ISSN: 2321-7782 (Online)
Volume 2, Issue 1, January 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Paper
Available online at: www.ijarcsms.com

Object Detection, Classification in Videos and its Parametric Evaluation Using MATLAB

R. Raj Bharath[1]
Assistant Professor
Department of CSE
Manakula Vinayagar Institute of Technology
Puducherry - India

G. Dhivya[2]
PG Scholar
Department of CSE
Manakula Vinayagar Institute of Technology
Puducherry - India

Abstract: Object detection and classification in video is an important and challenging task in the video surveillance. We propose efficient algorithm for moving object detection, classification and evaluate its parameter by alternating the algorithm in effective way. In proposed system image subtraction, threshold and foreground detection algorithms will be used for object detection and patterns are used for classification. Then frame by frame the objects are tracked and parameters like speed, velocity and object count of motioned object will be calculated. Finally the proposed method will be proved that object in both static and dynamic texture scenes over long time period is analyzed and parameter of moving object will be evaluated. It is mainly developed for boarder protection and monitoring, sports training.

Keywords: classification, dynamic texture, foreground detection, Object detection, threshold, pattern.

I. INTRODUCTION

Moving object detection in real time application is a important challenging task in video surveillance systems. There are three steps for automated video analysis. They are object detection, object tracking and classification. As the first step, object detection aims to locate and segment interesting objects in a video. Then, such objects can be tracked from frame to frame, and the tracked object can be analyzed to recognize object categorizes. Thus object detection in video acts as an first step for next processing such as tracking, classification of the detected moving object. Thus, object detection plays a critical role in practical applications and challenging task in real time.

The use of object detection, tracking and classification algorithms are not restricted in video surveillance. The major application of object detection and classifications are broader monitoring and protection, sports training, traffic control etc. Thus object detection play major role in various fields.

In most of the work, Optical flow method is used for detecting moving objects in video and image sequences. For example, adaptive optical flow method for tracking a person is dependent on being able to detect a person in particular across a series of frames in videos. Optical flow can be used to segment and extract a moving object from a scene, provided that estimated velocity of the moving object is known. But successful object detection in video also relies on being able to segment the background from foreground. However, the high computational time is needed to extract the object in optical flow and the lack of classification the foreground from the background. So this method is unsuitable for real time processing and misclassifies unmoved objects or large texture less regions as background.

Then Background subtraction is commonly used technique for segmentation of motion object in static scenes. It will attempts to detect moving object regions by subtracting the current image in each pixel-by-pixel from a reference background image that is created by averaging images over time period. Unfortunately, the derivation of background subtraction model is complex and computationally expensive.

Therefore, most of existing approaches for moving object detection and classification are computationally difficult and subject to more delays, adversely affect the performance of real-time surveillance.

II. OVERVIEW OF THE SYSTEM

In proposed system the aim is to build robust and novel moving object detection, classification algorithm that can detect object in a variety of challenging real world scenarios. The Overall system overview will be represented in figure 1

