International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
International Conference on Communication Computing & Virtualization 2016

Automatic Segmentation of Moving Object in Video Sequences

Shubhangi Vaikole
Research ScholarDatta
Meghe COEAiroli, Navi Mumbai

S.D.Sawarkar, PhD
PrincipalDatta Meghe
COEAiroli, Navi Mumbai

ABSTRACT
In content based video retrieval and concept detection systems video segmentation is the most important step. There are basically two methods for video segmentation, one is semiautomatic and other is automatic. A lot of work is already performed on this two approaches. Semiautomatic methodsrequires the user intervention to draw the boundary of object. Many applications require automatic segmentation methods but still there is a lot of scope for research because mostly the methods are application specific. The main focus of this paper is to identify the gaps that are present in the existingvideo segmentation system and also to provide the possible solutions to overcome those gaps so that the accurate and efficient system which can segment objects in video can be developed. The proposed system aims to resolve the issue of uncovered background, Temporary poses and Global motion of background.

General Terms
Pattern Recognition, Video Segmentation.

Keywords
Global Motion of Background (GMOB), Semiautomatic segmentation, affine model.

1. INTRODUCTION
Digital technologies and development of computer processing results in to the easier production of digital images and video and so there is a large increase in the quantity of visual information in the last decades. This was accompanied by newly emerging multimedia applications. The few examples of these applications include object-based coding, video repositories, and two way flow of information in video, video conferencing, video surveillance and home video editing. These newly developed applications and the large increase in amount of visual material raised the need for new methods of processing the visual information that will fulfil the requirements of these applications. Most of the requirements of these applications are pliable manipulation, usage of already present visual contents, efficient storage and fast transmission. To fulfil these requirements, video processing has moved from the conventional frame-based processing to content-based processing. The conventionalprocessing of frame based methods, which supports frameby frame coding, gives only restricted power in terms of access, manipulation and interaction with visual content. At the same time, content-based procedure gives morewidespread and stronger way of representing and coding content**.**

Moving object segmentation is a process used to separate object from the background. Video segmentation is the important application domain in the technology of video processing.The method of separating foreground objects from the background has a lot of applications in human-computer communication, compression of video, editing of contents in multimedia files and manipulation. In video compression and cinematographic effects moving foreground is separated from a stationary background given a video sequence. Digital composition is one of the important application of this, in which the required object is extracted from a video clip and pasted to a new background [1,12].

Object can be segmented based on the motion information called as temporal methods .When camera is stationary, the problem of segmentation of moving object becomes identifying the set of pixels which represent objects from a video in static background. In camera in motion sequences, the step of identifying video object is difficult, since object motions are disturbed by camera motion. This undesired motion should be first removed before separation of moving object is done.

Temporary posesor slow movements are also one of the issue in moving object segmentation. When parts of the object stops moving for some time or it is moving very slowly then it becomes very difficult to detect such type of movements. This paper focuses on solving this problem by adding region based segmentation method with the system.

2. REVIEW OF LITERATURE
A lot of work is already done on the video segmentation algorithms. In the review of literature various approaches available for video segmentation are disscussed.

Dong Zhang proposed a method in which primary object regions are extracted using spatially accurate method and temporally dense extraction method for video object segmentation.In this method the author has extracted the object proposal and used DAG approach which gives good segmentation performance. In order to find which image regions are objects and which is background, it makes these methods very slow [2]. Camille Courier proposed a method for Causal Graph based video segmentation. This method uses the graph based matching method, It is more robust to large camera displacements but spanning trees method takes more computation time[3] .MacFralane N.J.B. proposed a method in which piglets present in images aresegmented and tracked. This paper uses approximate median method which employs frame difference with constantly updated background model. Storage requirements of median filtering are reduced by this technique but it requires continuous updating of background model. [4].Ricardo proposed a Mixture of Gaussian model in this Background model is parametric instead of being a frame of values. Performance of this method depends on the modelling of background [5]. Using background registration method segmentation of moving object is proposed by Shao-Yi Chien.This method uses Frame difference, Registration of Background, Detection of object and Post Processing. Computational complexity of this method is low but slow

movements/temporary movements are not identified and works only for fixed camera [6]. In subtraction of background scheme, the important step is the selection of background motion which occurs globally or subtraction of background. In some approaches, assembled frame difference pictures are examined to again find stationary scene elements to compare with frames to detect change. In these approaches, there is a strong supposition of stationary background. Other methods associate consecutive frames to construct a reference background image before applying change detection. When there is a big deviation of the background, it is updated as required. Differences between two frames can also be identified by using two successive frames instead of using a reference background frame.

