INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

REVIEW PAPER ON MOVING OBJECT SEGMENTATION METHODS

Abhilasha[1], Bhawna Chauhan[2]

[1]M.Tech scholar, B.S.Anangpuria Institue of Technology & Management,

abhilashasangwan58@gmail.com

[2]Assistant Professor, B.S.Anangpuria Institute of Technology & Management,

bhawna.chauhan@faculty.anangpuria.com

Abstract

Segmentation of moving object in video sequence is important for many computer vision applications. Moving object segmentation is extraction of foreground from background. Segmentation includes steps as object detection and motion detection. Currently one of the most active research topics in computer vision is human and vehicle segmentation. Object segmentation is very useful for tracking object and recognition of the object in a video. Common methods for segmenting the moving objects are background subtraction, temporal segmentation, edge detection, contour and the combination of temporal-spatial segmentation. These methods are explained in this paper along with their advantages and disadvantages. These methods are used in animation, medical imaging, automatic surveillance –monitoring a scene to detect suspicious activities in shopping malls, offices, human computer interaction which include eye gaze tracking for data input to the computer.

Key terms – Segmentation, background subtraction, frame difference, contour, threshold, edge detection

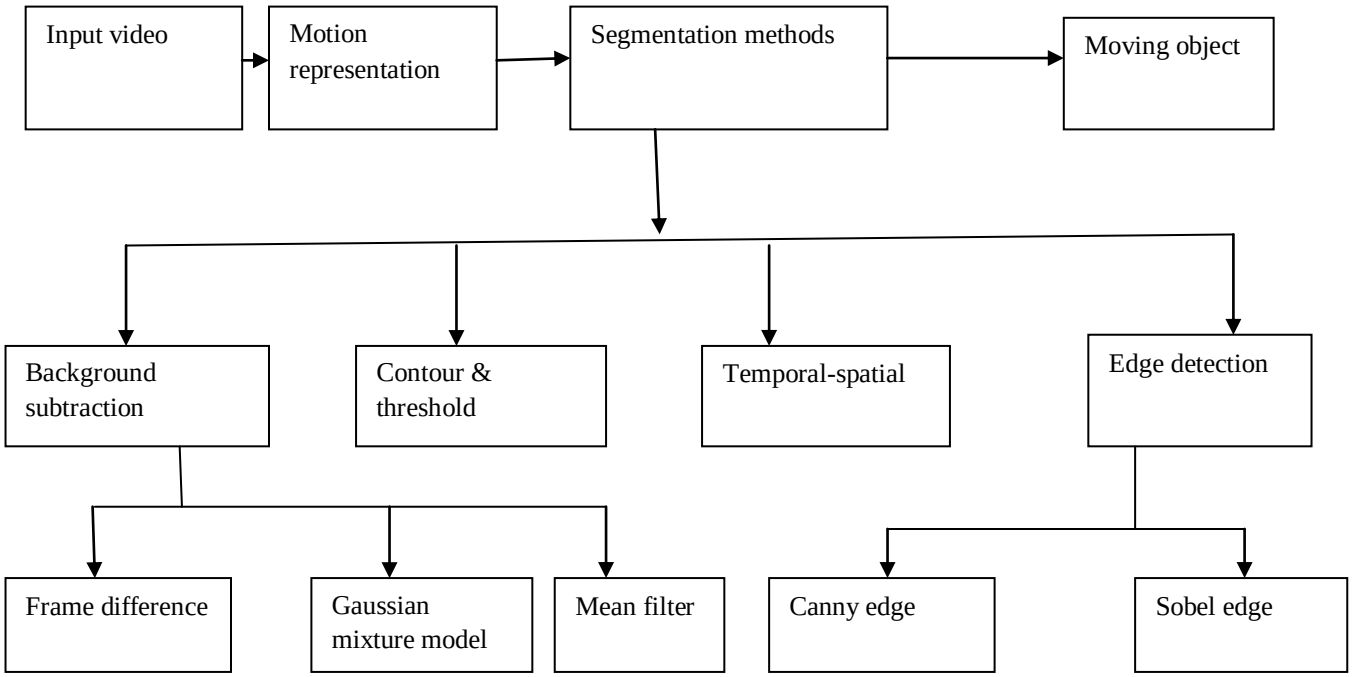
1. Introduction

Segmentation of objects in video is very important in many aspects of multimedia applications, such as traffic monitoring, human motion capture and video surveillance. Image segmentation is the partitioning of a image in objects of interest. There are several issues that need to be solved in motion segmentation particularly on noise, missing data and lack of a priori knowledge [2]. Presence of noise is one of the main problem. For some applications noise level can become serious. Some of the common issues that segmentation algorithm encounters are rapid light attenuation, strong reflections; back-scattering, non-uniform lighting and dynamic lighting that dramatically degrade the quality of the images [2]. Blurring is also a issue particularly when motion is involved. Another common difficulty is caused when moving objects create occlusions, or even worst, the whole object disappear and reappear in the scene [2]. Tracking and recognition of object in a video can be done with the help of the segmentation

