Detecting the Moving Object in Dynamic Backgrounds by using Fuzzy-Extreme Learning Machine

N.Keerthana[#1], K.S.Ravichandran[*2], B. Santhi[#3]
School of Computing, SASTRA University, Thanjavur-613402, India
keerthana.ns@gmail.com

ABSTRACT:- **Moving object detection in dynamic background is the important features in video surveillance systems. Detecting the moving object using the SOM in video streams are not suitable for dynamic background and it requires complex computation to adjust the threshold values based on HSV. This paper proposes Fuzzy-Extreme Learning Machine (FELM) for detecting the object in dynamic backgrounds. The proposed model involves Fuzzy-Extreme Learning Machine and Self Organizing Map (SOM) which are used to detect the moving objects as well as shadow elimination in dynamic background. Again it automatically determines the threshold values for various video sequences. The proposed approach identifies the moving objects automatically without human intervention and eliminates the shadows more effectively when compared to other existing methods in the recent literature.**

Keywords – Dynamic background, Fuzzy Extreme Learning Machine, HSV, Self Organizing Map, Shadow elimination, Surveillance systems, Video analysis.

I. INTRODUCTION

Surveillance is the monitoring of behavior. Object detection is the important task performed in video analysis to deal with surveillance [9] or security [10]. Video surveillance is a process of analyzing the video sequences. It is an active research area in computer vision. There are three types of video surveillance activities. Video surveillance activities can be manual, semi-automatic or fully-automatic. In the manual video surveillance system, the human wants to adjust large number of the parameters manually for each video sequence. Adjusting the parameter value for the each video is the complex task in this system. In Semi-automatic video surveillance system, it is having video processing with significant human intervention. In a fully-automatic system, the input is the video sequence that is taken from the surveillance system. In this system there is no human intervention [1].

In this system, monitoring the video for long duration by human operator is impracticable. The human wants to supervise the several screens to find the perilous situation. This leads to high failure to recognize the perilous situation in video surveillance system [2]. In this paper, Self Organizing Map is proposed to work in a dynamic environment. The proposed system also consists of Fuzzy Extreme Learning Machine stage, it determines the threshold values automatically and it does not need human involvement to process the different video streams. This method preserves the robustness in its performance.

II. RELATED WORK

The main aim to perform the background Elimination is to identify the moving object from the video in surveillance system. Chacon-Murguia and Gonzalez-Duarte proposed Neural Fuzzy technique to detect the object in the dynamic background without any human intervention. On fuzzy stage,sugeno and the mamdani fuzzy system are used. By using that fuzzy methods, video segmentation is done. The drawbacks in the sugeno fuzzy system are more complex and it takes more computation time and it also requires more memory to compute the parameters [11]. Joshi and Thakore proposed different techniques for detecting the moving objects and also tracking the detected object in video surveillance system for security purpose. In this method background subtraction with various parameter values are used to detect the moving objects. For detecting the moving object temporal frame difference method is used [1]. Du-Ming Tsai and Shia-Chih Lai proposed independent component analysis algorithm for background subtraction. It detects moving and motionless persons in indoor surveillance system quickly. It is used to detect the object in the Indoor scenes only and indoor applications such as homecare and healthcare monitoring system [3].

Rita Cucchirra et al proposed a technique for detecting the shadow and suppression are used in a system for detecting the moving object detection. HSV color space is used to improve the performance in detecting shadows, because moving shadow can affect the current localization and detection of moving objects [4]. Kalva et al proposed neural network architecture to form an unsupervised Bayesian Classifier for the application domain. It detects the moving object in natural scenes and it handles the segmentation [5]. Shih-chieh et al proposed adaptive Local-Patch Gaussian Mixture Model (LPGMM) method is used to detect the objects in the dynamic background from videos. Then SVM classifier is used to detect the foreground objects and shadow regions [6]. Maddalena and Petrosino proposed Self Organizing Background subtraction method to detect the moving object in dynamic background. The human wants to adjust the threshold values, learning rate parameters for every video sequence to detect the moving object. It achieves robustness in performance [7].