An algorithm which uses change detection method is proposed by Neri in this potential foreground regions are separated by applying a higher order statistics significance test to interframe differences. The previous methods compares successive frames by relying pixels. On a global level comparison could be done, so methods based on histograms were also proposed [7,17].

Change detection based methods proposed so far have applied frame difference information of two consecutive frames that is the current and the previous frame only. When object stop moving for some time or when move very slowly, it creates a problem for conventional change detector. If we consider only the frame difference the in this case, the motion information disappears. However, if we have information about the difference in background, we can see very clearly that these pixels represent the object region and should be included in the object mask.

In pixel based method which are not like other changedetectionbased approaches,decision criterion for motion does not come directly from the difference in frames whichare two successive frames. Instead an updated background information is maintained from the sequences of frames and analyzed each frame with the background. Any pixel is considered to be present in object region if it is significantly different from the background. The uncovered background is another region where the Pixel-based algorithm performs better than the conventional change detection algorithms. Because there is significant luminance change in uncovered background region and the moving object region, distinguishing the uncovered background from the object is not very easy if only the frame difference is available.

Most algorithms fail in segmenting the foreground with slow movements and temporary poses. Optical flow method of gradient-based methods have shown high performance but generally come with increased calculation burden. Block-based algorithms seem to be unbiased to slow and small object motion from frame to frame. There are lot of methodsforsegmentation of objects present in video, but the rapid video object segmentation techniques are based on change detection approach.

Automatic segmentation is simplified by methods based on detection of changes in frames as compared to finding motion of background globally approaches; they also have their own problems. Specifically, the entanglement of background generation and updating requires more scrutiny in these methods[13]. If there is a stationary referral background frame, the problem will be easier and accurate results may be obtained. However, when the background is in motion, and when initial background reference frame not present, the problem will be more complex, and the results of segmentation may not be accurate. This shows that there is a huge scope to obtain better segmentation system.

3. PROPOSED WORK

On the backdrop of the afore-mentioned review of literature and subsequent gaps identified from the summary of the literature review, the proposed work focuses on thedevelopment of a system to segment video objects automatically from the background given a series of video frames. The proposed work focuses on the issues of "moving camera/Global motion of background", "uncovered background" and "Temporary poses" and there solutions.

3.1 Global motion of background

Motion vectors describes the motion from one image plane to another. This motion vectors are distributed over the image plane which yields in to motion field.When we use temporal changes to analyze the image the motion is called as an apparent motion. Motion estimation aims at characterizing this apparent motion by motion field. There are two approaches by which we can calculate the motion field one is parametric and another is nonparametric. In this work focus is on the parametric method. Motion of the background can be estimated using different parametric methods such as translation,transformation,affine and projective models. Since affine model gives the optimum result with less complexity this model is used forbackground motion estimation.

3.1.1 Global Motion Estimation

Object motions are disturbed by camera motion. This undesired motion should be removed first before actually segmenting the moving object. This is done in three steps as motion vector estimation, Removal of motion vectors of background and finally frames warping. To find the motion vectors frame is divided in to blocks of n*n (8*8/16*16).Then motion vectors are found by searching for the best match in the reference or previous frame. To find best match criteria is used to minimize a measure of matching error between current block and blocks in previous frame.

$$MAD(m, n) = 1/n*m\ (\sum gl(I)-gl(I-1)) \quad (1)$$

$$(u,v)=min\ (MAD\ (m,\ n)) \quad (2)$$

Where MAD is mean absolute difference, gl is grey level and (u, v) is motion vector.

After finding motion vectors, motion vectors that greatly differ from their neighbors are rejected. The mean of 3*3 group of motion vector is calculated and compared with motion vector under test.

The affine model is able to model the zooming,rotation and shear along with simple translation. The affine model describes the displacement of a pixel of a region from frame n to frame n+1, bytranslation, rotation and linear scaling. In affine model there are 6 parameters a1 to a6 which are called as global motion parameters. These parameters are calculated from the motion vectors.

