International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 3, Issue 9, September 2014

Moving Object Tracking using Background Subtraction Technique and its Parametric Evaluation

P. Bhuvaneswari, T. Siva Kumar

***Abstract*— This paper proposes efficient motion detection and people counting based on background subtraction using dynamic threshold approach with mathematical morphology. Here these different methods are used effectively for object detection and compare these performance based on accurate detection. Here the techniques frame differences, dynamic threshold based detection will be used. After the object foreground detection, the parameters like speed, velocity motion will be determined. For this, most of previous methods depend on the assumption that the background is static over short time periods. In dynamic threshold based object detection, morphological process and filtering also used effectively for unwanted pixel removal from the background. The background frame will be updated by comparing the current frame intensities with reference frame. Along with this dynamic threshold, mathematical morphology also used which has an ability of greatly attenuating color variations generated by background motions while still highlighting moving objects. Finally the simulated results will be shown that used approximate median with mathematical morphology approach is effective rather than prior background subtraction methods in dynamic texture scenes and performance parameters of moving object such sensitivity, speed and velocity will be evaluated.**

***Index Terms*—Background subtraction, Dynamic thresholding, Video surveillance, Video segmentation, Video motion, Object tracking.**

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 [1] 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. 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 [2]. Firstly, video frames captured from a camera are input to the background subtractor. Pre processing stages are used for filtration and to change the raw input video to a processable format. Background modelling 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

Background modeling plays a vital role for object detection in surveillance system.

A. *Generalized Stauffer- Grimson Background Subtraction for Dynamic Scenes* [3]

Here an adaptive model for backgrounds containing significant stochastic motion (e.g. water). The new model is based on a generalization of the Stauffer– Grimson background model, where each mixture component is modelled as a dynamic texture. We derive an online K-means algorithm for updating the parameters using a set test2 of sufficient statistics of the model. Finally, we report on experimental results, which show that the proposed background model both quantitatively and qualitatively outperforms state-of-the-art methods in scenes containing significant background motions.

Techniques used:

- Dynamic textures.
- Background models.
- Background subtraction.
- Mixture models.
- Adaptive models.

ISSN: 2278 – 909X

B. Tracking and counting people in visual surveillance systems [4]

The greatest challenge on monitoring characters from a monocular video scene is to track targets under occlusion conditions. In this work, we present a scheme to automatically track and count people in a surveillance system. First, a dynamic background subtraction module is employed to model light variation and then to determine pedestrian objects from a static scene. To identify foreground objects as characters, positions and sizes of foreground regions are treated as decision features. Moreover, the performance to track individuals is improved by using the modified overlap tracker, which investigates the centroid distance between neighbouring objects to help on target tracking in occlusion states of merging and splitting. On the experiments of tracking and counting people in three video sequences, the results exhibit that the proposed scheme can improve the averaged detection ratio about 10% as compared to the conventional work significant background motions.

Techniques used:

- Intelligent surveillance System.
- People Tracking.
- People Counting.
- Overlap Tracker.
- Occlusion.

C. People Counting System based on BP Neural Network

A people-counting system based on a Back Propagation (BP) neural network is proposed in this paper. The proposed system uses cheap photoelectric sensor to collect data and introduces BP neural network for counting and recognition, and it is effective and flexible for the purpose of performing people counting. In this paper, new methods for segmentation and feature extraction are developed to enhance the classification performance. Promising results were obtained and the analysis indicates that the proposed system based on BP neural network provides good results with low false rate and it is effective for people-counting.

Techniques used:

- Segmentation.
- Feature Extraction.
- Back Propagation Neural Network.

D. Real Time Human Motion Detection and Tracking [5]

Human motion detection and tracking system is an automated video surveillance system for detecting and monitoring people in both indoor and outdoor environments. Detection and tracking are achieved through several steps: First, we design a robust, adaptive background model that can deal with lightning changes, long term changes in the scene and objects occlusions. This model is used to get foreground pixels using the background subtraction method. Afterwards, noise cleaning and object detection are applied, followed by human modeling to recognize and monitor human activity in the scene such as human walking or running.

Techniques used:

- Motion Detection.
- Tracking.
- Human Model.
- Surveillance.
- Image Processing.

