behavior analysis.

The k-means algorithm would estimate the distance measure on different objects of specific behavior to identify the class of activity. The accuracy of the k-means algorithm is depending on the feature being considered and the number of samples available in each class. Also, the accuracy depends on the variant features considered. The support vector machine has been used for different problems and can be used for the task of behavior analysis in video surveillance. Whatever the algorithm being used, the accuracy is highly depending on the feature considered and the methods of similarity estimation.

The process of behavior analysis from video surveillance is starts with object tracking. The video would contain number of objects in the sequence of frames. In order to perform behavior tracking, it is necessary to identify the objects present in the video frames. Once the objects of the frame have been detected then behavior can be identified. The objects can be identified by maintaining the templates of the objects of various activities. The template matching can be performed to identify the objects. Similarly, the background features plays vital role in behavior analysis. By removing the background features, the objects can be identified efficiently. The background features has been eliminated by using different background models but produces poor accuracy in removing the background features.

By considering all these, a multi variant feature analysis model has been presented in this paper. The method uses multiple features of variant form to perform behavior analysis model. The artificial neural network has been used for classification. The ANN has great impact in various scientific problems and the same can be used for behavior analysis in video surveillance. The Multi Variant Feature Similarity has been used by the neurons of the network for behavior analysis. The brief description of the method is presented in section 3. To eliminate the background features various segmentation algorithms available. Each algorithm would use different background models. Similarly, to identify the foreground objects, different features have been used. The objects sketch

II. RELATED WORKS

The problem of behavior analysis has been approached with different methods. Such methods are discussed and reviewed in this section. In [1], the author presented a detailed review in behavior understanding in surveillance videos. The author details various human activities and how it can be identified through surveillance video. Also, different methods have been studied and compared their performance methods and measures in detail.

In [2], the object tracking has been performed with Bayesian model. The method uses distance measures of objects in a structured environment. The method also used the packet filtering in object tracking.

To track multiple objects in surveillance video in online stream, a Bayesian model with top down approach is presented in [3]. The method focused on labeling multiple objects in video stream. The method adapts the occlusion features and finite sets of the multiple objects. The Bayesian model is approached recursively to track the objects of the video.

The correlation and dependency measures on feature for efficient multi label classification are presented in [4]. It is a Bayesian model which learns multiple class data streams. It measures the correlation between the streams and measures the relationship between them. According to the correlation measure and dependency measures, the classification is performed.

An HMM based behavior tracking model is presented in [5]. The model is a bottom up approach which transforms the image features into vector sequences. The vector sequences generated has been used to quantize the samples to perform classification on behaviors.

In DFP-ALC [6], the method extracts the patches of distinct frames. Such patches are indexed based on appearance based clustering (ALC). In classification, exact class is selected based on appearance measures.

Depth matrix based gesture recognition is presented in [7], which extract the image features and convert them into shape matrix from the posture sequences. Using the sequences, he method extracts the regional descriptors and classify them using Naïve Bayes Classifier.

The problem of recognizing the behavior of pedestrian, an reciprocal model to guide the pedestrian is presented in [8]. The method uses image and attribute features to measure the distribution weight. Estimated weight measure has been used to perform classification.

The behavior tracking has been approached with skeleton features in [9]. The method extracts the skeleton prototypes based on the shape features. Various pose of skeleton features are extracted and has been used to perform classification.

The problem of assessing autism spectrum disorder is handled with the pattern of robot and human interaction is presented in [10]. The method captures the behavior of children through RGB-D sensors. The method measure the behavior of the children based on the movements of body and head. Also the features of direction, magnitude of gazing's and energy of kinetic are used for classification.

In [11], the method segments the parametric and actions of primitives. The method combines the actions in to number of sequences and based on the similarity of motion sequences the classification is performed.

A knowledge based approach for controlling video surveillance has been presented in [12]. The model has leaned features of classes according to the features considered. The method performs classification according to the features available.

The above discussed algorithms suffer to achieve higher performance in behavior tracking and produces higher false classification ratio.

