Abstract: Object tracking is one of the most important components in a wide range of applications in computer vision, such as surveillance, human computer interaction, and medical imaging. In the proposed work we applied a object tracking algorithm to segmented object obtained from global motion estimation and compensation. In the segmentation of videos the main challenges occurs in the crowded areas and when the video is captured using moving camera. A fast and robust video segmentation technique, which aims at efficient foreground and background separation via effective combination of motion and color segmentation modules. Frame differencing and colour based segmentation methods were used for the identification of the objects in the videos. In the frame differencing process the current video frames were subtracted from the reference frames in order to obtain the objects in the video. The obtained objects were then segmented with the help of colour based segmentation methods. The performance of the process is measured with the help of performance metrics like Accuracy, Error rate, Precision and Recall. The proposed work will automatically segment, identify moving object and track the object in all the frames and also solve different problems in change detection like uncovered background, temporary poses, global motion of background.

Index terms: Moving Object ,Global Motion Estimation, Block Matching , object tracking, object detection, video object Segmentation

1.INTRODUCTION
Object tracking is an important task within the field of computer vision. The proliferation of high-powered computers, the availability of high quality and inexpensive video cameras, and the increasing need for automated video analysis has generated a great deal of interest in object tracking algorithms. There are three key steps in video analysis: detection of interesting moving objects, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior. Video segmentation generalizes this concept to the grouping of pixels into spatio-temporal regions that exhibit coherence in both appearance and motion. Such segmentation is useful for several higher-level vision tasks such as activity recognition, object tracking, content-based retrieval, and visual enhancement. Video clips are mostly found frame based. Raw video data is usually in the form of binary streams that are not well organized. For content representation, the raw video data must be decomposed into objects, each object representing particular meaningful content of video. The process of separating the meaningful video objects from the background in the video sequences is video object segmentation. In personal databases and web repositories use of data is continuously increasing. To extract the object from video, segmentation is often needed. The estimation of the objects can be done based on pixel based approaches and object based approaches. Motion vector estimation and flow estimation were object based approaches. Motion vector identifies the object location based on pattern of the motions in the set of video frames. Flow estimation identifies the object locations based on the gradient based estimation of the object movements in the videos. The object based approaches requires the informations regarding the camera calibrations and initial prediction of the object locations. For surveillance applications the informations regarding the camera calibrations were not always available. In the proposed work object
based approaches were employed for the identification of the objects in the video. Object tracking is posed as a binary classification problem in which the correlation of object appearance and class labels from foreground and background is modeled. Compared with state of art tracking methods, the proposed algorithm achieves favorable performance with higher success rates and lower tracking errors.

Optical flow or motion fields could theoretically be used, but they are extremely noise sensitive, and their accuracy is limited due to the aperture and occlusion problem. Change detectors or difference images, on the other hand, mark occlusion areas as changed, while the objects themselves are unchanged unless they contain sufficient texture. This makes exact boundary location difficult, and an additional mechanism is necessary to fill the holes inside objects. Our proposed algorithm is based on pattern recognition and object tracking principles, and thereby avoids many of the problems associated with motion estimation. The concept of moving connected components is introduced to enable automatic detection of moving objects, and a novel model update method allows for relatively large changes in shape. To improve the stability of the segmentation and to reduce the computational complexity, a filter is presented that removes stationary background.

2. REVIEW OF LITERATURE

A number of Object Tracking and video segmentation algorithms have been proposed for detecting and tracking object from the video. This section provides a review of the various approaches available for video segmentation and object tracking.

Andres Alarcon Ramirez1 and Mohamed Chouikha proposed a novel algorithm for automatic video object tracking based on a process of subtraction of successive frames, where the prediction of the direction of movement of the object being tracked is carried out by analyzing the changing areas generated as result of the object’s motion, specifically in regions of interest defined inside the object being tracked in both the current and the next frame [1].

Deqing Sun1Erik B. Sudderth1Michael J. Black A typical scene consists of very few moving objects and representing each moving object by a layer allows the motion of each layer to be described more simply. In this paper current layered motion models have not shown convincing layer segmentation results on challenging real-world sequences.[2]

N. A. Tsoligkas, D. Xu and I. proposed algorithm is based on pattern recognition and object tracking principles, and thereby avoids many of the problems associated with motion estimation. These projected regions are used to extract the markers for the next frame, which is then segmented by the watershed algorithm based on luminance. But drawback were it is not possible to represent that object by a single image that is warped from frame to frame.[3]