Ravichandran and Alsheyuhi proposed the Fuzzy Extreme Learning Machine and it automatically determines the parameter values for the given Input frames. This method is well suited to work in dynamic background [8]. Prati et al proposed the algorithm for detecting moving shadow objects in dynamic scenes. Deterministic and statistical method is used to detect the objects in Indoor and Outdoor scenes. It is a critical task to find the moving shadow objects in video sequences [12]. Zhou et al proposed Iris recognition system. This system is used in airport and Military systems [10]. Benedek and Sziranyi proposed novel adaptive shadow and Bayesian foreground model for detecting the moving object. The novel adaptive shadow model eliminates the shadow more effectively [13].

III. MATERIALS AND METHODS

III. A. Experimental Setup

The proposed model deals with dynamic background situations. The video samples used in our system are color videos. The video sequence 1 is taken from the webcam and it can have the resolution of 120×160pixels at 15 frames per second [11]. The video sequence 2 contains indoor scenes and have resolution of 320×240 at 24 frames per second [7]. The video sequence 3 contains outdoor scenes and have resolution of 352×288 pixels at 24 frames per second [7]. In this proposed model, video is given as the input and that video are converted into frames. The main goal of the background subtraction is to detect the moving objects from the video sequence taken from the stationary camera. In SOM, for each video streams want to adjust the manual threshold value Th1 and Th2. By using FELM, it automatically determines the threshold value Th1 and Th2. Th1 and Th2 are the segmentation threshold. Finally, the dynamic moving object is detected. It preserves the robustness in its performance as shown in Fig. 1.

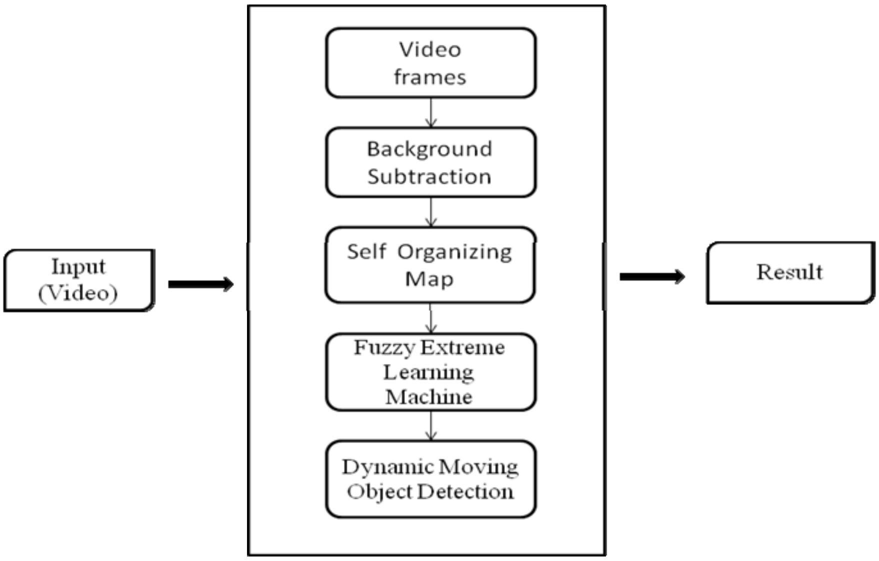


Fig. 1. Architecture of Proposed Model

(a) (b) (c)

Fig 2. Individual frames taken from the video used in our proposed approach

Fig 2. Represents a single frame taken from the original video. Fig. 2 (a) and 2 (c) video frames containing outdoor scenes. The Fig 2 (b) video frames containing Indoor scenes.