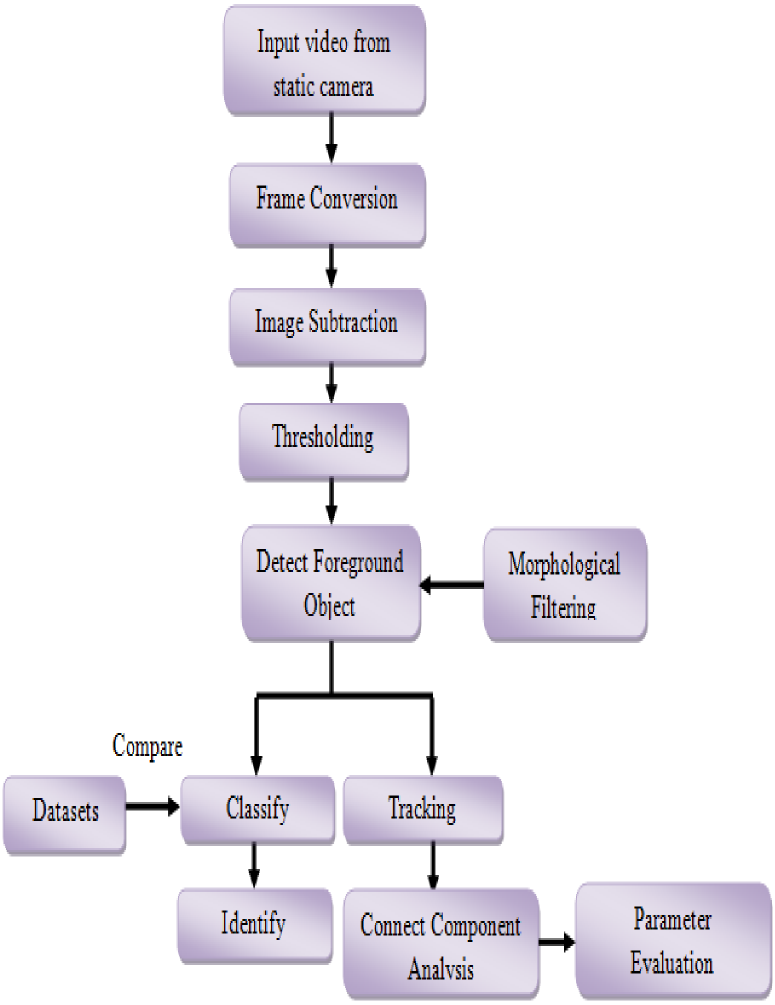


Fig.1 Overview of the system

The steps involved in overall system is

1. The first step is the acquired video is uploaded in application. The input video will be taken for both static and dynamic background. It will acquire from static camera. For processing the video files, convert video into frames and from frames to images.
2. Next step is by using image subtraction techniques, the difference between two images are estimated for evaluating the movement changes of an object.
3. Then segmentation is performed and foreground object is detected. While foreground detection the Unwanted backgrounds are filtered using morphological filtering.
4. After the object detection, the next step is to identify object by comparing with datasets. In datasets, the shapes of objects are maintained.
5. Then next one is parameters like speed, velocities are evaluated by tracking the object and numbers of objects are evaluated by using connected component analysis.

This is an overall system architecture and work flow for object detection, classification and calculates its parameter.

III. MOVING OBJECT DETECTION

Moving object detection is initial step in video surveillances. There are various algorithms used for detecting moving objects. The input video is taken using a static camera. It will be the extension of (E.g.: .Avi) format. It will acquire in both static and dynamic background. The challenges like view point variation, shape, occlusion etc are probably facing.

A. Frame Conversion

For processing an Input Video files, it has to convert it into frames by finding the information about .avi file. After that it has to convert into images. So videos are split into frames. The first frame, which is called reference frame, which represents the reference pixel values for comparing purpose and the second frame which is called the input frame, which contains the moving object. The two frames are compared and the differences in pixel values are determined. For each naming is assigned. The work flow of frame conversion will be given below in diagram form.

