Background Subtraction for Effective Object Detection by using M.O.H

Billipilli Rajasri
PG Scholar, Dept. of ECE
Sri Mittapalli College of Engineering

Z Vazraiah
Asst. Professor, Dept. of ECE
Sri Mittapalli College of Engineering

V.S.R Kumari
Professor & H.O.D, Dept. of ECE
Sri Mittapalli College of Engineering

Abstract - This paper proposes efficient object detection scheme for videos with fixed background and static cameras based on background subtraction . The scheme suggests two different algorithms; the first one models the background from initial few frames and the second algorithm extracts the objects based on local thresholding. Here these different methods are used effectively for object detection and compare these performance based on accurate detection. After the object foreground detection, the parameters like MSE, PSNR, entropy, mean and standard will be determined. In local threshold based object detection, morphological process and filtering also used effectively for unwanted pixel removal from the background. The strength of the scheme lies in its simplicity. The efficacy of the scheme is shown through comparative analysis with competitive methods. Both visual as well as quantitative measures show an improved performance and the scheme has a strong potential for applications in real time surveillance. The parameters of moving object such as MSE, PSNR, entropy, mean and standard will be evaluated by using M.O.H.

Keywords: Background modeling, M.O.H [multiple oblique histogram], back ground subtraction, video segmentation, video surveillance.

I.INTRODUCTION

In the past two decades object detection and tracking in video is a challenging problem and has been extensively investigated. It has applications in various fields such as video compression, video surveillance, human-computer interaction, video indexing and retrieval etc. Object detection involves locating object in the frames of a video sequence, while object tracking represents the process of monitoring the object‘s spatial and temporal changes in each frame. Object detection can be performed through region based image segmentation, background subtraction, temporal differencing, active contour models, and generalized Hough transforms. In order to allow high resolution images of the people in the scene to be acquired it is reasonable to assume that such people move about in the scene. The suggested background model initially determines the nature of each pixel as stationary or non-stationary and considers only the stationary pixels for background model information. In the background model, for a each pixel location a range of values are defined. Subsequently, in object extraction phase our scheme employs a local threshold, unlike the use of global threshold in conventional schemes.

To monitor the scene reliably it is essential that the processing time per frame be as low as possible. Hence it is important that the techniques which are employed are as simple and as efficient as possible. In surveillance system video sequences are obtained through static cameras and fixed background. A popular approach called background subtraction is used in this scenario, where moving objects in a scene can be obtained by comparing each frame of the video with a background [1]. Firstly, video frames captured from a camera are input to the background subtraction. Pre processing stages are used for filtration and to change the raw input video to a processable format. Background modeling then uses the observed video frame to calculate and update the background model that is representative of the scene without any objects of interest. Foreground detection is where the pixels that show a significant difference to those in the background model are flagged as foreground. Data validation is used to examine the found objects of interest and to eliminate any false matches. A foreground mask can then be output in which pixels are assigned as foreground or background. For effective object detection misclassified objects and shadows are removed.

II. RELATED WORK

For object detection in surveillance system, background modelling plays a vital role. Wren *et al.* haveproposed to model the background independently at each pixel location which is based on computation of Gaussian probability density function (pdf) on the previous pixel values [2].

In Stauffer and Grimson developed a complex procedure to accommodate permanent changes in the background scene [3]. Here each pixel is modeled separately by a mixture of three to five Gaussians. TheW4model presented by Haritaoglu *et al.* is a simple and effectivemethod [4]. It uses three values to represent each pixel in the background image namely, the minimum intensity, the maximum intensity, and the maximum intensity difference between consecutive frames of the training sequence. Jacques *et al.* brought a small improvement through the W4 model together with the incorporation of a technique for shadow detection and removal [5]. McHugh *et al.* Proposed an adaptive thresholding technique by means of two statistical models [6].One of them is nonparametric background model and the other one is foreground model based on spatial information. In ViBe, each pixel in the background can take values from its preceding frames in same location or its neighbor [7]. Then it compares this set to the current pixel value in order to determine whether that pixel belongs to the background, and adapts the model by choosing randomly which value to substitute from the background model. Kim and Kim introduced a novel background subtraction algorithm for dynamic texture scenes[8]. The scheme adopts a clustering-based feature, called fuzzy color histogram (FCH), which has an ability of greatly attenuating color variations generated by background motions while highlighting moving objects. Instead of segmenting a frame pixel-by-pixel, Reddy *et al.* used an overlapping block-by-block approach for detection of foreground objects [9]. The scheme passes the texture information of each block through three cascading classifiers to classify them as background or foreground. The results are then integrated with a probabilistic voting scheme at pixel level for the final segmentation.