methods such as background subtraction, temporal segmentation, edge detection and spatial segmentation. Background subtraction is a very popular technique for detection of moving object in a frame. It is used for separating foreground object from the background. In this method first the background is modeled and then foreground object is detected by subtracting the foreground scene from the background scene .Change in illumination of the environment, rippling water and waving trees are the problems that are faced during the background subtraction. Temporal video segmentation is first step towards automatic explanation of digital video sequences. Its goal is to divide the image stream into a set of significant and manageable segments that are used as basic ground rules for indexing. Each shot is represented by selecting Key frames and indexing by extracting spatial and temporal features [3].In edge information based video segmentation, first canny edge detector is applied to find edge information of each frame and then keep tracking these edges [4].Motion segmentation algorithm must be able to deal with noisy images, occlusions, missing data and multiple objects in a scene.

2. Methods of segmentation

Segmentation of moving objects can be done by various methods.

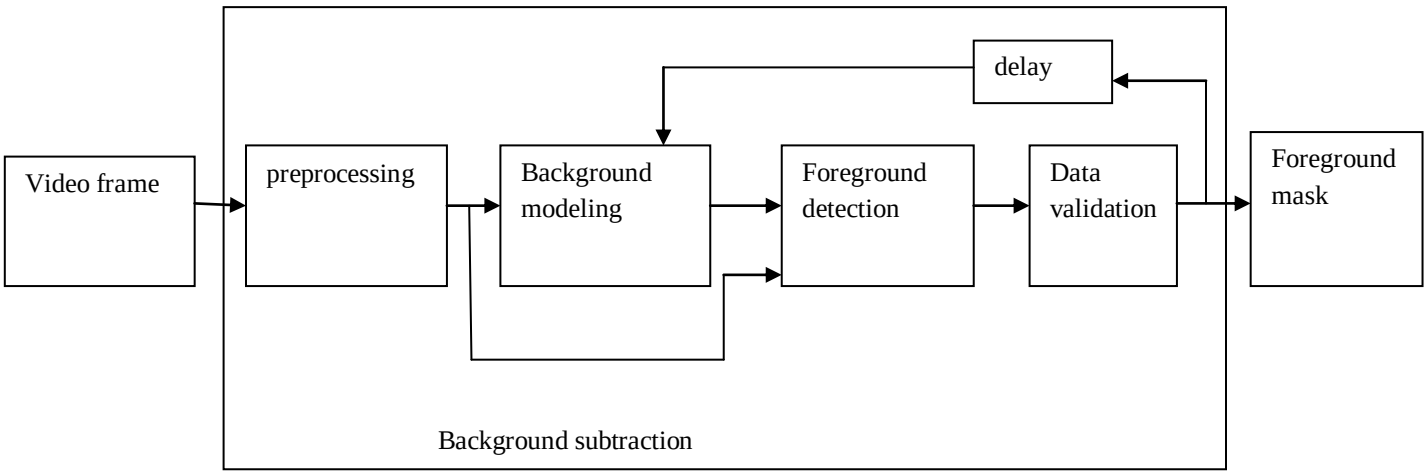


Description of various segmentation methods:

2.1 Background subtraction

Background subtraction is a widely used approach for detecting moving objects from static cameras. Identification of moving object from a video sequence is a very critical and important task. It is also called as foreground detection. It is done when image of interest is a part of video stream. Background subtraction is used for separating foreground object from the background. In this method first the background is modeled and then foreground object is detected by subtracting the foreground scene from the background scene. The name "background subtraction" comes from the simple procedure of subtracting the current image from the modeled image and then threshold the result to generate the foreground objects [1].Change of illumination of the environment, rippling water and waving trees are the problems that are encountered during background subtraction.