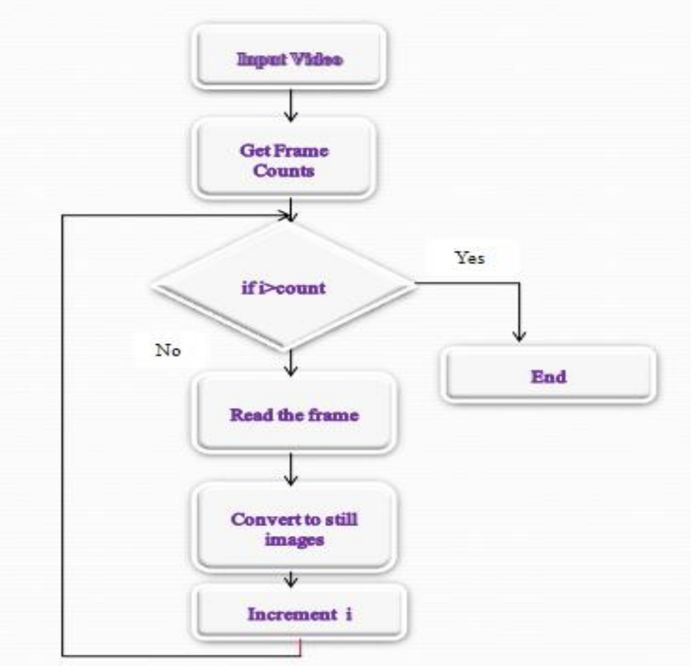


Fig. 2 Work flow of frame conversion

B. Difference Image

The difference between two frames is done with the Difference image techniques. It is nothing but subtraction of images acquired at different instants. That is by subtracting current frame and previous frame for detecting object from background in motion states.

A difference image d(i,j) is binary image where non-zero values represent image area with motion, that is, areas where there was a substantial difference between consecutive images f1 and f2:

$$d(i,j) = \begin{cases} 0 & \text{if } |f1(i,j) - f2(i,j)| <= € \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

Where € is small positive number and i,j represent the pixel value of images in row and column wise.

C. Threshold

The object detection will performed based on thresholding concept. Threshold is one of the segmentation process. It can be determined to segment objects and background. It is the process of separating an image into different portions of images by selecting a certain grayness level as a threshold and comparing each pixel value with that threshold, and then assign the pixel to the different portions, depending on whether the pixel's grayness level is below the threshold or above the threshold value.

So threshold is applied and the pixel is identified as background or foreground. A complete segmentation of an image R is a finite set of region $R_1, R_2, \ldots\ldots\ldots R_S$.

$$R=\bigcup_{i=1}^{S} R_i \qquad (2)$$

Thresholding is the transformation of an input image f to an output (segmented) binary image g as follows

$$g(i,j) = 1 \quad \text{for } f(i,j)>=T$$
$$= 0 \quad \text{for } f(i,j)< T \qquad (3)$$

Where T is the threshold, g(i,j) =1 for image elements of object and g(i,j) =0 for image element of background. If the difference average pixel value is smaller than a certain threshold value, then the output image background will be white (that is pixel value is 255). Otherwise the background will be black (pixel value is 0). Now the pixels are classified and the motion is tracked in the video.

Algorithm

1) Search all the pixels f(i,j) of the image f.
2) Set the threshold value based on image subtraction pixel value.
3) An image element g(i,j) of the segmented image is an object pixel if f(i,j)>= T and is a background pixel otherwise.
4) Extract the foreground object without background.

Finally the segmented objects are extracted without background. Suppose in foreground detection, if background is consider as foreground means then morphological filtering is used where erosion and dilation are performed to minimize the effects of noise and improve the detected regions.

Erosion

Erosion ϴ combines two sets using vector subtraction operation of set elements and is the dual operator of dilation. Neither erosion nor dilation is an invertible transformation.

$$X \ominus B=\{p \in Z^2: p+b \in X \text{ for every } b \in B\} \qquad (4)$$

This formula says that every point p from the image is tested.The result of the erosion is given by those points p for which all possible p+b are in X.

Example

The point set X eroded by the structuring elements B as,

X={(1,0),(1,1),(1,2),(0,3),(1,3),(2,3),(3,3),(1,4)}

B={(0,0),(1,0))

XϴB= {(0,3),(1,3),(2,3)}

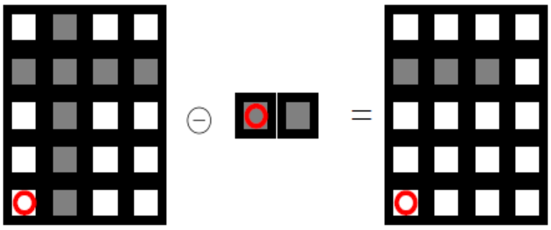


Fig.3 Example for erosion

Dilation

Dilation combines two sets using vector addition (or Minkowski set addition, e.g., (a,b)+(c,d) = (a+c, b+d)). The dilation X⊕B is the point set of all possible vector addition of pairs of elements, one from each of the sets X and B.