Generally, shadow removal algorithms are employed after object detection. Salvador *et al.* developed a three step hypothesis based procedure to segment the shadows [10]. It assumes that shadow reduces the intensities followed by a complex hypothesis using the geometrical properties of shadows. Finally it confirms the validity of the previous assumption. Choi *et al.* in their work of [11] have distinguished shadows from moving objects by cascading three estimators, which use the properties of chromaticity, brightness, and local intensity ratio. A novel method for shadow removal using Markov random fields (MRF) is proposed by Liu *et al.* in [12], where shadow model is constructed in a hierarchical manner. At the pixel level, Gaussian mixture model (GMM) and Local Illumination based Background Subtraction (LIBS) is used, whereas at the global level statistical features of the shadow is utilized.

From the existing literature, it is observed that most of the simple schemes are ineffective on videos with illumination variations, motion in background, and dynamically textured indoor and outdoor environment etc. On the other hand, such videos are well handled by complex schemes with higher computational cost.

B. Extraction of Foreground Object

After successfully developing the background model, a local thresholding based background subtraction is used to find the foreground objects. A constant c is considered that helps in computing the local lower threshold (TL) and the local upper threshold (TU). These local thresholds help in successful detection of objects suppressing shadows if any.

IV. MOH SYSTEM DESIGN

The block diagram of the system is as shown in **figure1**.

Furthermore, to remove misclassified foreground objects and shadows, additional computation is also performed. Keeping this in view, we suggest here a simple scheme called *Multiple Oblique Histogram (MOH)* that models the background by defining an intensity range for each pixel location in the scene. Subsequently, a local thresholding approach for object extraction is used. Simulation has been carried out on standard videos and comparative analysis has been performed with competitive schemes.

III. THE PRAPOSED M.O.H. SCHEME

The MOH scheme consists of two stages. The first stage deals with finding the stationary pixels in the frames required for background modelling, followed by defining the intensity range from those pixels. In the second stage a local threshold based background subtraction method tries to find the objects by comparing the frames with the established background.

A. Development of Background Model

Conventionally, the first frame or a combination of first few frames is considered as the background model. However, this model is susceptible to illumination variation, dynamic objects in the background, and also to small changes in the background like waving of leaves etc. A number of solutions to such problems are reported, where the background model is frequently updated at higher computational cost and thereby making them unsuitable for real time deployment. Further, these solutions do not distinguish between object and shadow. To alleviate these limitations we propose an intensity range based background model. Here the RGB frame sequences of a video are converted to gray level frames. Initially, few frames are considered for background modelling and pixels in these frames are classified as stationary or non-stationary by analyzing their deviations from the mean. The background is then modelled taking all the stationary pixels into account. Background model thus developed, defines a range of values for each background pixel location.

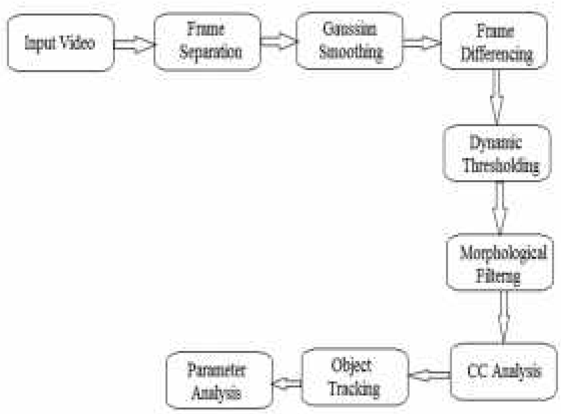


Figure 1: Block Diagram of M.O.H

A. Frame Seperation

An Input Video (.avi files) is converted into still images for processing it and to detect the moving objects. These sequences of images gathered from video files by finding the information about it through 'aviinfo' command. These frames are converted into images with help of the command 'frame2im'. Create the name to each images and this process will be continued for all the video frames.

B. Gaussian Smoothing Process

A Gaussian smoothing is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise and reduce detail. Gaussian smoothing is also used as a pre-processing stage in computer vision algorithms in order to enhance image structures at different scales—see scale space representation and scale space implementation. Mathematically, applying a Gaussian blur to an image is the same as convolving the image with a Gaussian function.

