Moving Object Detection Using Recursive Algorithm

P.Rajeswari, M.G.Sumithra, V.Premalatha, S.Deepak

***Abstract*— This paper deals with a real time visual surveillance system for detecting moving semantic objects of a certain class (such as humans or cars) from digital videos even in low-level illumination. Several existing systems used colour cues which is a major disadvantage in case of low-level illuminations. And in some existing systems results were not very accurate. In this method the prior identification of object is not needed whereas some existing system requires it. Here recursive algorithm is used to overcome such drawbacks. The main advantage of the proposed method is that shape analysis is used to detect objects without much noise. Both recursive algorithm and shape analysis produce very accurate results. This paper finds major application in case of security purposes.**

***Index Terms*—Digital videos, Low level illumination, Recursive algorithm, Shape analysis.**

I. INTRODUCTION

Object detection is a computer technology related to computer-vision and image processing that deals with detecting instances of semantic objects of a certain class with more information in digital images and videos. In some of the domains of object detection include face detection and walking person detection. Applications of object detection are vast in many areas of computer vision, including image retrieval, security and video surveillance. Moving object detection is a real time system for detecting people and their body parts in monochromatic imagery.

Moving object detection employs a combination of shape analysis and robust techniques for detecting to detect people, and to locate and track their body parts. It builds "Appearance" models of people so that they can be identified after occlusions or after other interactions during which moving object detection cannot track them individually. Moving object detection has been designed to work with only monochromatic video sources. While most previous work on detection and detecting of people has relied heavily on colour cues, moving object detection is designed for outdoor surveillance process, and particularly for night time situations and also like low light level situations. These cases, colour will

not be seen clearly and people need to be detected and tracked based on those weaker appearance and motion cues. Moving object detection is a real time system. It currently is implemented on a dual processor Pentium PC and can process between 20-30 frames per second depending on the image resolution (typically lower for IR sensors than video sensors) and the number of people in its field of view. In the long run, moving object detection will be extended with models to recognize the actions of the people it tracks. People are interested in interactions between human beings and objects - e.g., people exchanging objects, leaving objects in the scene, taking objects from the scene. The descriptions of people – their global motions and the motions of their parts - developed by Moving object detection are designed to support such activity recognition.

Moving object detection currently operates on video taken from a fixed camera, and many of its image analysis algorithms would not generalize easily to images taken from a moving camera. At this point, the surveillance system might stop and invoke a system like Moving object detection to verify the presence of people and recognize their actions.

In Moving objects detection, foreground regions are detected in every frame by a combination of background analysis and simple low level processing of the resulting binary image. The background scene is statically modelled by the minimum and maximum intensity values and maximal temporal derivative for each pixel recorded over some period, and is updated periodically. Each foreground region is matched to the current set of objects using a combination of shape analysis and tracking. These include simple spatial occupancy overlap tests between the predicted locations of objects and the locations of detected foreground regions, and "dynamic" template matching algorithms that correlate evolving appearance models of objects with foreground regions. Second-order motion models, which combine robust techniques for region detection and matching of silhouette edges with recursive least square estimation, are used to predict the locations of objects in future frames.

II. SYSTEM DESCRIPTION

This paper describes the computational models employed by moving object detection to detect the people and their parts. The proposed method is designed to allow moving object detection to determine types of interactions between people and objects and to overcome many inevitable errors and ambiguities that arises in dynamic image analysis. It uses motion cues and shape analysis to detect persons even during low level illumination.

ISSN: 2278 – 909X

The Fig.1 shows the flow diagram of the proposed method. An input video is taken. Then it is converted into its constituent frames. It is followed by Frame differencing where successive frames are compared and the result is threshold to obtain binary images. Now the binary images are subjected to Morphological filtering in order to remove the noise (i.e. unwanted pixels). Then the processed frames are merged into a video. Thus the detected video shows only the moving object without noise.

