# A Novel Statistical Method for Detecting and Segmenting Moving Objects in Video

A. Syed Kalith[1], D.Mohanapriya[2], Dr.K.Mahesh[3]

[1] *M. Phil Scholar, Department of Computer Applications ,Alagappa University, Karaikudi, Tamil Nadu.*
[2]*Ph.D Scholar, Department of Computer Applications, Alagappa University, Karaikudi, Tamil Nadu.*
[3]*Professor ,Department of Computer Applications , Alagappa University, Karaikudi, Tamil Nadu, India*

## ABSTRACT

*Video segmentation plays an important role in the MPEG-4 standard for multimedia applications. Segmentation of videos into their respective foreground and background articulate its importance in Video compression, human-computer interaction, video editing and manipulation etc. Video sequences are converted into frames and processing is done. The key perspective consider in this paper is the moving object detection with noise reduction. The video segmentation is done by detecting the moving objects on each frames and then labeling on it. A hybrid algorithm is proposed that quickly and efficiently extract the moving objects from the video sequence. Background difference is involved so as to acquire the initial object masking and to solve the uncovered background problem in the frames. The combination of noise reduction and background difference will yield the moving object within the video sequences with accuracy. The proposed algorithm is evaluated with varying input video sequences and results are produced. The experimental results show the method defers low computational complexity and better results in real time applications.*

**Keywords:** - *MPEG-4 Standard, Video Segmentation, Filtering, Noise Removal, Background Difference, Object Detection*

## 1. INTRODUCTION

MPEG-4 standard is the standard used by many multimedia applications.MPEG-4 is a method of defining compression of audio and visual (AV) digital data. This is the standard designed for a group of audio and video coding format .MPEG-4 standard provides solution for coding of natural or synthetic video and audio or multimedia files. This video coding standard MPEG-4 relies on a content based representation of multimedia files. In recent years, the paradigm of video coding has shifted from that of frame-based approach to a content-based approach, particularly with the finalization of the IOS multimedia coding standard MPEG-4. MPEG-4 defines syntax for a set of content based functionalities, namely, content– based interactivity, compression and universal access. However, it does not specify how the video content is to be generated. To generate the video content, video has to be segmented into video objects and tracked as that transverse across then video frames. so, the video segmentation is used to extract the content of the video .

Video segmentation has been a key technique for visual information extraction and plays an important role in digital video processing, pattern recognition, and computer vision .A wide range of video-based applications will benefit from advances in video conferencing, and personal entertainment. Video segmentation is the process of segmenting or partitioning the videos into meaningful shots. Shot is the collection of similar consecutive frames, where frame is the single still of the video .Though video segmentation the contents can be extracted easily where it makes the process of indexing and retrieval of video, video summarization, detection of objects that can serve as cues for event recognition; region based coding process efficient in the MPEG-4 standard.

Video segmentation can partition the videos in many ways where this proposed video segmentation process involves the extraction of foreground objects from the background of the video. Segmenting the videos can be in three ways: Image segmentation based video segmentation, Edge information based video segmentation and Change detection based video segmentation. Image segmentation based video segmentation involves first apply image segmentation algorithms such as watershed transform and color segmentation each frame to separate a frame into many homogeneous regions. By combining motion information desired with motion estimation , optical flow or frame difference ,regions with motion vectors different from the global motion are merged as foreground regions

algorithms ,often can give segmentation results with accurate boundaries , but the computation load for image segmentation and motion information.

Change detection based segmentation algorithms [1], threshold the frame difference to form change detection mask. Even though the processing speed is high, it is not robust. It suffered from shadow, noise, light changing. Edge information based algorithm, implemented here first apply canny edge detector to find the edge information of the frame and keep tracking it for finding the edges for the video or the Consecutive frames. Then secondly the morphological motion filter is used to find the edges of the foreground objects with the motion information. Next, a filling technique is used to fill the foreground moving objects to get the final object mask. Then finally the post processing is implemented to enhance the final object mask. Now the object is extracted successfully from the background without lose of the concerned edge of object. The works done related to this paper has been discussed in the section 2.

## 2. LITERATURE SURVEY

Panagiotis Sidiropoulos et al.[1] have proposed a technique, where the low-level and high-level features extracted from the visual and the aural channel have been used jointly. The proposed technique has been built upon the renowned method of the Scene Transition Graph (STG) for overcoming the difficulties of existing scene segmentation techniques. Firstly, a STG approximation has been introduced for reducing the computational cost, and then the unimodal STG-based temporal segmentation technique has been extended to a method for multimodal scene segmentation. The latter has exploited the results of numerous TRECVID-type trained visual concept detectors and audio event detectors using a probabilistic merging process that merges several individual STGs while at the same time reducing the need for selecting and adjusting many STG construction parameters. Their proposed approach has been analyzed using three test datasets, such as TRECVID documentary films, movies, and news-related videos.