Background subtraction algorithm:



There are four steps under background subtraction method:

a) Pre-processing

b) Background modeling

c) Foreground detection

d) Data validation

Background subtraction generally consist of four steps.

Preprocessing , background modeling , foreground detection and data validation .

a) Preprocessing is a collection of tasks that are applied on raw data video in order to convert raw input video into a format that can be processed further properly.

b) In background modeling, background is modeled from the preceding scenes.

c) In foreground detection. Foreground object is detected by performing subtraction .

d) In data validation, extra pixels are eliminated that are not a part of moving object and it then outputs the foreground mask.

Background subtraction approaches can be classified as-

2.1.1) Frame difference.

2.1.2) Mean filter.

2.1.3) Gaussian mixture model.

2.1.1) Frame difference:

In this approach difference between two frames is calculated.

$D(t+1)= |V(x,y,t+1) – V(x,y,t)|$

Where x,y are the pixel location variable and t is the time dimension .

In this threshold function plays a very important role. Threshold function is used in order to identify whether a pixel is a foreground pixel or background pixel.

This approach provides fast recovery.

2.1.2) Mean filter:

In this background scene is calculated by taking average of a series of preceding images.

$$B(x,y) = 1/N \quad \sum_{i=1}^{n} V(x,y,t-i)$$

Where N is the number of preceding images which are taken for averaging.

Once the background scene is calculated, subtract it from the foreground scene and then apply threshold to it.

If $|V(x,y,t) – B(x,y)| > Th$

Then the pixel is considered as foreground pixel otherwise it is considered as background pixel .

2.1.3) Gaussian mixture model

Background modeling is a very important step for all background subtraction algorithm. Background model is enough to identify all moving objects in a sequence of video frames. In this approach, the values of a particular pixel are showed as a mixture of Gaussians. At each Iteration Gaussians are evaluated, in order to determine which one is mostly likely compared to the background. Pixels that do not match with the "background Gaussians" are classified as foreground pixels. Noise in the image can be removed by this approach.

2.2) Contour and threshold method:

2.2.1) Contour method

Active contour model is also called as snakes. It is a framework for delineating an object outline in a video sequence of frames. It has a representation of bounding contour that is updated over time . Initial ACM (active contour model) investigation is a very important step of moving target segmentation and image difference method is used for this purpose. Once the moving regions in the scene are found out, the edges of moving regions can be used as the initial ACM [6]. The contour method is useful for many segmentation algorithms to get moving object clearly. For example: In background subtraction method counters are used to get the foreground object clearly. Existing active contour models can be categorized into edge-based models and region-based models. Region-based models are more robust when weak edges exist and less sensitive to initialization, but more computationally expensive than edge-based models. Active contour models are commonly used in image segmentation problems, especially for medical images with a lot of noise. Object with clear background can be segmented using this method.

2.2.2) Threshold method:

Threshold is one of the methods used for image segmentation. The simplest method of Object segmentation is called as thresholding method. It is useful in separating foreground from the background. The basic step of this method is to select the threshold value. By selecting a threshold value T, the grey level image can be converted to binary image. The binary image should contain all of the necessary information about the position and shape of the objects of interest. The advantage of obtaining first a binary image is that it reduces complexity of the data and simplifies the process of recognition and classification. The purpose of threshold algorithm is to segment the motion objects from the background pixels according to their gray value differences. Clear segmentation of moving object is done by this method.

2.3) Temporal and spatial method

Main objective of Spatial-temporal segmentation is to create a layered image representation of the video for image coding applications whereby video data is just described as a set of moving layers [7]. For example: In Recognizing and quantifying human movement requires spatial segmentation followed by temporal segmentation. The spatial segmentation method is effectively a tracking process which determines a motion vector encapsulating a set of joint angles for each frame. The temporal segmentation is a CHMR (Continuous Human Movement Recognition) system which attempts to understand the movement skill that could have produced the observed sequence of motion vectors [8]. A shot may be defined as an unbroken sequence of frames taken from one camera. In this method the segmented objects are dirt-free.

