IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 4, Ver. VI (Jul.-Aug. 2016), PP 26-30
www.iosrjournals.org

Video Segmentation Using Global Motion Estimation and Compensation

Pallavi R. Wagh[1], Shubhangi Vaikole[2], Dr.Sudhir Sawarkar [3]
[1]*(ME Student, Department of C.E., Datta Meghe College of Engineering, Airoli, Navi Mumbai, India)*
[2]*(Research Scholar, Datta Meghe College of Engineering, Airoli, Navi Mumbai, India)*
[3]*(Principal, Datta Meghe College of Engineering, Airoli, Navi Mumbai, India)*

Abstract : *Video has to be segmented into objects for content-based processing. A number of video object segmentation algorithms have been proposed such as semiautomatic and automatic. Semiautomatic methods adds burden to users and also not suitable for some applications. Automatic segmentation systems are still a challenge, although they are required by many applications. The proposed work aims at contributing to identify the gaps that are present in the current segmentation system and also to give the possible solutions to overcome those gaps so that the accurate and efficient video segmentation system can be developed. The proposed system aims to resolve the issue of uncovered background, Temporary poses and Global motion of background*
Keywords - *Moving Object, Region based Segmentation, Block Matching, Global Motion Estimation and Compensation*

I. Introduction

Moving object segmentation is a process used to separate object from the background. In digital video processing technology video segmentation generated by objects is an important application domain. Segmentation of foreground objects from background has a lot of applications in human-computer interaction, video compression, multimedia content editing and manipulation. The extraction of the moving foreground from a stationary background from a general video sequence has various applications such as compression of videos and also in the cinematographic effects. One of its important applications is digital composition, in which the object of interest is extracted from a video clip and pasted to a new background [1].

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

A large number of video object Segmentation methods have been proposed, most aiming to specific applications, and trying to fulfill specific constraints. Good results have been obtained so far in semiautomatic methods, since there is also human interference in the segmentation process. However, the human assistance involved in these methods is not required because it unnecessary adds work of users and also it is not suitable for some applications. On the other hand, fully automatic video segmentation systems are still a summons, although they are required by many applications.

Many automatic segmentation systems are developed for particular problems and with simplified assumptions like videos with fixed background. So it is very important to have flexible automatic segmentation system for different types of videos. Most of the existing automatic segmentation systems involve complicated techniques. Also each stage of the segmentation process involves computationally acute operations to obtain good results. Thus, reducing the complications of the techniques involved is required while keeping good performance of segmentation results. This can be done by selecting efficient algorithms with reduced complicated methods in each step of the segmentation process. Accuracy of segmentation can be improved by applying post-processing.

II. REVIEW OF LITERATURE

A number of video segmentation algorithms have been proposed. This section provides a critical review of the various approaches available for video segmentation.

Dong Zhang proposed a method for video object segmentation through spatially accurate and temporally dense extraction of primary object regions. In this method the author has extracted the object proposal and used DAG approach which gives good segmentation performance. To find out which image regions are objects vs. 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 for segmentation and tracking of piglets in images. This paper uses approximate median method which employs frame difference with constantly updated background model. Storage requirements of median filtering are alleviated 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].

Efficient moving object segmentation algorithm using background registration technique is proposed by Shao-Yi Chien. This method uses Frame difference, Background Registration, Object detection 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].

An algorithm based on change detection is proposed by Neri which separates potential foreground regions employing a higher order statistics (HOS) significance test to inter-frame differences. The earliest methods were comparing successive frames by relying pixels. Comparison could be performed on a global level, so methods based on histograms were also proposed [7]

.A. Summarized findings of literature review

Change detection based methods proposed till date has applied frame difference information of two successive frames (the current and the previous frame) only. One of the problems that confuse the conventional change detector is that of the temporary poses or slow movements. In these cases, the motion information disappears if we use the difference in the frames only. But, if we use background difference information, we can see very clearly that these pixels belong to the object region and should be included in the object mask.

Most algorithms fail in segmenting the foreground with slow movements and temporary poses. Optical flow methods based on gradient have shown good results but generally come with increased computational overhead. Block-based algorithm gives satisfactory results for slow movements and small object motion from frame to frame. There are various methods of segmentation of video object, but the faster video object segmentation techniques are based on change detection approach.

If videos are captured using a fixed camera the segmentation will be easier and accurate results may be obtained. However, when the videos are captured through moving camera, and when no initial background reference frame is present, the segmentation will be difficult, and the results of segmentation may not be that good. This shows that there is still a lot to be done to obtain better segmentation system.

III. Proposed Work

On the backdrop of the afore-mentioned review of literature and subsequent gaps identified from the findings of the literature review, the proposed work aims at contributing to develop a system to segment video objects automatically from the background given a sequence of video frames. The proposed work aims to resolve the issue of “moving camera/Global motion of background”, “uncovered background” and “Temporary poses”.