E. Fuzzy Color Histogram and its use in Color Image Retrieval [6]

A conventional color histogram (CCH) considers neither the color similarity across different bins nor the color dissimilarity in the same bin. Therefore, it is sensitive to noisy interference such as illumination changes and quantization errors. Furthermore, CCHs large dimension or histogram bins require large computation on histogram comparison. To address these concerns, this paper presents a new color histogram representation, called fuzzy color histogram (FCH), by considering the color similarity of each pixel's color associated to all the histogram bins through fuzzy-set membership function. A novel and fast approach for computing the membership values based on fuzzy -means algorithm is introduced. The proposed FCH is further exploited in the application of image indexing and retrieval. Experimental results clearly show that FCH yields better retrieval results than CCH. Such computing methodology is fairly desirable for image retrieval over large image databases.

Techniques used:

- Conventional color histogram.
- Fuzzy means.
- Fuzzy histogram.
- Illumination changes.
- Image indexing and retrieval.
- Membership matrix.

III. SYSTEM DESIGN

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

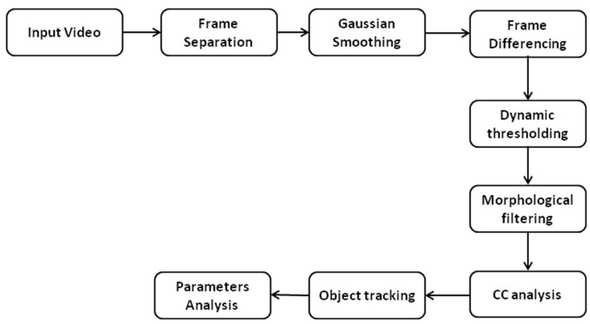

Figure1: Block diagram.

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

ISSN: 2278 – 909X

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. Figure2 represents the process flow of frame separation.

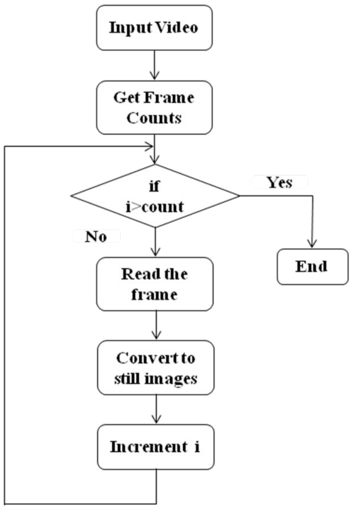

Figure2: Process flow.

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. Figure3 represents the segmentation process.

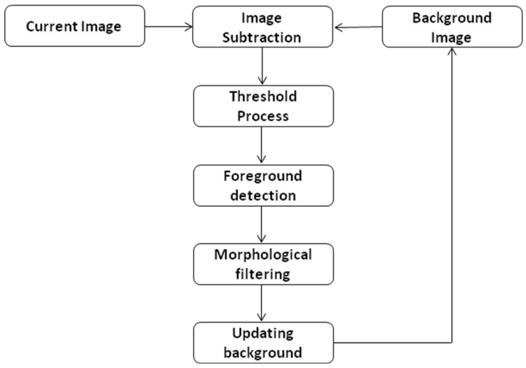

Figure3: Segmentation Process.

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

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. Figure4 represents the morphological process.

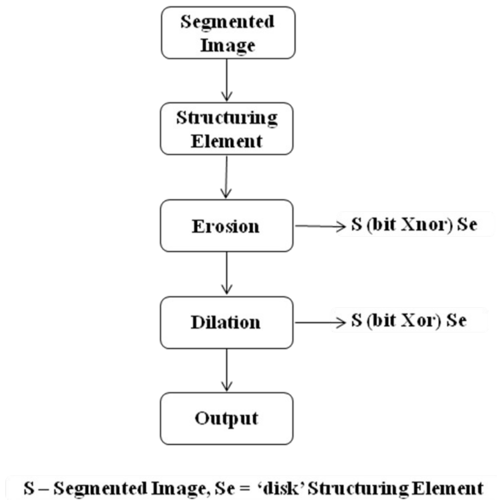

Figure4: Morphological process.

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.

The rule to find output pixel is the maximum of input pixels neighbourhood matrix. Finally the output image is smoothened for reducing distortion from background and edge sharpness.

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 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.

IV. PARAMETER EVALUATION

- *Velocity:* The velocity of object is evaluated based on distance travelled by an object and frame rate

  Velocity = Distance travelled / Frame rate

- *Sensitivity:* It measures the proportion of actual positives which are correctly identified

  Sensitivity = Tp. / (Tp + Fn)

  Where,

  Tp = True Positive: Object pixels correctly classified as object.

  Fn = False negative: Object pixels incorrectly classified as background.