3.1.2. Global Motion Compensation

From the above step affine parameters are estimated then the motion of the camera is compensated according to the motion model.Frame warping is used to align the previous or next frame to current frame. New frame is calculated from previous frame by transforming the co-ordinates of previous frame into new position defined as

X'=a1*X + a2* Y+ a3 and

Y'= a4*X + a5* Y+ a6 (3)

Where a1 to a6 represents to transformations, 2 scaling and 2 rotation camera parameters.

The calculated pixel positions using frame warping method may not be integer values. These non-integer values are converted in to pixel intensities using different methods of interpolation such as linear, bilinear, spline, nearest neighbor etc.

The proposed work focused on the simple form of interpolation that is linear interpolation where the mean of two adjoining points is taken to find the middle point.

3.2 Uncovered Background

Uncovered background is the region which will be visible only when the object in the scene moves. In video which are captured by using the moving camerabackground motion is compensated first, before change detection is applied. To solve theproblem of uncovered background (Fig.1) for video sequences whose background frame isdifficult to obtain, two change detection results are combined.

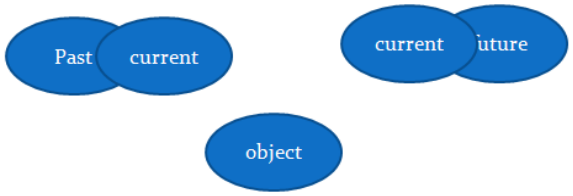

Fig. 1: Removal of Uncovered Background

In this system uses three successive frames as past, current and future. The past frame is normalized with respect to current frame and future frame is normalized with respect to current frame .This two are combined by a logical AND operator. The mask eliminates all areas except the foreground object detected which is the region that overlaps in two masks[18].

3.3 Temporary Poses/Slow Movements

Block Matching algorithm is more tolerant to slow movements so this method is usedto calculate the motion vectors.To resolve the issue of temporary poses we have integrated region based segmentation with our system. Region based segmentation partitions the frame in to regions which are uniform with respect to some characteristics such as colour and intensity. Result of region based segmentation is "OR" with the object detection result to give final output.

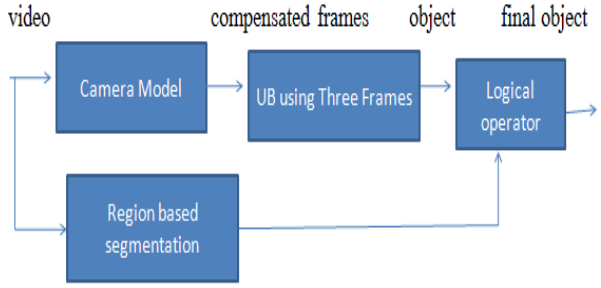

Fig.2: Block Diagram of the Proposed System

4. EXPERIMENTAL RESULTS

The system is tested on a Setback standard dataset. It consists of 14 videos some of which are having interacting objects, slow movements, deformation, motion blur, and occlusion. Algorithm is applied on a humming bird sequence. First video is given as an input to Global Motion and estimation, here motion vectors and camera parameters are calculated. Then Frame warping is done to compensate the motion of background. In the next step compensated frames are given to uncovered background step. In this object is detected and uncovered background is removed. The result of this step is 'OR' with the result of region based segmentation to give final output. The experimental results includes detected object and performance analysis.

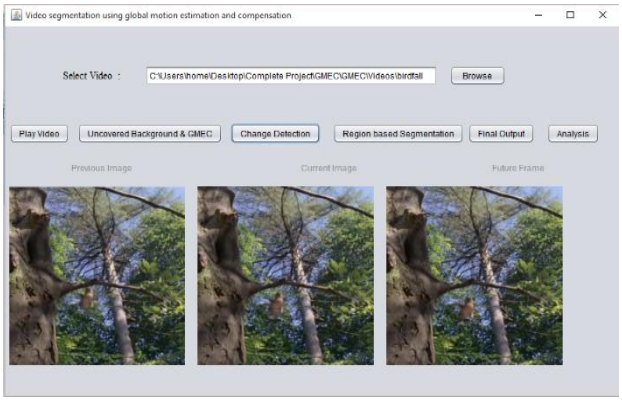

Fig.3: Input Video

(a) (b) (c)

Fig. 4:Result of (a) Detected object (b) Region based segmentation output (c) Final output