$X \oplus B = \{p \in Z^2 : p = x+b,\ x \in X \text{ and } b \in B\}$ (5)

Example

The point set X dilated by the structuring elements B as,

X={(1,0),(1,1),(1,2),(2,2),(0,3),(0,4)}

B={(0,0),(1,0)}

X⊕B={(1,0),(1,1),(1,2),(2,2),(0,3),(0,4),(2,0),(2,1) ,(2,2) ,(3,2) ,(1,3) ,(1,4)}

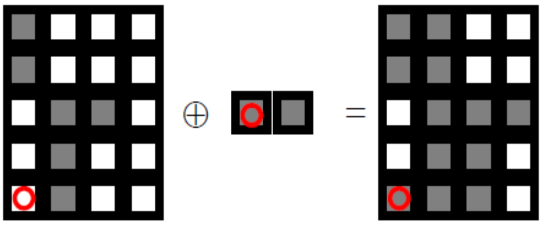


Fig.4 Example for dilation

After the object detection in video the accuracy measures done for output performance. For that the ground truth images are maintained in separate folder for measuring the accuracy. The sample ground truth images are,

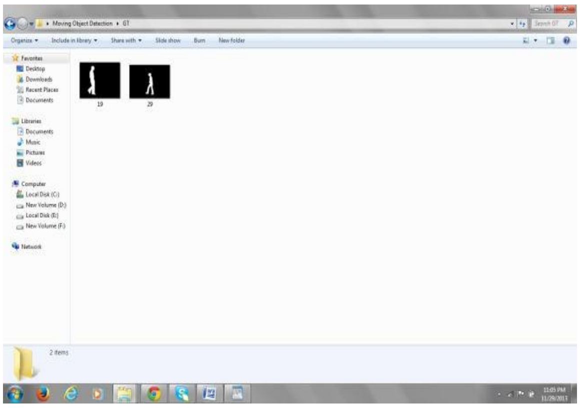


Fig.5 Example for ground truth images

Then the measures like Root Mean Square Error(RMSE) and Peak Signal to Noise Ration(PSNR) are measured.

The formula to measure RMSE as,

RMSE=sqrt(sum(sum(a-b).^{2}))./(r*c) (6)

The formula to measure PSNR as,

$$PSNR=10*\log10(255.^2/RMSE) \quad (7)$$

Where a represents a segmented image and b represents ground truth images and r and c are size of the segmented images.

By comparing the segmented images with ground truth image the sensitivity measure is evaluated in percent. If the sensitivity measure is low then the performance of detected object is minimum or else the performance is good.

IV. Object Classifications

In video analysis the second step is object classification. In proposed work the detected object is classified as whether it may be human being or not. It will base on pattern matching techniques. Patterns are nothing but arrangement of descriptors with certain properties.

A. Pattern Matching

The prior shape patterns for objects under different classes are maintained separately for identify the segmented object. Based upon the shape retrieved from the segmented object and compare that shape with predefined pattern datasets under the different object classes. The mean shapes of human beings ten classes in data sets will be given below.

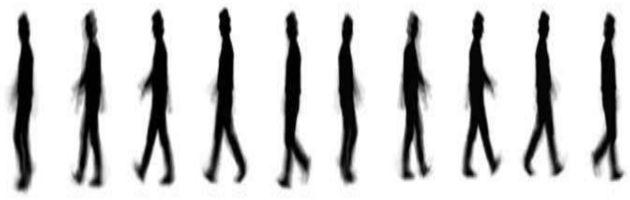


Fig.6 Walk posture mean shapes of ten classes in data sets

Then based upon arm posture of human beings also it will get classify. The data sets with different arm posture are given below for sample.