Kuo Liang chungi et al.[2] have developed a promising predictive watershed-based video segmentation algorithm using motion vectors. The proposed algorithm has much computation-saving merit because the next frame could take over the segmented results of the current frame based on the motion vector information of the next frame. Their proposed algorithm has achieved a better execution-time performance as compared to the existing Chien et al.'s video segmentation algorithm whose input video sequence was assumed to have no motion vector information. Moreover, the mitigation of the over-segmentation problem occurred in both algorithms has been examined.

Chasanis et al. [3] have proposed herein, local invariant descriptors were used to represent the key-frames of video shots and a visual vocabulary was created from these descriptors resulting to a visual words histogram representation (bag of visual words) for each shot. A key aspect of our method was that, based on an idea from text segmentation, the histograms of visual words corresponding to each shot were further smoothed temporally by taking into account the histograms of neighboring shots. In this way, valuable contextual information was preserved. The final scene and chapter boundaries were determined at the local maxima of the difference of successive smoothed histograms for low and high values of the smoothing parameter respectively. Numerical experiments indicate that our method provides high detection rates while preserving a good tradeoff between recall and precision

The algorithm proposed by Tripty Singh et.al, [4], a method for moving object detection, in high secured arenas, which is deployed both in the static position and in the dynamic position is provided. Author's main objective is to increase the efficiency of the moving object detection both at the online and offline processing of video sequences. Their implementation is carried out by reducing the noise at background, calculating frame difference and producing snapshots for every sequence.

In the method proposed by Yasira Beevi and S.Natarajan [5] a video segmentation algorithm for MPEG-4 camera system by means of change detection, background registration methods and real time adaptive threshold techniques are analyzed. Their algorithm provides better segmentation results with low computation complexity. It uses a shadow cancellation mode, that can able to handles both light changing effect and shadow effect. Moreover, the algorithm has also applied real time adaptive threshold techniques through which the parameters can be determined automatically.

## 3. PROPOSED WORK

In video segmentation, the extraction of individual objects from the frame is very crucial issue and therefore our proposed video segmentation method extracts the objects from each frame. Also the proposed technique segments both the dynamic and static foreground objects without considering global motion. The motion segmentation process is carried out by both the frame difference algorithm and intersection method subsequently the most common and accurate segmented objects are retrieved from both the segmented results whereas the static foreground are

segmented using the intersection of consecutive frames. Here we use Fuzzy fusion method for extracting features from the object. The below figure shows the process used in our proposed method.

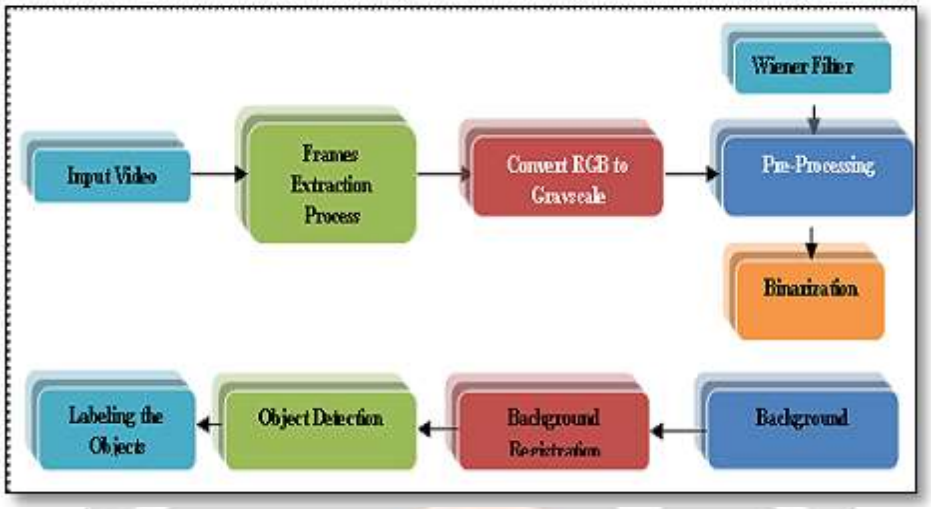


**Fig 1** Proposed Diagram Flow

## 4 METHODOOGIES

### 4.1 Video Acquisition

The proposed algorithm considers the AVI (Audio – Video Interleave) file as the input file. The AVI file format is capable for storing the audio as well as the video file container under Resource Interchange File Format. Basically the video file is compressed since it holds large space for storage. The input file is converted into respective frame sequences. The frame rate is usually 15 fps in order to reduce the frame size of the video clip involved. This single video frame is again divided into multiple frames to reduce the computational complexity of the algorithm