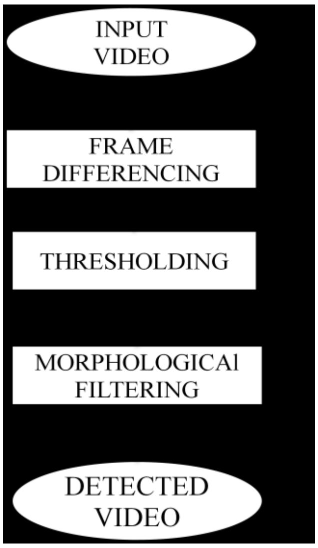

Fig: 1 The Proposed Object Detection Process

A. Input Video

The video from surveillance system is taken .This video contains a person whose motion is being detected

B. Video To Frame Conversion

The total input video is converted into multiple frames which are used for further processing. The total number of frames obtained depends on the length of the input video.

C. Frame Differencing

Frame differencing is based on a model of background variation obtained while the scene contains no people. The background scene is modelled by representing each pixel by three values; its minimum and maximum intensity values and the maximum intensity difference between consecutive frames observed during the motion of the video. These values are estimated over several seconds of video and are updated periodically for all the frames. Used to detect background variation and Effective Segmentation of foreground from background. First, the scene is converted to an array of pixel values. Then the pixel values of the previous frame are subtracted from the current frame's pixel values, and the absolute value of the values is obtained. The resultant array of values that represent how much each pixel has changed between the two frames, with higher values representing more change. The amount of change in a region of pixels can be interpreted as the amount of motion that is taking place in that region. These data can then be used to determine where in the scene the most motion is taking place.

D. Thresholding

The simplest method of image segmentation is called the thresholding method. This method is based on a clip-level or a threshold value to turn a gray-scale image into a binary image. The key of this method is to select the threshold value. Several popular methods are used in industry including the maximum entropy method, Otsu's method and k-means clustering. Recent methods have been developed for thresholding computed tomography (CT) images. The key idea is that, unlike Otsu's method, the thresholds are derived from the radiographs instead of the (reconstructed) image.

Thresholding may be viewed as an operation that involves tests against a function T of the form

$$T = T\ [(x,y),p(x,y),f(x,y)]$$

Thresholding is defined as g(x,y) is 1 if $f(x,y) < T$ or else 0 if $f(x,y) >= T$.

1).OTSU'S method

In computer vision and image processing , Otsu's method is used to automatically perform histogram shape-based image thresholding,[1] or, the reduction of a gray level image to a binary image. The algorithm assumes that the image to be threshold contains two classes of pixels or bi-modal histogram (e.g. foreground and background) then calculates the optimum threshold separating those two classes so that their combined spread (intra-class variance) is minimal. The extension of the original method to multi-level thresholding is referred to as the Multi Otsu method. Otsu's method is named after Nobuyuki Otsu.

Based on a very simple idea: Find the threshold that minimizes the weighted within-class variance. This turns out to be the same as maximizing the between-class variance. Operates directly on the gray level histogram [*e.g.* 256 numbers, P (i)], so it's fast (once the histogram is computed). I've used it with considerable success in "murky" situations.

2).OTSU'S assumption

Histogram (and the image) is bimodal. No use of spatial coherence, or any other notion of object structure. Assumes stationary statistics, but can be modified to be locally adaptive. Assumes uniform illumination (implicitly) , so the bimodal brightness behaviour arises from object appearance differences only. The class probabilities and class means can be computed iteratively. This idea yields an effective algorithm.

Separability factor S is given by $S = \sigma_B^2(t)/\ \sigma^2$ Range of S is $0<S<1$. For good threshold S should be near to 1.

E. Morphology

Mathematical morphology is a set- and lattice-theoretic methodology for image analysis, which aims at

ISSN: 2278 – 909X

quantitatively describing the geometrical structure of image objects.

The structuring element is a binary array. It has an origin. The shape of the structuring element is chose first. With the help of given structuring element erosion and dilation is performed to achieve morphological filtering. The structuring element is a matrix consisting of 0's and 1's.Can has any arbitrary size and shape.eg: flat, diamond, rectangle etc The Origin of a Structuring Element can be determined by, origin = floor ((size (nhood) +1)/2)

1). Erosion

Erosion is one of two fundamental operations (the other being dilation) in morphological image processing from which all other morphological operations are based. It was originally defined for binary images, later being extended to grayscale images, and subsequently to lattices. It can be used to eliminate unwanted pixels it removes a foreground pixel when it is surrounded by its 8 neighbourhood background pixels. It is used to remove white noise. It is based on the logical AND relationship.Erode the image with the structuring element. This removes all the lines, also shrinks the rectangles.