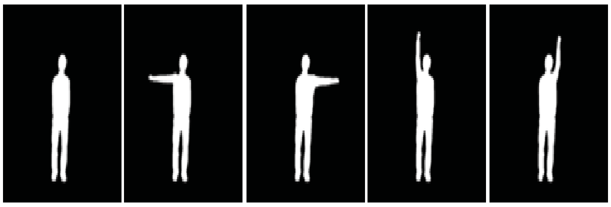


Fig.7 Basic arm posture of human beings in data sets

If it is matches then those objects belongs to particular class and finally get identified in which class the object was. And identify the segmented object belongs to human or not. These data sets are collected from external sources.

V. Object Analysis

The parameters like object count are measured based on connected component analysis and parameter like speed, velocity are evaluated.

A. Object Counting

Connected components analysis (CCA) is used to count the objects in an image. It will scan an image and groups its pixels into components based on pixel connectivity. Once all groups have been determined, each pixel is labeled with a gray level or a color labeling according to the component it was assigned to. E.g. Consider the simple scene which is given below in that *8* objects yield *8* different classes.

Fig.8 Example for CCA. (a) object grouped with color labels. (b) object grouped with gray level

B. Parameter Measures

The speed and velocity will be evaluated as,

Speed: The speed of object will be determined by ratio between distance travelled by object and total travel duration.

Speed = Distance travelled / Time. (6)

Distance=Sqrt($(x2-x1)^2 + (y2-y1)^2$). (7)

X_2 - previous pixel position.

X_1 - current pixel position in width.

Y_2 - previous pixel position.

Y_1 - current pixel position in height.

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

Velocity = Distance travelled / Frame rate. (8)

VI. Results and Analysis

This method was tested on several video sequences. All images were of size 324×244 with dynamic background. The proposed system was able to detect and human identification in almost all of the moving objects in the sequences. And it is implemented with the help of MATLAB tool and run on desktop PC with a 3.09GHz Intel and 1.90 GB RAM.

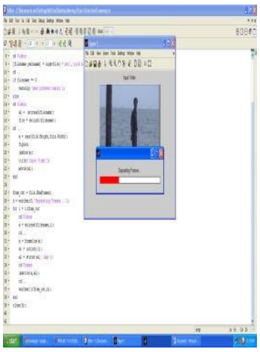
Figure 9 shows a sample input video is uploaded with dynamic background and the frame conversion is applied and finally the output is obtained and parameters are evaluated.

In matlab we have implemented this image subtraction, threshold and pattern matching algorithm. In the figure 9, the man is moving behind the tree it will detect and track movement finally classified as human or not. Then we calculated speed and velocity of object using matlab command and then found out the distance according to reference point.

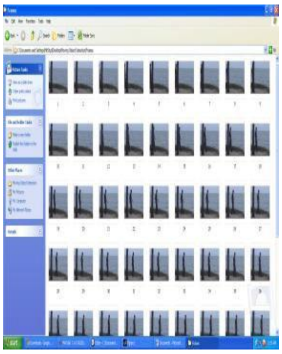
The detection output is more accurate and thus can significantly improve the performance of video surveillance based applications.

 ISSN: 2321-7782 (Online)

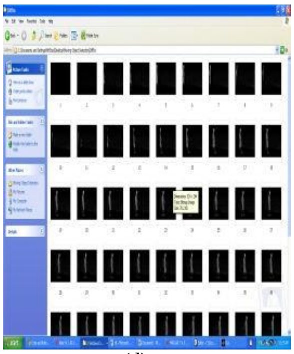
(a)



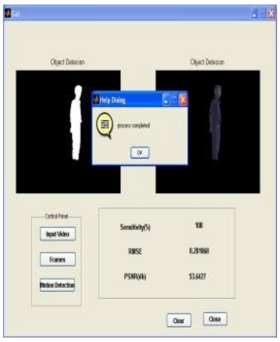
(b)



(c)



(d)



(e)

Fig.9 Sample Output Result. (a) Sample input video. (b) Frame conversion process. (c) Converted frames (d) Image Subtraction. (e) object detection with accuracy.