2.4) Edge detection:

Edges are major local changes of intensity in an image. Edges in general occur on the boundary between two different regions in an image. Main goal of edge detection is to produce a line drawing of a scene from an image of

that scene. The important features can be extracted from the edges of an image. These features are then used by higher-level computer vision algorithms.
Edges can be classified as step edge, ramp edge, roof edge and ridge edge. Edge detection techniques are also used as the base of another segmentation technique. To segment an object from a video, one needs closed region boundaries. Looked-for edges are the boundaries between such objects. Segmentation methods can also be applied to the edges obtained from edge detectors. Edge information plays a important role in the segmentation process. For tracking and recognizing the object from the video these methods are used with any segmentation technique. Segmenting moving objects is a challenging task in computer vision.

The four steps of edge detection
(1) Suppress as much noise as possible, without destroying the true edges.
(2) Apply filter to enhance the quality of the edges in the image.
(3) Find out which edge pixels should be discarded as noise and which should be retained.
(4) Find exact location of an edge. Edge thinning and linking are required in this step.

Some of these techniques that come under edge detection:
2.4.1) Canny edge detector.
2.4.2) Sobel edge detector.

2.4.1) canny edge detection:

This is perhaps the most widely used edge detector in computer vision. Canny has shown that the first derivative of the Gaussian closely approximates the operator that optimizes product of *signal-to-noise* ratio and localization. Moving-edge points are generated based on the processing of the frame difference edge map, current frame edge map, and background edge map. These spatial domain edge maps are obtained using Canny edge detector, which involves Gaussian convolution to avoid noise.

The algorithm has 5 steps for image:
1. Blurring the image in order to take away the noise.
2. Find the gradients.
3. Apply Non-maximum suppression.
4. Apply thresholding.
5. Edge tracking is done by hysteresis.

2.4.2)Sobel edge detection

The Sobel operator performs 2-D spatial gradient measurement on an image and so emphasizes the regions of high spatial gradient that correspond to edges. In general it is used to find the approximate absolute gradient magnitude at each point in an input grey scale image. In contour based segmentation, edge detection techniques are the most generally used. Canny and Sobel operator are the commonly used edge detection methods due to the flexibility in performance [5]. The 2-D Sobel operator was the most well-liked edge detection operator until the development of edge detection techniques called Canny edge detection. It proved popular as it gave a better performance than other edge detection operators, such as Prewitt operator. Here there is no pre-processing or smoothing stage before the Sobel operation since we aim to observe performance of the operator with respect to increasing noise. The thresholds for the 2-D Sobel, to achieve the binary image, are determined by a root mean square (RMS) estimate of the noise. It is very simple method. Detection of edges and their orientations is done.

Table 1: Comparison table of segmentation methods

Methods	Advantage	Disadvantage
Background subtraction a) Frame difference b) Mean filter c) Gaussian mixture model	Time adapting, provides fast recovery.	Illumination changes, rippling water and waving trees create problem
Contour method	Used to track dynamic objects in spatial dimensions.	Longer computation time.
Threshold method	Can segment moving objects clearly	No guarantee that the pixel identified are contiguous.
Temporal and spatial	Segmented objects are dirt-free.	Long evaluation time
Edge detection a) Canny edge detection	Improving signal to noise ratio, better detection specially in noise conditions	Complex computation, time consuming.
Edge detection b) Sobel edge detection	Simplicity , detection of edges and their orientation	Sensitivity to noise , inaccurate

3. Conclusion

In this paper we have studied various segmentation methods. Segmentation methods can be classified as background subtraction, contour and threshold, spatio-temporal and edge detection methods. These methods are described with their advantage and disadvantage. These method are used in animation, medical field, automatic surveillance – monitoring a scene to detect suspicious activities in shopping malls , offices , human computer interaction which include eye gaze tracking for data input to the computer .

REFERENCES

[1] “Background Subtraction techniques” Alan M.

[2] “Motion segmentation: a review” Luca Zappella

[3] “Temporal video segmentation; a survey” Irena koprinska

[4] “An efficient video segmentation algorithm with real time adaptive threshold technique” Yasira Beevi

[5] “Segmentation of moving objects in image sequence: a review” Dengsheng Zhang and Guojun Lu Gippsland

[6] “Automatic Segmentation of Moving Objects By Fusing Multi-Information” Lin Zheng Chongzhao Han Dongguang Zuo Hongyan Zhu

[7] “Spatio-Temporal segmentation of video data” John Y.A. Wang and Edward H. Adelsonz

[8] “ Spatial and temporal segmentation of continuous human motion from monocular video images” Richard D. Green