Proposed Method

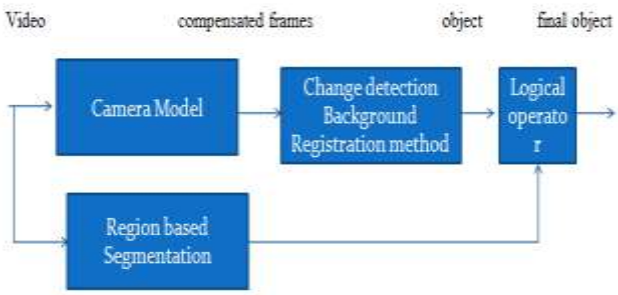

Fig 1: Block diagram of proposed method

Global motion of background

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)) (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 neighbours are rejected. The mean of 3*3 group of motion vector is calculated and compared with motion vector under test. Then 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.

Removal of Uncovered Background

In this system uses three consecutive 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 operator removes all areas except the foreground object detected which is the region that overlaps in two masks

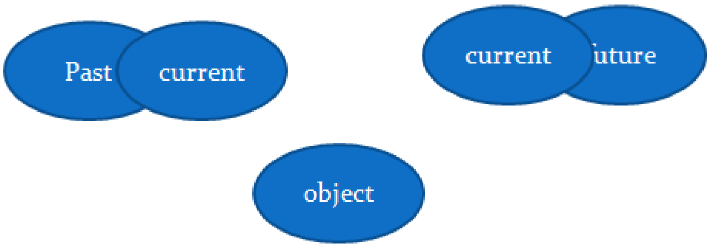

Fig2: Removal of Uncovered Background

Temporary Poses/Slow Movements

Block Matching algorithm is more tolerant to slow movements so we have used this method to 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.

IV. Result And Discussion

Result of Proposed Method

The system is tested on a Segtrack 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

Proposed method is also evaluated on Segtrack v2 dataset [20]. There are 14 videos in this dataset, and also a pixel-level segmentation ground-truth for each video is available. It consists of 14 videos with camera in motion some of which are having interacting objects, slow movements, deformation, motion blur, and occlusion. In SegTrack dataset there are 6

Videos in addition to those 8 more videos are added in SegTrack v2 dataset. Also there are changes in the ground truth for videos which contains multiple objects. In SegTrack dataset only one object is detected whereas in SegTrack v2 dataset multiple objects are detected.

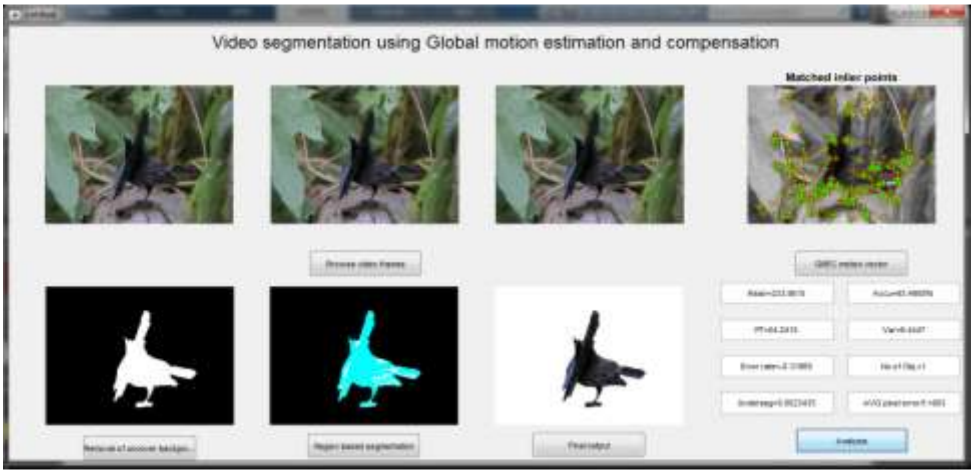

Fig 3: Snapshots of proposed method 1

TABLE I Evaluation Parameters

Mean	233.8615
Variance	6.4447
Processing time	54.2415
Under/over segmentation	0.0023435
Error rate	0.31965
Accuracy	93.4906
Avg.per frame pixel error rate	903

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 .we have compared our result with the other methods and found that pixel error is reduced by our proposed method.

TABLE II Average per frame pixel error

Video	Ours	[2]	[4]	[5]	[8]
Birdfall	25	155	189	288	252
Chetaah	401	633	806	905	1142
Girl	251	1488	1698	1785	1304
Monkeydog	21	365	472	521	563

From Table II we can say that average per frame pixel error is improved as per other methods
Table III shows segmentation accuracy of the proposed methods and the state of the art methods including tracking and graph based approaches [22, 23, 24, 25, 26, 27and 28]. The proposed method's segmentation accuracy is improved as compared to other methods.