*Gauss Coeff = (1/sqrt (2*pi*sig^2))*
*(exp(x^2+y^2/2*sig^2))*

Where, *x*, *y*, *sig* - input coordinates corresponds to the target and standard Deviation.

C. Segmentation Process

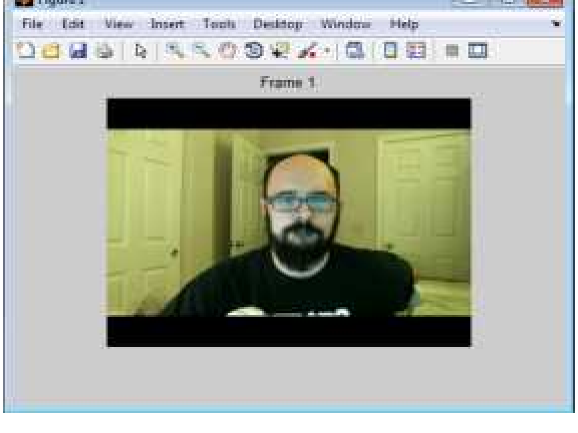
The moving object will be detected by frame subtraction and segmentation algorithms. The frame subtraction is done by subtracting current frame and previous frame for detecting object from background. The moving object extraction from subtracted frames is done by dynamic thresholding method for foreground detection. Then background will be updated by comparing the process frame and background frame.

This will be continued for all consecutive frames. After this process, morphological filtering will be applied for reducing background noise for accurate object detection.

using the morphological process to obtain desired objects. Morphological process involves holes filling and opening the region beyond certain given area. From the labelled image, number of objects, each object region features are obtained for object tracking and counting from each frame sequence.

V. RESULTS

The Input is selected in **Figure 2**



D. Morphological Process

Morphological operations are applied on segmented binary image for smoothening the foreground region. It processes the image based on shapes and it performs on image using structuring element. The structuring elements will be created with specified shapes (disk, line, square) which contains 1's and 0's value where ones are represents the neighbourhood pixels. Dilation and erosion process will be used to enhance (smoothening) the object region by removing the unwanted pixels from outside region of foreground object. After this process, the pixels are applied for connected component analysis and then analysis the object region for counting the objects.

E. Dilation and Erosion

Dilation and Erosion morphological operations are performed on images based on shapes. It is formed by structuring element. It is a matrix containing 1's and 0's where 1's are called neighbourhood pixels. The output pixel is determined by using these processing pixel neighbours. Here, the 'line' structuring element is used to dilate and erode the image for smoothing.

Dilation: It is the process of adding a pixel at object boundary based on structuring element. The rule to find output pixel is the maximum of input pixels neighbourhood matrix.

Erosion: It is to remove the pixel from the object boundary depends on structuring element.

F. CC Analysis

Connected component analysis is a process to label the segmented image foreground pixels with 4 or 8 neighbourhood connectivity. It will be used to separate the image into n number of local objects from grouping of similar pixels. Then the irrelevant background object with maximum area was removed