- *Correlation Coefficient:* It is used to find the similarity between two different images with their intensities. It will be described by,

  *Cor_coef = [sum (sum (u1.*u2))] / [sqrt(sum(sum(u1.*u1))*sum(sum(u2.*u2)))];*

  Where,

  u1 = F1 – mean of F1, u2 = F2 – mean of F2

  F1 – Obtained result and F2 – Ground truth

- *PSNR (Peak Signal to Noise Ratio)*

$$PSNR = 10\, Log_{10} \frac{255^2}{MSE}$$

- *RMSE (Root Mean Square Error)*

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M} \sum_{j=0}^{N} (a_{i,j} - b_{i,j})^2$$

RMSE = sqrt (MSE)

Where,

M, N are Number of Rows and Columns

a_{ij} – Actual result and b_{ij} – Desired result

V. SOFTWARE REQUIRED

- MATLAB 7.5 and above versions
- Video Processing toolbox

VI. RESULTS

A. Input Video

Input video can be selected is as shown in figure5.

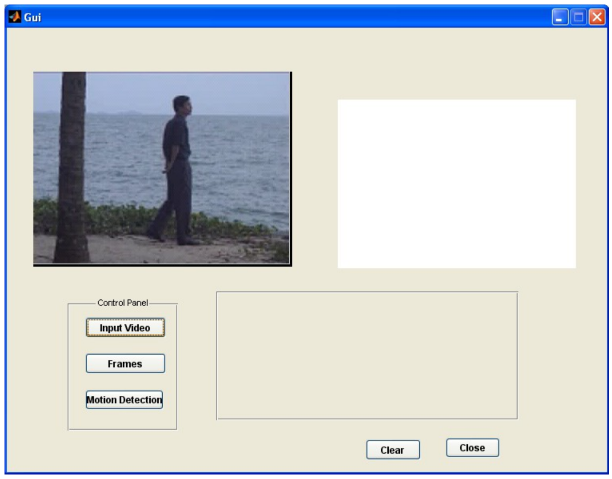

Figure5: Input video.

B. Object Detection

The objects in the input video are detected as shown in figure6.

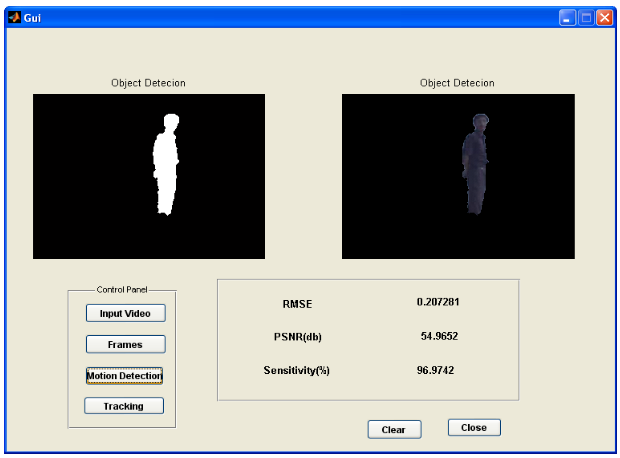

Figure6: Object detection.

C. Object Traction

The objects in the input video is identified as shown in figure7

ISSN: 2278 – 909X

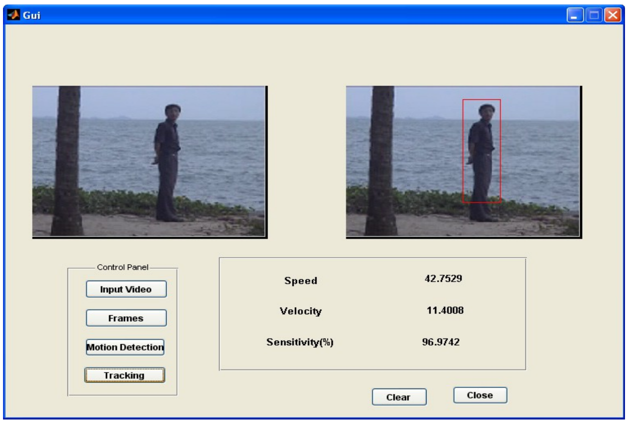

Figure7: Object tracking.

VII. Conclusion

The paper presented an efficient motion detection based on background subtraction using frame difference with thresholding and mathematical morphology. It will be enhanced with futures of connected component analysis and morphological filtering for tracking and counting moving objects. After the foreground detection, the parameters like Count, velocity of the motion was estimated and performance of object detection will be measured with sensitivity and correlation using ground truth. Finally the proposed method will be proved that effective for background subtraction in static and dynamic texture scenes compared to prior methods.

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

[4] 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.

[5] 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.

[6] 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.

[7] 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. 189–19 6.