III. MVFS BEHAVIOR TRACKING WITH ANN

The proposed multi variant feature similarity measure based behavior tracking algorithm reads the input video. A set of video frames are generated and for each image the method generates interest points where huge features are available. Based on the interest points on different video frames, the method estimates multi variant feature similarity between the interest points. Based on the value of MVFS, the method identifies the list of objects and eliminates them to perform background subtraction. In the same way, the method identifies list of varying human features to perform behavior analysis. The classification is performed using ANN. The detailed approach is presented below:

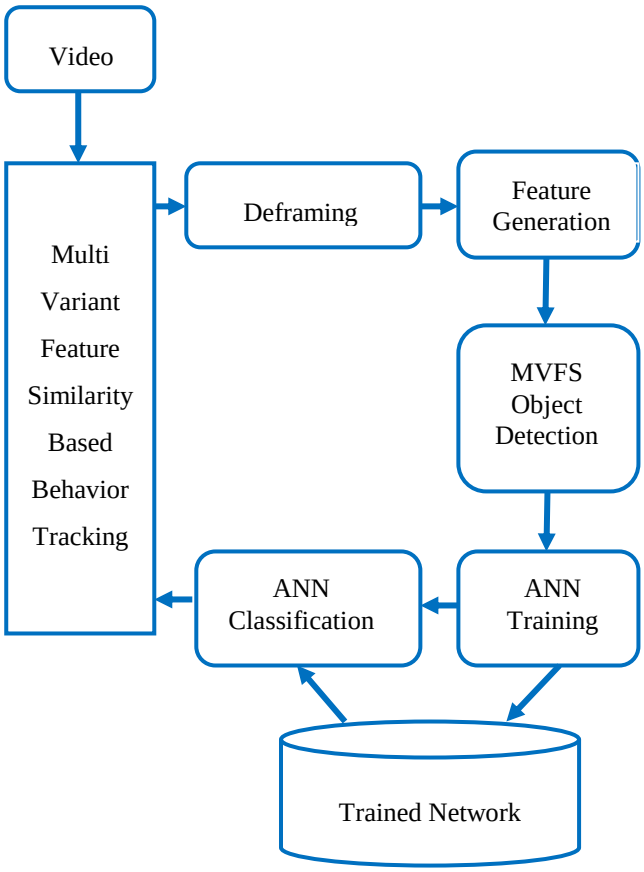

Figure 1: General Architecture of Proposed MVFS Behavior Analysis Model

The Figure 1, present the architecture of proposed MVFS behavior analysis model and shows various functional components in detail.

3.1 De-framing /Feature Extraction:

The input video being received and the video have been split into number of frames. For each second of the video, the method generates number of frames. From each frame, the method generates interest points. The interest points are identified based on the availability of maximum features. The Speed up Robust Feature Based approach has been used to generate interest points. The method generates hessian matrix and based on that the list of interest points are identified.
Identified interest points are extracted and converted into feature vector. Generated feature vector has been used to perform behavior tracking.
Algorithm:

Input: Video v

Output: Fd

Start

Read input video v.

For each second s

Generate image frame

imf = $\int_{i=1}^{size(v)} \sum Frames \in v(i)$

End

For each frame f

Generate Integral image

Iimg = $\int \sum IntegralImage(f, b)$

B – Window size.

For each integral image i

H = Generate Hessian Matrix ($Iimg_i$)

Convert feature descriptor Fd.

Add to Fd = $\sum(FeatureDescriptor \in Fd) \cup fd$

End

End

Stop

The above discussed algorithm extracts the frames from the video and generates feature descriptor based on interest points identified.

3.2 MVFS Object Detection

In this stage, the input feature set has been read belongs to different frames. For each feature descriptor of images, the method estimates multi variant feature similarity (MVFS) towards the feature descriptors of other or next images feature set. The MVFS measure has been measured between the feature set of different frames. Based on the variance, the method identifies the background objects and human objects. Identified objects set has been used to perform behavior analysis.
Algorithm:

Input: Feature Descriptors Set Fds.

Output: Object set Os, Human Set Hs.

Start

Read feature descriptor set Fds.