Figure 2: Input

The output of M.O.H is shown in **Figure 3**

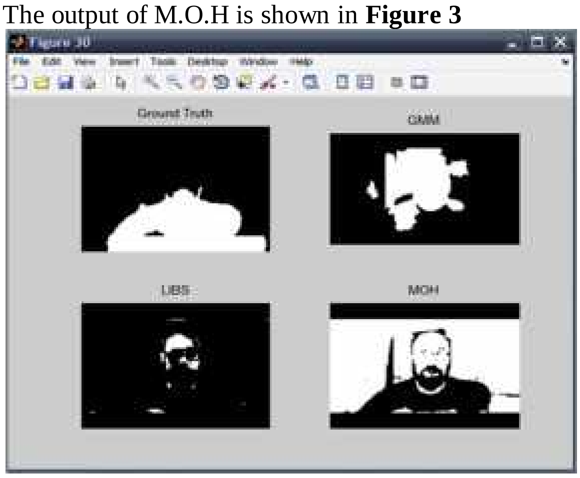


Figure 3: M.O.H Output

The Parameter values are shown in **Figure 4**

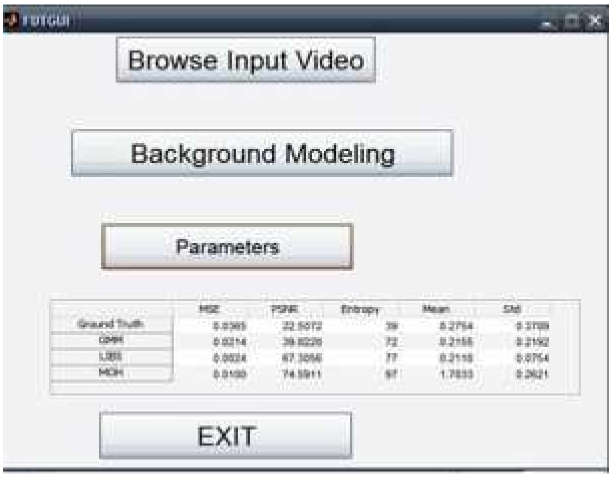


Figure 4: Parameters

CONCLUSION

In this work we have proposed a simple but robust scheme of background modelling and local threshold based object detection. Videos with variant illumination background, textured background, and low motion background are considered for simulation to test the generalized behaviour of the scheme. Recent schemes are compared with the proposed scheme (Ground truth, GMM, LIBS), both qualitatively and quantitatively and parameters like MSE, PSNR, entropy, mean, standard are also calculated. In general, it is observed that the suggested scheme outperforms others and detects objects in all possible scenarios considered.

REFERENCES

[1] Kalyan Kumar Hati, Pankaj Kumar Sa, and Banshidhar Majhi, “Intensity Range Based Background Subtraction for Effective Object Detection”, IEEE Signal Processing Letters, Vol. 20, No. 8, pp.759- 762 August 2013.

[2] A. Yilmaz, O. Javed, and M. Shah, “Object tracking: A survey,” ACM Comput. Surv., vol. 38, Dec. 2006.

[3] C.Wren, A. Azarbayejani, T. Darrell, and A. Pentland, “Pfinder: Realtime tracking of the human body,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 19, no. 7, pp. 780–785, Jul. 1997.

[4] C. Stauffer and W. Grimson, “Adaptive background mixture models for real-time tracking,” in *IEEE Comput. Soc. Conf. CVPR*, 1999, pp. 246–252.

[5] I. Haritaoglu, D. Harwood, and L. Davis, “W4: Real-time surveillance of people and their activities,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 22, no. 8, pp. 809–830, Aug. 2000.

[6] J. Jacques, C. Jung, and S. Musse, “Background subtraction and shadow detection in grayscale video sequences,” in *Eighteenth Brazilian Symp. Computer Graphics and Image Processing*, Oct. 2005, pp. 189–196.

[7] J.McHugh, J. Konrad,V. Saligrama, and P. Jodoin, “Foreground-adaptive background subtraction,” *IEEE Signal Process.Letters*, vol. 16, no. 5, pp. 390–393, May 2009.

[8] O. Barnich and M. Van Droogenbroeck, “ViBe: A universal background subtraction algorithm for video sequences,” *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1709–1724, Jun. 2011.

[9] W. Kim and C. Kim, “Background subtraction for dynamic texture scenes using fuzzy color histograms,” *IEEE Signal Process. Lett.*, vol. 19, no. 3, pp. 127–130, Mar. 2012.

[10] V. Reddy, C. Sanderson, and B. Lovell, “Improved foreground detection via block-based classifier cascade with probabilistic decision integration,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 1, pp. 83–93, Jan. 2013.

[11] E. Salvador, A. Cavallaro, and T. Ebrahimi, “Cast shadow segmentation using invariant color features,” *Comput. Vis. Image Understand.*, vol. 95, no. 2, pp. 238–259, 2004.

[12] J. Choi, Y. Yoo, and J. Choi, “Adaptive shadow estimator for removing shadow of moving object,” *Comput. Vis. Image Understand.*, vol. 114, no. 9, pp. 1017–1029, 2010.

[13] Z. Liu, K. Huang, and T. Tan, “Cast shadow removal in a hierarchical manner using MRF,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 1, pp. 56–66, Jan. 2012.

[14] L. Li,W.Huang, I.Y.H.Gu, andQ. Tian, “Foreground object detection from videos containing complex background,” in *Proc. Eleventh ACM Int. Conf. Multimedia*, Nov. 2003, pp. 2 10.