BW2 = imerode (BW1, SE); imshow (BW2)

2). Dilation

Dilation is one of the basic operations in mathematical morphology. Originally developed for binary images, it has been expanded first to grayscale images, and then to complete lattices. The dilation operation usually uses a structuring element for probing and expanding the shapes contained in the input image. It is used to add the pixels in the foreground when it is damaged. Uses logical NAND operation. Here we use erosion followed by dilation method. It is done with the opening Operator. Adds the necessary pixels. To restore the rectangles to their original sizes; dilate the eroded image using the same structuring element, SE.

BW3 = imdilate (BW2, SE); imshow (BW3)

F. Frames to Video Conversion

All the processed frames are again being converted into video. The significance of video is, it has more information than images and also understanding the content present in the video is easy. The length of the video depends on the number of frames present. Finally the output video contains the segmented foreground object without noise.

III. EXPERIMENTAL RESULTS & DISCUSSION

The result of the paper detects the object in the video removing all the noises. It is used in the surveillance system to detect the moving objects. The video is first converted into number of frames, and then the frames are compared with the consecutive frames for frame differencing. Then each frame is subjected to thresholding process. Morphological filtering is done to remove all the noises present in the frames. Then all frames are joined to compose the required output video. The output video contains the detected moving object.

A. Video to Frame Conversion

Fig.2 Input video

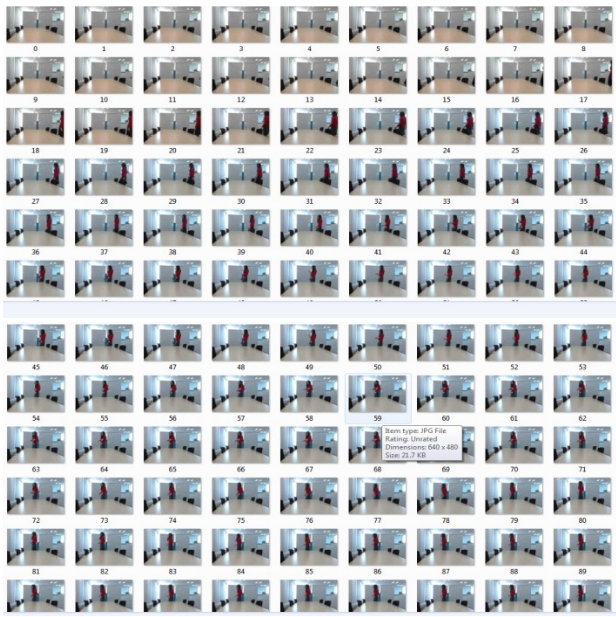

Fig.3 Constituent frames from the video

The video is being converted into frames where zeroth frame shows the background, frame 18 shows the entry of the foreground object and frame 75 shows the foreground object at the middle of the image. The number of frames obtained depends on the length of the video.

B. Thresholding

Thresholding process is used to segment the foreground object from the background. Here the threshold value is selected based on maximum between class variance.

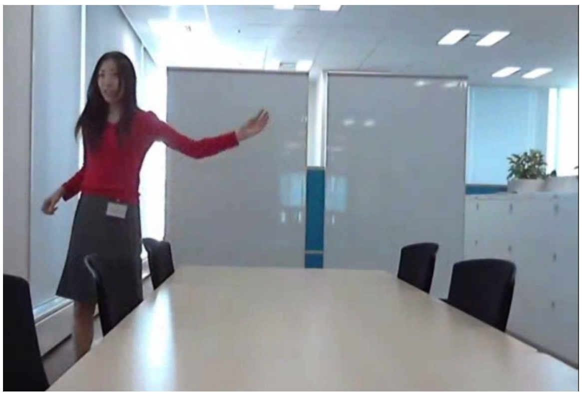

Fig.4 Input Image (frame no 145)

ISSN: 2278 – 909X

Fig.5 Threshold output (frame no 145)

Fig.6 Input Image (frame no 100)

Fig.7 Threshold output (frame no 100)

C. Frame Differencing

Here the frame differencing is done to detect the foreground object by comparing each frame with the reference frame.