Thang Dinh* and Gérard Medioni propose a novel approach which can yield both the motion segmentation and the motion estimation in the presence of discontinuities. Note that the unit used to calculate the threshold here, is not in pixels but in points which are in sparse space. Drawback of Tensor Voting is that it suffers from a lengthy processing time.[4]

Adrian Ulges1, Christoph H. Lampert2, Daniel Keysers3, Thomas M. Breuel present two probabilistic formulations of the problem and carry out optimization using the RAST algorithm. Though greedy search procedures may be fast, attractive solutions for online processing, they do not guarantee global optimality.[5]

Vasileios Mezaris, Ioannis Kompatsiaris Michael G Strintzis proposed novel unsupervised video object segmentation algorithm is presented, aiming to segment a video sequence to objects: spatiotemporal regions representing a meaningful part of the sequence. It cannot qualify for representing a semantic object on their own.[6]

Thomas Meier and King N. Ngan proposed algorithm is based on pattern recognition and object tracking principles, and thereby avoids many of the problems associated with motion estimation. The concept of moving connected components is introduced to enable automatic detection of moving objects. But drawback found is it is not possible to...
represent that object by a single image that is warped from frame to frame.[7]

3. PROPOSED WORK
The present study aims at developing a system to segment objects from the video automatically and to track them accurately. In the proposed work we will try to solve the issue of moving camera for this we will use motion estimation and compensation.

A number of video object segmentation algorithms have been proposed for specific applications and specific requirements. Many automatic segmentation systems are designed for specific assumptions like videos with fixed background. It is necessary to have flexible automatic segmentation system for different types of videos and those which segment and track the objects from moving camera.

3.1 VIDEO:
The input video is converted into frames. Each video consist of n frames. Video Captured using moving camera creates the global motion of background. In order to compensate the global motion of background motion estimation and compensation should be applied. A robust motion detection method based on the frame difference is used to identify whether homogeneous regions are moving or not.

For each homogeneous region, if 85% pixels are identified as moving pixels, the region is identified as moving. Only for moving regions, motion vector field is estimated by hierarchical block matching methods inside the regions. The obtained parameters are tested for the whole region. When the error is above a certain threshold, the region will be split according to the motion information. Regions with similar motion and color are merged together. Moving objects are projected to the next frame according to their affine motion models.

3.2 GMEC (Global Motion Estimation & Compensation)
For video Sequences that contain camera motion, this global motion is compensated by the motion estimation and compensation.

Motion vector Estimation: Block Matching Algorithm. Block matching is analyzed as an alternative for detecting moving objects. Besides detection of moving objects, block matching also provides motion vectors (location of motion) which can aid in tracking objects.

In block matching, blocks in the current frame are matched to blocks in a reference frame (an earlier frame). The Block Matching is as illustrated in Fi. For each block in the current frame, the reference frame is searched for the best matching block. A matching criterion determines the best match from candidate blocks in the reference frame. If the matched block is not in the same location in the reference frame as in the current frame, the block has moved. A foreground mask of the moving blocks is generated. Blocks with the same motion can be combined to form moving objects. Block matching adds the additional information of block motion, making block matching attractive for tracking applications.

For the foreground extraction the main algorithm used is the Block Matching Algorithm (BMA) which is the motion estimation algorithm employed in the MPEG.

The noise is estimated first and then using BMA the motion is estimated i.e., BMA partitions the current frame in small, fixed size blocks and matches them to the previous frame, in order to estimate blocks.
displacement (Referred as Motion Vectors) between two successive frames.

There are different Kinds of Block Matching Algorithms. They are as follows:

1. Full-search block matching algorithms: A block in the anchor frame is compared to all the possible blocks that can be found within a defined search range in the target frame.

2. Fixed-search pattern algorithms: The main idea is to check only some points within the search range.

3. Multi-Resolution Motion Estimation: The main idea behind the multi-resolution method is to start from a low resolution image and to estimate the motion through different steps, each one implying a larger resolution of the picture. In this proposed work, full-search Block Matching method is used.

![Figure 3 Motion Vector Estimation by Affine Model.](image)

3.3. CHANGE DETECTION METHOD

Conversion to Gray Scale: Many approaches, such as the Change detection, require only grayscale images to perform the operations. Grayscale images are related to the luminance component in the YCbCr color space, where Y is the luminance component that gives the average brightness of the image.
image. Since working with the luminance component is enough to process images without chrominance components, conversion from the RGB color space to the YCbCr color space is done in order to work only with the Y component as a grayscale sequence of images. Working only with one component instead of three, leads to a system to be computationally more efficient since less memory resources are used. Fortunately, MATLAB has the function rgb2gray to extract automatically the grayscale image from the RGB color space. The conversion to grayscale was also useful to reduce the computational load to 1/3. Instead of processing three color components, the process is made using only the luminance component.