### 4.2 Frame Extraction Process

The input file thus separated to frames with 15 fps is applied for extraction process. Generally, the object detection in dynamic environment is the process of locating or detecting the moving object within the given video frame. This will change continuously from frame to frame. There placed much method in literature such as inter- frame subtraction, optical flow method etc., to detect the moving objects between the frames. In the proposed algorithm, the background subtraction and registration method is applied for the frames.

### 4.3 Conversion of RGB to Grayscale

Video segmentation is categorized into two major types named as spatial and temporal segmentation. The digital image processing falls under the spatial segmentation. The video sequence is collected in static background, with RGB color pixels. For accurate results the RGB pixel is separated to autonomous red channel, green channel and blue channel. After the respective separation, the conversion to gray scale is taken place. All the frames are completely converted into grayscale intensity images. Gray scale images are acquired without color, and levels from 0. The conversion takes place to reduce the complexity when the segmentation takes place.

### 4.4 Pre – Processing

The Pre – Processing is the step where the noise on the image should get removed. The method for removing the noise from video signal is termed as video de-noising. Because of the noise in images, it will disturb and there is a chance for image quality degradation. Hence, the noise reduction should be taken place so as to enhance the video quality. Based on the interpretation of the pixels the noise level varies. Mainly two type of filtering algorithms were imposed to remove noise from the images, named as linear and non-linear filtering. The background noise present in the frames is reduced using filter technique in the video sequences. In the proposed algorithm the “Weiner Filter: is used for de-noising. The Wiener filter is applied to filter out the noise that has corrupted a signal. The Wiener filter approaches filtering from a different angle. Wiener filters are characterized based on the Performance criteria such as minimum mean -square Error, assumption and requirement.

### 4.5 Binarization

Binarization is a technique by which the gray scale images are converted to binary images. Binarization separates the foreground text to the background information. The most common method for binarization is to select a proper threshold for the intensity of the image and then convert all the intensity values above the threshold to one intensity value and all intensity values below the threshold to the other chosen intensity (i.e From White to Black). The noise free image is now through binarization converted to binary image without noise.

### 4.6 Background Difference and Registration

The next step with the noise reduced binary image is the computation of background. Once images are in use, the algorithm performs a background subtraction of the image to isolate the images and create a mask. The background subtraction involves in two steps. First, is to subtract the pixels between the foreground and background objects channel wise. They are summed up and threshold is calculated. The key step algorithm is that, the background modeling and updating. The basic idea of the background difference method is to do differential between current image and background image. When the pixels values between moving target and the background are not very close, the complete moving target can be detected. The method is very sensitive to light and shadow objects. The frames with the background difference are registered and summed up to indicate the foreground and background objects that are above the threshold.

### 4.7 Object Detection

The difference between the successive frames along with the sensitivity is calculated in the background difference step. The values that is greater than the threshold is detected in this step.

### 4.8 Object Labeling

The final step is the object labeling step. The threshold values that are calculated in the above steps are labeled according to the shape based retrieval. Shape based image retrieval is the measuring of similarity between shapes represented by their features. Shape content description is difficult to define. Feature extraction and similarity measurement between the extracted features are the two steps taken place in similarity measure calculation. Shape descriptors can be divided into two main categories: region based and contour-based methods. In the proposed system the labeling is made through the rectangular box.

## 5. RESULT AND DISCUSSION

The Experimentation data consists of a different videos collected from various sources. The video is preprocessed by de-noising technique using the wiener filter then the segmentation has been applied on the video frames of individual moving object detection and labeling is done. Noise reduction is done as the pre process in video in order to improve its quality. The results show the moving object through the rectangular box shape. under the region based method that uses complete area for the shape description.