Fig.8 Reference frame (frame no 1)

Fig 9 Frame containing object (frame no145)

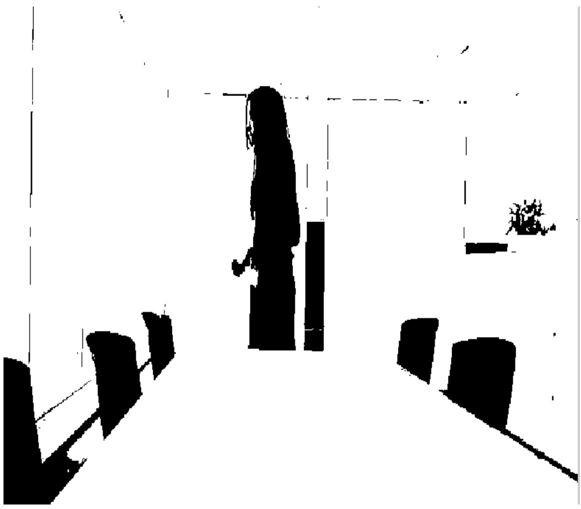

Fig.10 Frame containing object (frame 100)

ISSN: 2278 – 909X

Fig.11 Frame differenced output (145)

Fig.12 Frame Differenced output (100)

D. Morphological Filtering

In the input frames filtering process is being done to remove the noises in the frame. Erosion, dilation, opening, closing process is done to detect the object from the background by removing all the noises.

Fig.13 Filtered output (frame no 145)

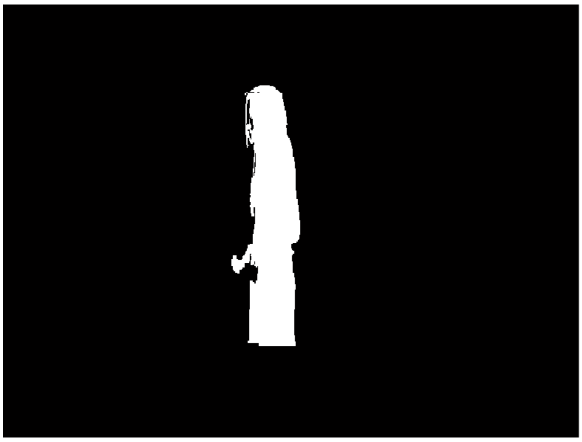

Fig.14 Filtered output (frame no 100)

E. Frames to Video Conversion

All the processed frames are again being converted into video. The length of the video depends on the number of frames present. The output video contains the detected moving object.