For each feature set fs

For each feature descriptor fd

For each feature set of other image Fos

For each feature descriptor fod

Compute MVFS = $\frac{\sum_{i=1}^{size(fd(fs))} fd(fs(i)) \equiv fod(fos(i)) / size(fd(fs))}{size(fd(fs))}$

End

End

If MVFS>Th then

Add to Hs = $\sum(fd \in Hs) \cup Fd$

Else

Add to Os = $\sum(fd \in Os) \cup Fd$

End

End

End

The above discussed algorithm briefs how the objects of the image has been identified and split into background and foreground objects.

3.3 MVFS-ANN Training

In this stage, the method reads the input video data set. For each video data set, the method extracts the MVFS feature. Extracted objects and human feature set has been used to generate number of neurons and initialize them with the features extracted. The method has classified number of behaviors. For each behavior different layer of neurons has been initialized and trained. The trained neural network has been used to perform behavior analysis.

Algorithm:

Input: Video Set Vs.

Output: ANN

Start

Read input video set vs.

Initialize neural network ANN.

For each video v

Fds = Feature extraction (Frame set Fs)

[Os, Hs] = MVFS-ObjectDetection(Fds)

Generate Neuron N and Initialize with Hs.

End

Stop

The above discussed algorithm generates number of neurons and initializes them with the features of objects. Generated neural network has been used to perform behavior analysis.

3.4 MVFS-ANN Behavior Classification

In this stage, the method reads the input test video and the network being trained earlier. From the video, the method generates the frame set and extracts the features. Extracted features has been feed with ANN generated, where each neuron estimates MVFS measure on each class of behavior features. Based on the outcome of MVFS measure, the method estimates the MVFS behavior weight. Using MVFS-BW, the method classifies the behavior of the human to produce result.

Algorithm:

Input: Test video Tv

Output: Behavior B.

Start

Read input video Tv.

Fds = Deframing-Feature Extraction (TV)

[Os,Hs] = MVFS-FeatureExtraction(Fds)

[MVFSs] = Apply Neural Network on Hs.

For each class of behavior B

Compute MVFS-W = $\frac{\sum MVFSs(B)}{No\ of\ Neurons} \times \frac{\sum MVFSs(B)>Th}{No\ of\ Neurons}$

End

B = choose the class with maximum MVFS-W value.

Stop

The above discussed algorithm estimates behavior weight based on the result produced by the neural network on the features extracted. Based on the value of behavior weight a single one has been selected.

IV. RESULTS AND DISCUSSION

The proposed multi variant feature similarity based behavior analysis model on video surveillance has been implemented using MatLab. The proposed approach has been measured for its efficiency in different parameters. The results produced by the proposed method have been presented below.

Table 1: Details of evaluation

Parameter	Value
Data Set	VIRAT 2.0
Number of Videos	6
Number of Behaviors	10
Tool	MatLab

The Table 1, present the details of data set being used for the evaluation of the proposed algorithm. According to the details of data set, the method has been measured for its efficiency in the following parameters. The VIRAT data set maintains 10 numbers of behaviors or activity and for each activity 6 number of videos available.

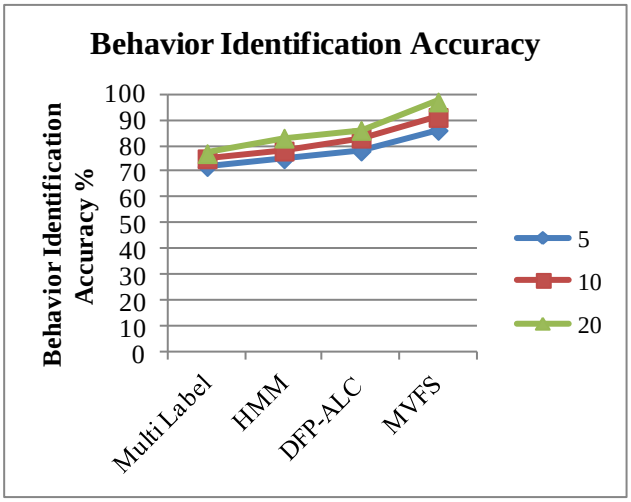

Figure 2: Comparison on behavior identification accuracy

The performance on behavior identification has been measured and compared with the result of previous algorithms. The proposed MVFS algorithm has produced less efficiency than other methods.