III.B. Methods

1. Moving object detection in dynamic background

The proposed Model aimed to detect the moving objects from a video sequence in Dynamic Backgrounds with stationary cameras. The Self Organizing Map (SOM) decreases the computational load to improve the video segmentation. During the training phase, the output provides the highest activation unit to the given input pattern and declared as the winner. The output node whose incoming weight is the shortest Euclidean distance from the input vector is considered as the winner-take-all. The process of the SOM is to select the output layer topology and train the weights from input to output layer. Unsupervised learning is used in the SOM and it automatically produces the classes in the dataset.

In this model, input is given as the video file in AVI format. First step is to separate the frames from the video. Then convert it from RGB plane to HSV plane and it forms the neuronal map for the first frame (HSV) and that frame is considered as the background. The HSV components mapped with each pixel (x, y) of the first sequence frame and update weights. Finally, the remaining frames will be converted to HSV and it will be compared to the background and detect the moving objects from the remaining frames by using the values Th1 and Th2 [7]. The Th1 and Th2 segmentation threshold value. The shadow detection is performed well in Hue-Saturation-Value (HSV) color space.

2. Procedure

Step 1: Transmute the video streams into a number of frames.

Step 2: Set first frame as a background.

Step 3: Calculate HSV for the frame and form Neuronal map ‘NM’ using weight vectors.

Step 4: Calculate HSV for the next frame.

Step 5: Calculate Euclidean Distance between HSV and Neuronal map.

Step 5 (a): Set the segmentation threshold value Th1 and Th2.

Step 6: If Minimum distances $D<Th1$, Set the corresponding pixel as background.

Step 6 (a): Else the distance $D\geq Th2$, Set the corresponding pixel as foreground.

Step 7: Update the weight vectors.

Step 8: Else set as object.

Step 9: Repeat Step 4 to Step 8 up to the last frame.

III.C. Fuzzy Extreme Learning Machine

Huang et al., proposed the new algorithm called Fuzzy Extreme Learning Machine for Single Hidden layer Feed- forward neural networks (SLFNs). It automatically determines the threshold value for the given video sequence to detect the moving object. This model is appropriate for Dynamic Environment. SLFNs randomly choose the weight for the input and analysis the output weights. The SLFN have n pairs of approximate input and output values and also have P hidden nodes namely, z= (x_i,y_i) where $x_i \in R^n$, $y_i \in R^m$, for i = 1, 2, …, n, then standard SLFNs with P hidden nodes and output function k(x) are modeled as $\sum_{i=1}^{p}\beta_i S(a_i,x_j,b_i)=t_j$, where j=1,2,….,n.$(a_i,b_i)$ are the hidden node parameters and β_i are the weight vector .The β_i connects the ith hidden node and the output node. The mathematical model is given as $\sum_{i=1}^{p}\beta_i S(a_i,x_j,b_i)=t_j$, where j=1,2,….,n and it is equivalent to $H\beta=T$ [8].where

$$H=\begin{pmatrix} h(x_1) \\ . \\ . \\ h(x_n) \end{pmatrix} = \begin{pmatrix} S(a_1,x_1,b_1) & \ldots\ldots\ldots & S(a_p,x_1,b_p) \\ . & & . \\ . & & . \\ S(a_1,x_n,b_1) & \ldots\ldots\ldots & S(a_p,x_n,bp) \end{pmatrix}_{n\times P} \quad . \tag{1}$$

$$\beta=\begin{pmatrix} \beta_1^T \\ . \\ . \\ \beta_P^T \end{pmatrix}_{P\times m} \quad \text{and} \quad T=\begin{pmatrix} t_1^T \\ . \\ . \\ t_p^T \end{pmatrix}_{n\times m} \qquad (2)\ \&\ (3)$$

The output matrix H is calculated for hidden layer. The output of the ith node is from ith column of H with respect to input x1,x2,….,xn.The output weight vector is $\beta = H^{\#} T$. $H^{\#}$ is the Moore-penrose [14] generalized inverse of Hidden layer H.