VII. Conclusion and Further Work

This paper implemented a novel efficient algorithm for object detection and classification in video for surveillance applications. The strength of our approach lies in the ability to separate background and foreground in accurate and implemented in simple way with low time consumption and less noise, and proved effectively in both static and dynamic background texture scenes.

The velocity and speed of the moving object can also be calculated. In future work, we will incorporate the complex video like disaster videos for object detection and classification present in the scene for more sophisticated vision based applications like fire accident, earthquake disaster etc.

Acknowledgement

I convey my sincere thanks and in depth gratitude to my guide Mr.R.Raj Bharath M.E., Assistant Professor (CSE) of Manakula Vinayagar Institute of Technology for her valuable guidance of my work.

References

1. Xiaowei Zhou, Can Yang, and Weichuan Yu, "Moving Object Detection by Detecting Contiguous Outliers in the Low-Rank Representation," IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 35, No. 3, March 2013.
2. Fei Chen, Huimin Yu, and Roland Hu, "Shape Sparse Representation for Joint Object Classification and Segmentation," IEEE Transactions On Image Processing, Vol. 22, No. 3, March 2013.
3. F. Lecumberry, A. Pardo, and G. Sapiro, "Simultaneous object classification and segmentation with high-order multiple shape models," *IEEE Trans. Image Process.*, vol. 19, no. 3, pp. 625–635, Mar. 2010.
4. Lijing Zhang, Yingli Liang "Motion human detection based on background subtraction," IEEE Trans. Second International Workshop on Education Technology and Computer Science, March 2010.
5. J. Mike McHugh, Janusz Konrad, Venkatesh Saligrama, and Pierre-Marc Jodoin, "Foreground-Adaptive Background Subtraction," IEEE Signal Processing Letters, Vol. 16, No. 5, May 2009.
6. Bineng Zhong, Shaohui Liu, Hongxun Yao, Baochang Zhang, "Multi-Resolution Background Subtraction For Dynamic Scenes," IEEE Trans., Nov. 2009.

7. Daniel Cremers, Stanley J. Osher, Stefano Soatto, “Kernel Density Estimation and Intrinsic Alignment for Shape Priors in Level Set Segmentation,” International Journal of Computer Vision 69(3), 335–351, 2006.
8. A. Mittal and N. Paragios, “Motion-Based Background Subtraction Using Adaptive Kernel Density Estimation,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, June 2004.
9. F. Ranchin, and F. Dibos, “Moving objects segmentation using optical flow estimation”, in *Proc. Workshop on Mathematics and Image Analysis*, Paris, France, Sep. 2004.
10. Liyuan Li, Weimin Huang,Irene Yu-Hua Gu, and Qi Tian, “Statistical Modeling of Complex Backgrounds for Foreground Object Detection,” IEEE Transactions On Image Processing, Vol. 13, No. 11, November 2004.
11. Bir Bhanu and Jing Peng, “Adaptive Integrated Image Segmentation and Object Recognition,” IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 30, No. 4, November 2000.
12. Chris Stauffer, W.E.L Grimson, “Adaptive background mixture models for real-time tracking,” IEEE Trans on Computer Vision, 1999.
13. Ahmed S. Abutaleb, “Automatic Thresholding of Gray-Level Pictures Using Two-Dimensional Entropy,” Computer Vision, Graphics, And Image Processing 47, 22-32 (1989).
14. Milan Sonaka, Vaclav Hlavac, Roger Boyle “Image Processing, Analysis, and Machine Vision” Second Edition, PWS Publishing.
15. R. C. Gonzalez University of Tennessee, Richard E. Woods MedData Interactive, “*Digital Image Processing,”* Second Edition, Prentice Hall.