The performance on false classification ratio produced by different methods has been measured. The MVFS algorithm reduced the ratio of false classification compare to other methods.

The performance on time complexity has been measured and presented in Figure 1. The proposed MVFS algorithm has produced less time complexity than other methods.

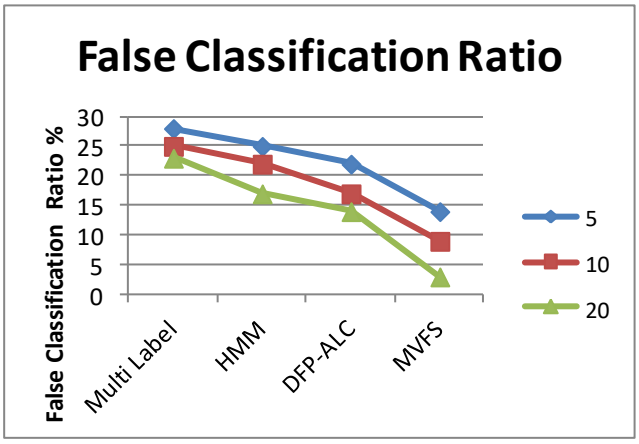

Figure 3: Comparison on false classification ratio

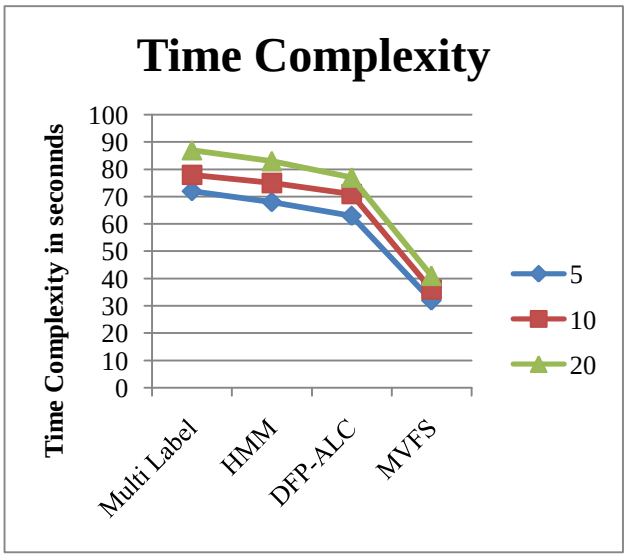

Figure 4: Comparison on time complexity

V. CONCLUSION

In this paper, an efficient multi variant feature similarity based behavior analysis model has been presented. The method reads the input video set and extracts the features of interest points. Extracted features have been used to estimate the MVFS measure to identify the feature of human and other background objects. Identified feature set has been used to generate the neural network and initialize them. Second, the test video has been used to extract the features and tested with neural network. Based on the result of ANN, the method estimates the MVFS behavior weight MVFS-BW. Based on the weight estimated, the method identifies the behavior class as result. The proposed method improves the performance of behavior identification and classification with false classification ratio with less time complexity.