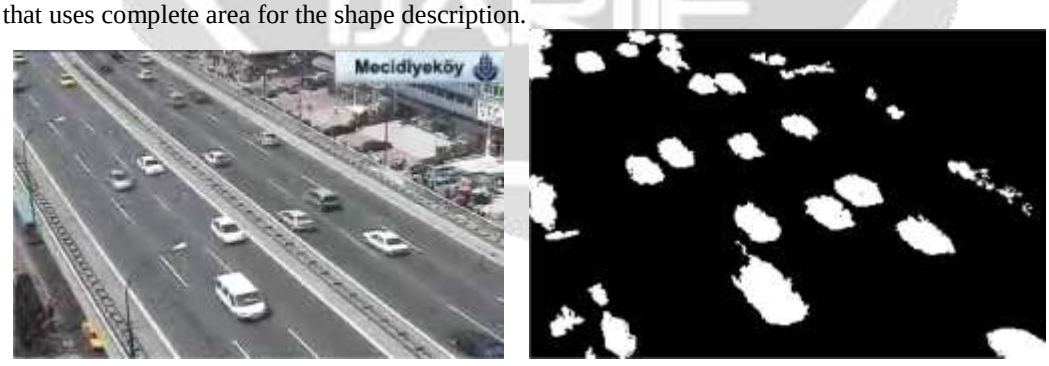


**Figure 2. Moving Object Detection and Segmentation**

## 6. CONCLUSIONS

The proposed research work, gives a hybrid algorithm for video segmentation based moving object detection. It uses the filtering technique to remove noise from the frames and background difference method to calculate the threshold. Wiener filter is best suited for Gaussian noise to remove. The implementation is carried out by reducing

the noise in the background, checking frame difference and generating the labeling for each object that are detected. Hence, in this paper a new hybrid algorithm for detection of moving objects in video is designed and implemented, which is useful in various applications like security and video surveillance etc.

## 7. REFERENCES

[1] Panagiotis Sidiropoulos1, Vasileios Mezaris, Ioannis Kompatsiaris, Hugo Meinedo, Miguel Bugalho, Isabel Trancoso,"Video scene segmentation system using audio visual features", Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), 2010.

[2] Kuo liang chung , Yan syun lai and Pei-ling huang,"An Efficient Predictive Watershed Video Segmentation Algorithm Using Motion Vectors", Journals of information science and engineering, vol.26, no.699-711, 2010.

[3] Vasileios Chasanis, Argyris Kalogeratos, Aristidis Likas, "Movie Segmentation into Scenes and Chapters Using Locally Weighted Bag of Visual Words", ACM International Conference on Image and Video Retrieval, 2009.

[4] Tripty Singh, Sanju S and Bichu Vijay, "A New Algorithm Designing for Detection of Moving Objects in Video", in International Journal of Computer Applications (0975 – 8887) Volume 96– No.2, June 2014, pp:4-11.

[5] Yasira Beevi and Natarajan. S, "A Hidden Markov Model Framework for Video Segmentation Using Audio and Image Features", citeseerx.ist.psu.edu/viewdoc/ .

[6] D.Mohanapriya,Dr.K.Mahesh "A novel foreground region analysis using NCP-DBP teture pattern for robust visual tracking" , Springer Multimedia Tools and Appications –An International Journal, Volume: 76 Issue No: 24, December 2017. pp:25731-25748.

[7] D.Mohanapriya, Dr.K.Mahesh,"Robust Video Tracking System with shadow suppression based on Feature Extraction",Australian Journal of Basic and Applied Sciences, Vol.10, No.11 (July), 2016 pp 307-311.

[8].D.Mohanapriya and Dr.K.Mahesh, "A video target tracking using shadow suppression and feature extraction," IEEE Xplore Digital Library.

[9] D.Mohanapriya, Dr.K.Mahesh "Robust Visual Target Tracking Via Nearest Sequential Boundary Pattern" International Journal of Pure and Applied Mathematics Volume 118 No. 7 2018, 207-213.

[10] D.Mohanapriya, Dr.K.Mahesh, "Video Tracking System by suppressing shadow and Feature Extraction- A Review", International Journal of Computer Engineering and Applications(IJCEA), Vol.10, No 6 ,June , 2016

[11] K.Rajalakshmi, Dr.K.Mahesh, "Video Steganography Based on Embedding the Video Using PCF Technique "Proceedings of IEEE Conference (ICICES 2017) 978-1-5090-6135-8/17 ©2017 IEEE.

[12] K.Rajalakshmi, Dr.K.Mahesh,"ZLBM:Zero Level Binary Mapping Technique for Video Security" Multimedia Tools And Applications Springer, DOI 10.1007/s11042-017-4942-0, ISSN:1380-7501,Online:13 July 2017.

[13] K. Rajalakshmi, K. Mahesh, "Video Embedding with Compression Based on Patchwise Code Formation", Australian Journal of Basic and Applied Sciences, vol. 10, no. 13, pp. 2309-8414, August 2016, ISSN 1991-8178.