1. ELM Algorithm

For the given set of training input / output values $\{z= (x_i,y_i)$ where $x_i \in R^n$, $y_i \in R^m$, for $i = 1, 2, \ldots, n\}$, the activation function k(x) and the hidden nodes P.

Step 1: Generate random hidden node parameter (ai,bi) for i = 1,2, …, n by using continuous distribution.

Step 2: The output matrix H is calculated for hidden layer.

Step 3: The output weight β is calculated based on the connection $\beta = H^{\#} T$.

2. The Fuzzy-Extreme Learning Machine Features

1) Quick Learning Speed.

2) Good generalization performance.

3) Decision making is fast.

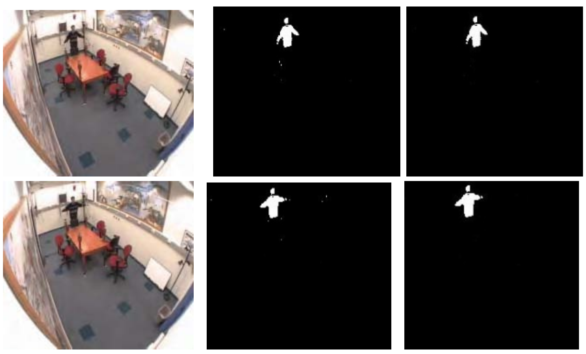
4) It reduces the computation time.

IV. RESULTS AND ANALYSIS

The moving object detection is determined by using the value of Th1 and Th2 generated by the human and that values are compared with FELM. Th1 and Th2 are the segmentation threshold value. The original video is converted into frames and detects the background/foreground moving object. The original frame is given as the input as shown in fig. 3 (a). By using SOM the manual parameter is adjusted by a human viewer for each video sequence and they detect the moving object from the video as shown in fig. 3 (b). In fig. 3 (c) FELM, it automatically adjusts the main parameter contains in the SOM model and detect the moving object more efficiently in both Indoor and outdoor scenes.

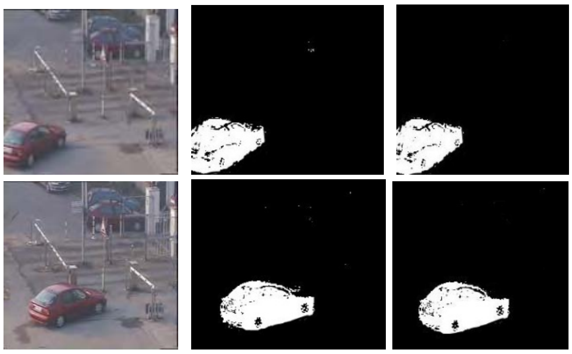
(a) (b) (c)

Fig.3. Detection of moving object in video 1: (a) Original video frame (b) SOM (c) FELM



(a) (b) (c)

Fig.4. Detection of moving objects video 2: (a) Original video frame (b) SOM (c) FELM



(a) (b) (c)

Fig.5. Detection of moving objects in video 3: (a) Original video frame (b) SOM (c) FELM

IV.A. Analysis

TABLE I
Th1 and Th2 values determined by Human Viewer

Video	**Th1**	**Th2**
1	0.06	0.05
2	0.10	0.08
3	0.03	0.02

The video sequence 1 containing outdoor scenes. The corresponding foreground is calculated by SOM and FELM. In SOM, the human assigns the parameter values of Th1, Th2, L1, L2.Th1 and Th2 are the segmentation threshold values. The c1 and c2 are the learning rate to measure and update the background quickly. For video sequence 1, the foreground value calculated by SOM are Th1=0.06, Th2=0.05 as shown in Table I. Finally the moving object is perfectly detected. The video sequence 1 and 3 containing outdoor scenes. The video sequence 2 containing indoor scenes. The FELM automatically determine the parameter value and determine the object more effectively. Likewise the other video sequences are computed and determine the moving object in surveillance system. The FELM randomly generating training and testing dataset. The average training accuracy is 0.9733.The average testing accuracy is 0.9496.By using this method, it reduces the time to train the neural network [12].