### Detailed Change Detection Algorithm

#### Frame/Background Difference

The differencing includes frame differencing and background differencing. In the frame difference, the frame difference between current frame and previous frame is calculated and thresholded.

\[
FD(x,y,t) = |I(x,y,t) - I(x,y,t-1)|
\]

\[
FDM(x,y,t) = \begin{cases} 
1 & \text{if } FD \geq Th \\
0 & \text{if } FD < Th 
\end{cases}
\]

It can be presented as Where I is frame data, FD is Frame Difference, and FDM is Frame Difference Mask. The previous frame is substituted by background frame. After background differencing, another change detection mask named Background Difference Mask is generated. The operations of Background Differencing can be presented as

\[
BD(x,y,t) = |I(x,y,t) - BG(x,y,t-1)|
\]

\[
BDM(x,y,t) = \begin{cases} 
1 & \text{if } BD \geq Th \\
0 & \text{if } BD < Th 
\end{cases}
\]

where BD is background difference, BG is background frame, and BDM is Background Difference Mask, respectively.

### 3.4 OBJECT DETECTION

Both of FDM and BDM are input into Object Detection to produce Initial Object Mask. The procedure of Object Detection can be presented as the following equation.

\[
IOM(x,y,t) = \begin{cases} 
BDM(x,y,t), & \text{if } BI(x,y,t) = 1 \\
FDM(x,y,t), & \text{else}
\end{cases}
\]

### 3.5 HAUSDORFF OBJECT TRACKER

#### Detailed Description of Algorithm

1. **Edge Detection by Canny Edge Tracker**
   - The algorithm runs in 5 separate steps:
     a. Smoothing: Blurring of the image to remove noise.
     b. Finding gradients: The edges should be marked where the gradients of the image have large magnitudes.
     c. Non-maximum suppression: Only local maxima should be marked as edges.
     d. Double thresholding: Potential edges are determined by thresholding.
     e. Edge tracking by hysteresis: Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

2. **Hough Transform & Hausdorff Distance**
   - The Hough transform is a technique which can be used to isolate features of a particular shape within an image. Because it requires that the desired features be specified in some parametric form, the classical Hough transform is most commonly used for the detection of regular curves such as lines, circles, ellipses, etc.

3. **Segmentation by Color based approach**
   - A new color-based approach for automated segmentation is used for more accurate object detection. The method comprises three stages: (a) color normalization to reduce the quality variation of tissue image within samples from individual subjects or from different subjects; (b) automatic sampling from tissue image to eliminate tedious and time-consuming steps; and (c) principal component analysis (PCA) to characterize color features in accordance with a standard set of training data.

### 3.6 FINAL OUTPUT

The Output Obtained from Change Detection is summed up with output of object detected in hausdroff object tracker. Criterion for motion does not come directly from the frame difference of two consecutive frames. Instead, we construct and maintain up-to-date background information from the video sequence and compare each frame with the background.
4 SIMULATIONS AND EXPERIMENTAL RESULTS
The proposed work have been done using segtrack dataset. The summing of Output Obtained from Change Detection with output of object detected in hausdroff object tracker achieve an outstanding performance in detecting moving objects in moving camera.

The Evaluation Factors used are Accuracy Error rate, Precision, Recall.

![Figure 6 Accuracy & Error rate of Object Detected](image1)

![Figure 7 Performance graph of Precision](image2)

![Figure 8 Performance graph of Recall](image3)

5. CONCLUSION
An algorithm for Block-based Video Segmentation has been implemented using block Matching and recursive tracing. It has demonstrated favourable results. Slow movements and temporary poses, where other segmentation algorithms fail to give correct segmentation, the block –based method of Video Segmentation using Block Matching Algorithm turned out to be a suitable option as it outperforms others. Block-based method is more tolerant to slow and small object motion from frame to frame.

There are various methods of video object segmentation, but the faster video object segmentation techniques are based on change detection approach. Both pixel-based and block-based algorithms implemented in this dissertation have demonstrated potential and can be used by the researchers without giving it a second thought about their computational complexity.

6. REFERENCES
Conference Proceeding Info

2nd International Conference on Recent Development in Engineering, Science, Management and Humanities (ICRESMH-16)
at Indian Federation of United Nations Associations (IFUNA) C-6, Qutab Institutional Area, New Delhi, India on 23rd October 2016


[10] Hybrid Object-Based Video Compression Scheme Using a Novel Content-Based Automatic Segmentation Algorithm. Author: N. A. Tsoligkas, D. Xu and I. French. 2010