TABLE III segmentation Accuracy

Video	VS-3FA	[22]	[23]	[24]	[25]	[26]	[27]	[28]
Girl	89.139	83.7	**89.2**	87.7	31.9	53.6	52.4	87.9
Birdfall	**90.652**	77.5	62.5	49	57.4	56	32.5	57.4
Parachute	**94.753**	94.9	93.4	96.3	69.1	85.6	69.9	94.5
Cheetah-Deer	**67.920**	63.1	37.3	44.5	18.8	46.1	33.1	33.8
Cheetah-Cheetah	66.9382	35.3	40.9	11.7	24.4	47.4	14	**70.4**
Monkeydog-Monkey	**98.515**	82.2	71.3	74.3	68.3	61	22.1	54.4
Monkeydog-Dog	**55.845**	21.1	18.9	4.9	18.8	18.9	10.2	53.3
Penguin-#1	71.7662	92.7	51.5	12.6	72	54.5	20.8	**93.9**
Penguin-#2	43.4248	**91.8**	76.5	11.3	80.7	67	20.8	87.1
Penguin-#3	69.7539	**91.9**	75.2	11.3	75.2	7.6	10.3	89.3
Penguin-#4	44.2337	**90.3**	57.8	7.7	80.6	54.3	13	88.6
Penguin-#5	50.4218	76.3	66.7	4.2	62.7	29.6	18.9	**80.9**
Penguin-#6	73.0292	**88.7**	50.2	8.5	75.5	2.1	32.3	85.6
Drifting-#1	44.2213	67.3	74.8	63.7	55.2	62.6	43.5	**84.3**
Drifting-#2	54.4297	**63.7**	60.6	30.1	27.2	21.8	11.6	39
Hummingbird-#1	**87.962**	58.3	54.4	46.3	13.7	11.8	28.8	69
Hummingbird-#2	**73.405**	50.7	72.3	74	25.2	-	45.9	72.9
BMX-Person	**96.957**	88.9	85.4	87.4	39.2	2	27.9	88
BMX-Bike	**79.43**	5.7	24.9	38.6	32.5	-	6	7
Frog	**91.33**	61.9	72.3	0	67.1	14.5	45.2	81.4
Worm	75.7265	76.5	82.8	84.4	34.7	36.8	27.4	**89.6**
Soldier	**91.116**	81.1	83.6	66.6	66.5	70.7	43	86.4
Monkey	80.4877	86	84.8	79	61.9	73.1	61.7	**88.6**
Bird of Paradise	93.4908	93	94	92.2	86.8	5.1	44.3	**95.2**

V. Future Scope

The use of two change detection mask has introduced a delay ,this can be overcome by using a single change detection mask. More efficient data structure and software can be used for implementation to improve the performance

VI. 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.

References

[1]. Shao-Yi Chien, Yu-Wen Huang, Bing-Yu Hsieh, Shyh-Yih Ma, 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 Video Technology., vol. 6, pp. 732- 748, no. 5, Oct. 2004.

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

[3]. Camille Couprie,"Causal Graph based video segmentation"(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" Florida Atlantic 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).

[20]. Hanqing Jiang, Guofeng Zhang, Huiyan Wang, Hujan Bao," Spatio-Temporal Video Segmentation of Static Scenes and its Applications." IEEE Transaction on Multimedia ,VOL.17,2015

[21]. Stephen Liwicki,Stefanos P. Zafeiriou,Maja Pantic, "Online Kernel Slow Feature Analysis for Temporal Video Segmentation and Tracking", IEEE Transaction on Image Processing ,VOL.24,2015

[22]. L. Wen, D. Du, Z. Lei, S. Z. Li, and M.-H. Yang. Jots: Joint online tracking and segmentation. In CVPR, 2015. 1, 2, 6, 7,

[23]. F. Li, T. Kim, A. Humayun, D. Tsai, and J. M. Rehg. Video segmentation by tracking many figure-ground segments. In ICCV, 2013. 1, 2, 5, 6

[24]. Y. J. Lee, J. Kim, and K. Grauman. Key-segments for video object segmentation. In ICCV, 2011.

[25]. M. Grundmann, V. Kwatra, M. Han, and I. Essa. Effi- cient hierarchical graph-based video segmentation. In CVPR, 2010.

[26]. M. Godec, P. M. Roth, and H. Bischof. Hough-based tracking of non-rigid objects. In ICCV, 2011.

[27]. S. Wang, H. Lu, F. Yang, and M.-H. Yang. Superpixel tracking. In ICCV, 2011.

[28]. Yi-Hsuan Tsai,"Video segmentation via Object Flow",CVPR201**6.**