TABLE II
Comparison of different methods and proposed method

Method	Shadow Detection	Foreground evaluation from the frame	Indoor/outdoor scenes	Dynamic Background update
A. prati 2003 [12]	Yes	Yes	Both	No
C.Benedek 2008 [13]	Yes	Yes	Both	No
D-M.Tsai 2009 [3]	No	Yes	Indoor scene	No
Proposed Work	Yes	Yes	Both	Yes

V. CONCLUSION

The proposed method is the automatic model, it automatically determines the threshold values Th1 and Th2 without human intervention and it doesn't need previous training. It automatically tuned the threshold values for each video without any previous training. By using SOM, human want to tune the threshold parameter Th1 and Th2 for every video to find the moving object in the dynamic background. The Fuzzy Extreme Learning Machine approach proposed in this paper determines the threshold value automatically and eliminates the shadow more effectively in surveillance system. This proposed method works well in dynamic environments.

REFERENCES

[1] Kinjal A Joshi, Darshak and G. Thakore, "A survey on Moving object detection and tracking in video surveillance system," IJSCE.,vol. 2,no. 3,pp. 44-48, July 2012.

[2] H. M. Dee and S. A. Velastin, "How close are we to solving the problem of automated visual surveillance?" *Mach. Vis. Appl.*, vol. 19, no. 5/6, pp. 329–343, Sep. 2008.

[3] D. - M. Tsai and S. -C. Lai, "Independent component analysis-based background subtraction for indoor surveillance," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 158–167, Jan. 2009.

[4] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti, "Improving shadow suppression in moving object detection with HSV color information," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, Aug. 2001, pp. 334–339.

[5] D. Culibrk, O. Marques, D. Socek, H. Kalva, and B. Furth, "Neural network approach to background modeling for video object segmentation,"*IEEE Trans. Neural Netw.*, vol. 18, no. 6, pp. 1614–1627, Nov. 2007.

[6] Shih-chieh wang,Te-Feng su and Shang-Hong Lai, "Detecting moving objects from dynamic backgrounds with shadow removal,"IEEE Trans., pp 925-928, June 2011.

[7] L. Maddalena and A. Petrosino, "A self-organization approach to background subtraction for visual surveillance applications," *IEEE Trans.Image Process.*, vol. 17, no. 7, pp. 1168–1177, Jul. 2008.

[8] K. S. Ravichandran and Alsheyuhi, S.S.S. ,"FELM based Intelligent Optimal Swithing Capacitor Placement," IEEE Trans., pp 366 – 371, Jul 2011.

[9] T. Celik and H. Kusetogullari, "Solar-powered automated road surveillance system for speed violation detection," *IEEE Trans. Ind. Electron.*, vol. 57, no. 9, pp. 3216–3227, Sep. 2010.

[10] Z. Zhou, Y. Du, and C. Belcher, "Transforming traditional iris recognition systems to work in non-ideal situations," *IEEE Trans. Ind. Electron.*, vol. 56, no. 8, pp. 3203–3213, Aug. 2009.

[11] Mario I. Chacon-Murguia and Sergio Gonzalez-Duarte," An adaptive neural-fuzzy approach for object detection in dynamic backgrounds for surveillance systems," *IEEE Trans. Ind. Electron* , vol. 59, no. 8, pp. 3286–3298, Aug. 2012.

[12] A. Prati, I. Mikic, M. Trivedi, and R. Cucchiara, "Detecting moving shadows: Algorithms and evaluation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 7, pp. 918–923, Jul. 2003.

[13] C. Benedek and T. Szirányi, "Bayesian foreground and shadow detection in uncertain frame rate surveillance videos," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 608–621, Apr. 2008.

[14] D.serre,"TMatrices: Theory and Applications", Springer, Berlin, 2002.