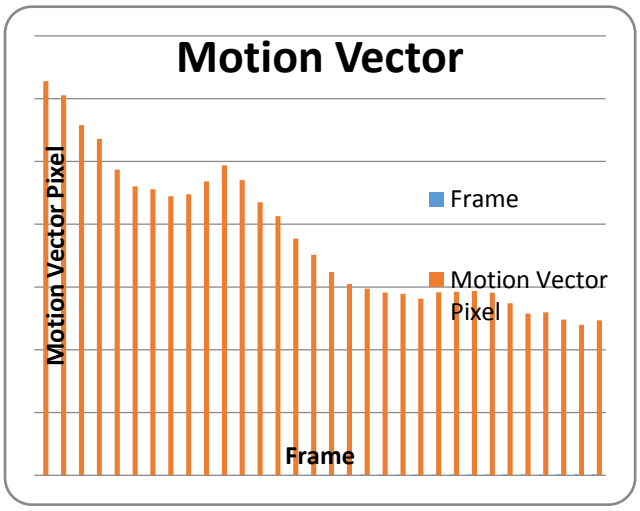

Fig.5: Graphical Representation of Motion Vector

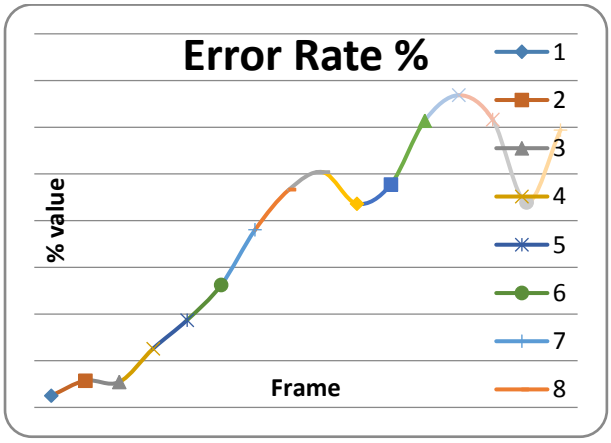

Fig. 6 Graphical Representation of Error Rate

For each video average per frame pixel error is calculated by dividing the XOR result of ground truth and detected object by number of frames of that video .The result of the proposed method is compared with the other methods and found that pixel error is reduced by proposed method.

Table 1. Average per frame pixel error

Video	Proposed	[8]	[9]	[10]	[11]
Birdfall	42	155	189	288	252
Cheetah	216	633	806	905	1142
Girl	611	1488	1698	1785	1304
Monkeydog	256	365	472	521	563

Table 2.Analysis of videos

Compared with Ground Truth			
Image Sequence	**Number Of Frames**	**Number Of Object Detected**	**Missed Object**
Girl	21	1	0
Birdfall	30	1	0
Cheetah	29	2	0
Drifting Car	74	2	0

Table 3 Performance Analysis of Proposed method for Video Segmentation

Methods	Processing Time(Sec)	Mean	Variance	Frame Rate(Frame per Second)	Accuracy
Frame Difference (FD)	24.43	4.2	3.5	4.16	H
Approximate Median (AM)	15.56	11.84	10.2	4.06	M
Mixture of Gaussians (MGM)	267.63	11.8	2.0	0.30	L
Proposed method	23.67	0.03	0.04	3.62	H

The system is analyzedby calculating Precision, Recall and F-Measure. Fraction of retrieved instances which are relevant represents the precision; it is calculated by dividing the true positive instances by total positive instances. Fraction of relevant instances which are retrieved represents the recall; it is calculated as dividing true positive instances by false negative and true positive instances.

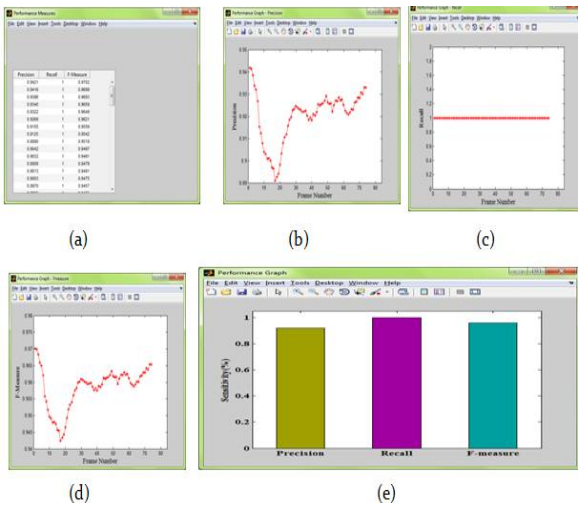

Fig 7.Performance Analysis (a) performance measure (b) Precision (c) Recall (d) F-Measure (e) Sensitivity