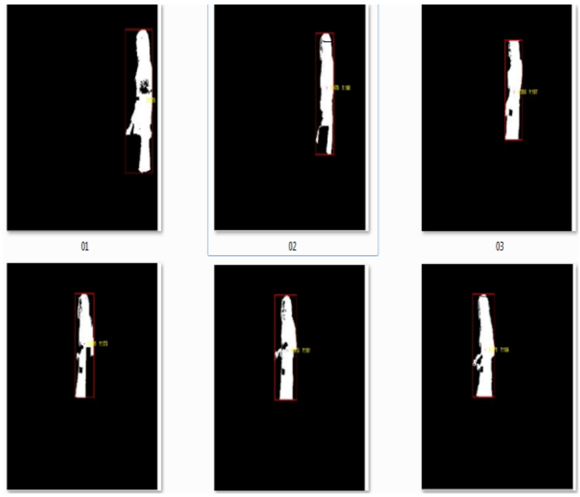

Fig.15 Processed frames

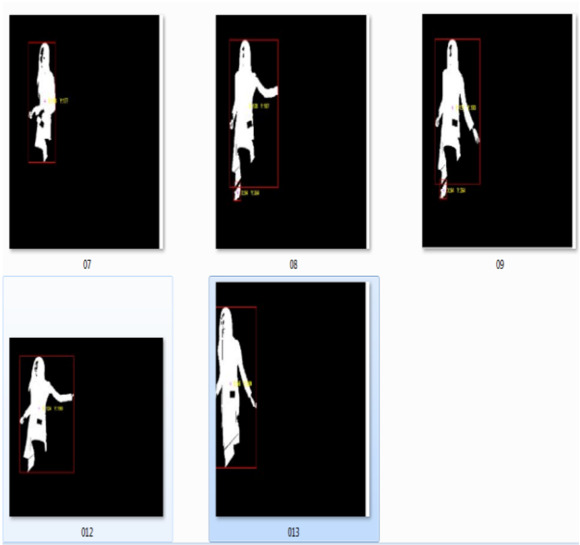

Fig.16 Processed frames

ISSN: 2278 – 909X

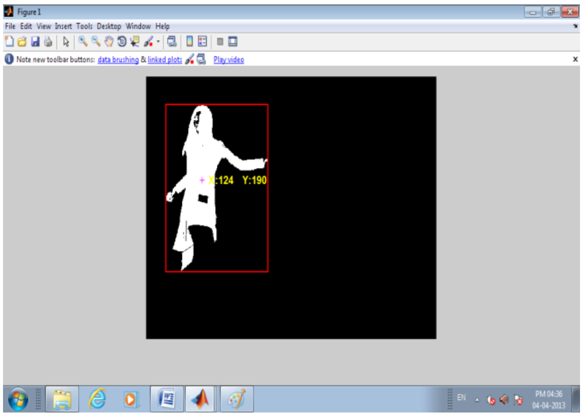

Fig.17 Video with the detected object

.

IV. CONCLUSION

The proposed algorithm extracted the foreground object from the background effectively. It uses shape analysis and motion cues for accurate detection. The complexity is being reduced by using recursive algorithm. This algorithm detected the moving object without noise even in low level illumination. It finds its applications in moving object detection in the prohibited areas in military, traffic areas for security purpose. Future enhancement will be multiple person detection and tracking even through occlusions.

REFERENCES

[1] Li cao, Yi jiang "An effective background reconstruction method for video objects detection" Third international conference on=networking and distributed computing on 2012
[2] Li He Hui wang, hong zhang "Object detection by parts using appearance, structural and shape features" International conference on Bio Engineering & communication on 2011
[3] Weilun lao, Jungong Han,Peter H.N.de"Automatic Video-Based Human Motion Analyzer for Consumer Surveillance System" IEEE Transaction on consumer economics,vol 55,No 2, may 2009.
[4] Christopher Richard Wren, Ali Azarbayejani, Trevor Darrell, and Alex Paul Pent land (1997), "Pfinder: Real-Time Tracking of the Human Body", IEEE Transaction on pattern analysis and machine intelligence, vol. 19, No. 7, pp.780-785.
[5] DimaDamen, David Hogg(2012)," Detecting Carried Objects from sequences of walking Pedestrians" IEEE Transaction on Pattern analysis and machine intelligence ,vol 34, No 6,pp.1056-1067.
[6] DimaDamen, David Hogg(2012)," Detecting Carried Objects from sequences of walking Pedestrians" IEEE Transaction on Pattern analysis and machine intelligence ,vol 34, No 6,pp.1056-1067.
[7] Ismail Haritaoglu, David Harwood and Larry S.Davis (1998) W^4: who? When? Where? What? A Real time system for detecting and detecting people" IEEE Conference on face and gesture recognition,pp.1-6.
[8] Dongming Hao and David G.Daut (1991), " Morphological Hit-or-Miss Transformation for Shape Recognition" journal of visual communication and image representation, Vol. 2,No. 3,pp.230-243.
[9] Bing-Fei Wu, Yen-Lin Chen and Chung-Cheng Chiu(2005), "Recursive algorithms for Image Segmentation Based on a Discriminant Criterion", International Journal of Information and Communication Engineering 1:2,pp.55-60.

P.Rajeswari, *PG Scholar, ECE Department,*
Bannari Amman Institute of Technology, Erode, India.
M.G.Sumithra, *Professor, ECE Department*
Bannari Amman Institute of Technology, Erode, India.
V.Premalatha, *PG Scholar, ECE Department,*
Bannari Amman Institute of Technology, Erode, India.
S.Deepak, *PG Scholar, ECE Department,*
Bannari Amman Institute of Technology, Erode, India.

ISSN: 2278 – 909X