REFERENCES

1. Sarvesh Vishwakarma, Anupam Agrawal, A Survey on Activity Recognition and Behavior Understanding in Video Surveillance, Springer, Visual Computing, Volume 29, pp 983-1009, September 2012.
2. Junda Zhu, Object Tracking in Structured Environments for Video Surveillance Applications, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 20, No. 2, February 2010.
3. Du Yong Kim, Ba Ngu Vo, A Labeled Random Finite Set Online Multi-Object Tracker for Video Data, Elsevier, Pattern Recognition, Volume 90, pp 377-389, June 2019.
4. Tien Thanh Nguyen, Thi Thu Thuy Nguyen, Multi-Label Classification via Label Correlation and First Order Feature Dependence in a Data Stream, Elsevier, Pattern Recognition, Volume 90, pp 35-51, June 2019.
5. Jungi Yamato, Recognizing Human Behavior Using Hidden Markov Models, Springer, Video Computing, pp 90-131, 2002.
6. Sivapriyaa Kannappan, Yonghuai Liu, Bernard Tiddeman, DFP-ALC: Automatic Video Summarization Using Distinct Frame Patch Index And Appearance Based Linear Clustering, Elsevier, Pattern Recognition, Vol.120, pp: 8-16, April 2019.
7. Lalit Kane, Pritee Khanna, Depth Matrix And Adaptive Bayes Classifier Based Dynamic Hand Gesture Recognition, Elsevier, Pattern Recognition, Vol. 120, pp 24-30, April 2019.
8. Zhongji, Erlu He, Haoran Wnag, Image-Attribute Reciprocally Guided Attention Network For Pedestrian Attribute Recognition, Elsevier, Pattern Recognition, Vol 120, pp 89-95, April 2019.
9. Alessia Saggese, Nicola Strisciuglio Learning Skeleton Representations for Human Action Recognition, Elsevier, Pattern Recognition, Vol. 118, pp 23-32, February 2019.
10. Salvatore Maria Anzalone, Jean Xavier, Quantifying Patterns Of Joint Attention During Human-Robot Interactions: An Application For Autism Spectrum Disorder Assessment, Elsevier, Pattern Recognition, Vol. 118, pp 42-50, February 2019.
11. Juan Pedro Bandera, Rebeca Marfil, A New Paradigm For Autonomous Human Motion Description and Evaluation: Application to the Get Up& Go Test Use Case, Elsevier, Pattern Recognition, Vol. 118, pp 51-60, February 2019.
12. B. Georis, F. Brémond, M. Thonnat, Real-Time Control of Video Surveillance Systems With Program Supervision Techniques, Springer, Machine Vision And Applications, Volume 18, Issue 3–4, Pp 189–205, August 2007.
13. Angelica Lozano, Giuseppe Manfredi, Luciano Nieddu, An Algorithm for the Recognition of Levels of Congestion in Road Traffic Problems. Springer, Mathematics and Computers in Simulation, Volume 79, pp 1926-1934, February 2009.
14. Jie Zhou, Dashan Gao, David Zhang, Moving Vehicle Detection for Automatic Traffic Monitoring. IEEE Transactions on Vehicle Technology, Volume 56, pp 51-59, January 2007.
15. Xutong Niu, A Semi-Automatic Framework for Highway Extraction and Vehicle Detection Based on a Geometric Deformable Model. Springer, Journal of Photography and Remote Sensing (ISPRS), Volume 61, pp 170-186, December 2006.
16. Zhang W, Fang X Z, Yang X, Moving Vehicles Segmentation Based on Bayesian Framework for Gaussian Motion Model. Elsevier, Pattern Recognition Letters, Volume 27(1), pp 956-967, July 2006.
17. Xainyi Liu, Zhi Qiang Liu, Detection of Vehicles from Traffic Scenes Using Fuzzy Integrals. Pattern Recognition, Volume 35(4), pp 967-980, April 2002.
18. D Pokrajac, V Zeljkovic, L J Latecki, Spatial-Temporal Algorithm for Moving Objects Detection in Infrared Video Sequences, IEEE, September 2005.
19. Carlos Filipe Paulo, Paulo Lobato Correia, Automatic Detection and Classification of Traffic Signs", WIAMIS, June 2007.
20. Boon Chong Chee, Mihai Lazarescu, Tele Tan, Detection and Monitoring of Passengers on Image Analysis and Processing, pp. 143–148, September 2007.

AUTHORS PROFILE

Mrs. Shana L received her B. Tech Degree in Applied Electronics and Instrumentation from Kerala University, India and M.E. Degree in Applied Electronics from Anna University, India. She has four years of teaching experience. Her area of interest includes Image Processing, Medical Image Processing and Biomedical Engineering.

Dr. C Seldev Christopher received his B.E. Degree in Electronics and Communication Engineering from Manonmaniam Sundaranar University, India in 1997 and M.E. Degree in Communication Systems from Madurai Kamaraj University, India in 1999. He is in the teaching field since January 2000. He completed his Doctorate Degree in 2010 from Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, India. He is a Senior Member of IEEE and member of ACM, IET and ISTE. He has more than thirty five papers in International Journals and more than thirty papers in International Conferences to his credit. His area of interest includes Image Processing and Signal Processing. He is a resource person in various FDP and Workshops in the field of Image Processing and MATLAB. He is the approved supervisor for guiding Ph.D students under Anna University Chennai.