5. CONCLUSION

The proposed work solved the issue of Moving camera which adds the unwanted disturbance in the video and also solved the problem of temporary poses or slow movements. The performance analysis shows that the system gives good precision and recall. The work also improved the accuracy of Video segmentation.The proposed method is compared with the existing methods with respect to average per frame pixel error and it is observed that the proposed method resultsin comparatively less error. The processing time of the proposed method is also improved as compared to the existing methods. The proposed method has been tested on a standard dataset of segtrack and it has been observed that the method gives the good results.

6. REFERENCES

[1] Shao-Yi Chien, Yu-Wen Huang, Bing-Yu Hsieh, Shyh-YihMa, and Liang-Gee Chen,"Fast Video Segmentation algorithm with Shadow Cancellation, Global Motion compensation, and Adaptive Threshold Techniques,"IEEE Trans. on Circuits and System for VideoTechnology., vol. 6, pp. 732- 748, no. 5, Oct. 2004.

[2] Dong Zhang1, Omar Javed2, Mubarak Shah1," VideoObject Segmentation through Spatially Accurate andTemporally Dense Extraction of Primary Object Regions," 2013 IEEE Conference on Computer Vision and Pattern Recognition

[3] Camille Couprie,"Causal Graph based videosegmentation"(2012)

[4] McFralane, N. J. B. and Schofield, C.P. "Segmentation and tracking of piglets in images". Machine Vision and Applications, Vol. 8, No. 3, 187-193. 2005.

[5] Ricardo Augusto Castellanos Jimenez "Event Detection In Surveillance Video" FloridaAtlantic UniversityBoca Raton, Florida May 2010

[6] Shao-Yi Chien, Shyh-Yih Ma, and Liang-Gee Chen, "Efficient Moving Object Segmentation Algorithm Using Background Registration Technique," IEEE Trans. on Circuits Syst. Video Technol., vol. 12, no. 7, pp. 577-586, 2002.

[7] Tung-Chien Chen "Video Segmentation Based on Image Change Detection for Surveillance Systems".

[8] Dong Zhang, Omar Javed, Mubarak Shah," Video Object Segmentation through Spatially Accurate and Temporally Dense Extraction of Primary Object Regions," 2013 IEEE Conference on Computer Vision and Pattern Recognition

[9] T. Ma and L. Latecki. Maximum weight cliques with mutex constraints for video object segmentation. In CVPR, pages 670–677, 2012.

[10] Y. Lee, J. Kim, and K. Grauman. Key-segments for video object segmentation. In ICCV, pages 1995–2002, 2011.

[11] D. Tsai, M. Flagg, and J. Rehg. Motion coherent tracking with multi-label mrf optimization. In BMVC, page 1, 2010.

[12] H. Jiang, A.S. Helal, A.K. Elmagarmid, and A. Joshi. "Scene change detection techniques for video database systems". multimedia Systems, 6(3):186–195.

[13] A.Dailianas, R.B. Allen, and P. England. Comparison of automatic video segmentation algorithms". In SPIE Conference on Integration Issues in Large Commercial Media Delivery Systems, volume 2615, pages 2–16, Philadelphia, PA

[14] M.K. Mandal, F. Idris, and S. Panchanathan" A critical evaluation of image and video indexing techniques in the compressed domain". Image and Vision Computing, 17(7):513–529.

[15] S. Y. Chien, Y. W. Huang, and L. G. Chen, "Predictive watershed: a fast watershed algorithm for video segmentation," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, May 2003, Page(s):453-461

[16] R. Zabih, J. Miler, K. Mai, "A feature-based algorithm for detecting and classifying production ejects", Multimedia Systems 7 (1999) 119}128.

[17] K. Zhang and J. Kittler, "Using background memory for efficient video coding," in Proc. IEEE Int. Conf. Image Processing, 1998, pp. 944–947.

[18] R. M. Haralick and L. G. Shapiro, Computer and Robot Vision. Reading, MA: Addison-Wesley, 1992, pp. 28–48.

[19] Chen T-H, Liau H-S and Chiou Y-C. (2005) "An Efficient Video Object Segmentation Algorithm Based on Change Detection and Background Updating." Kun Shan University,National Computer Symposium, MIA1-2 (MI14).