IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-4, Special Issue-3, May 2016 E-ISSN: 2347-2693

Saliency Aware Video Object Detection and Tracking

Rakshitha N[1], Mangala C N[2]

[1,2]*Dept. of Computer Science and Engineering,*
East West Institute of Technology, Bengaluru, India

Available online at: www.ijcseonline.org

Abstract—Detection and tracking of moving objects in a video has been emerging as a demanding research in the domain of computer vision and image processing in the resent years. It has been used in various applications like visual surveillance, traffic monitoring etc for tracking interested objects. An efficient method for object detection and tracking is proposed in this work. Two discriminative visual features like spatial edges and temporal motion boundaries as indicators for foreground object locations are considered. Initially frame wise spatiotemporal saliency maps by making use of geodesic distance indicators are created. Geodesic distance also provides an initial estimation for background and foreground by building on the observation that foreground areas are surrounded by the regions with high patio temporal edge values. Coherent object segmentation is done by combining all this spatio temporal maps. Finally the segmented object is tracked using Kalman filter get efficient result

Keywords—Spatial edges, Temporal motion boundaries, Spatiotemporal saliency maps, Geodesic distance, Kalman filter, Visual surveillance, Pixel segmentation, super pixels.

I. INTRODUCTION

In the field of security automated surveillance system plays a prominent role due to its immense interest in this field. Due to which the traffic statistics, scene motion, human identification, object detection and tracking offers a context for the extraction of significant information. Especially Video surveillance has generated great deal of interest in the field of surveillance. Detecting moving objects and tracking in a video is a challenging task as differentiating the interested object from the background becomes a typical problem.

Videos are nothing but the sequences of images. Each of these images is called as frames. Usually these frames are displayed in fast enough frequency so that human eyes can predict the changes in the continuity of the contents present in the frame. The contents present between the consecutive frames are however closely related. Hence statuses of the objects are tracked making use of the two adjacent frames as still or moving. Video analysis can be done by applying all pre-processing steps to the generated frames. Tracking of these objects from frame to frame to recognize the object behavior is done after object detection as two are closely related process.

Divyani Prajapati et.al [12] proposed an efficient method for object detection and tracking. This involved steps like detection of moving objects and tracking frame to frame to obtain temporal information. Comparative and fundamental analyses for different techniques are also provided in the literature. Houari Sabirin et.al [03] presented an efficient method for moving object detection and tracking based on spatio temporal graph based approach in H.264/AVC bit streams. By first clustering the encoded blocks of the potential object block the spatio temporal graph is constructed. Using which object blocks are enabled and effective detection for the small sized objects and the objects with small motion vectors and residues are done even under occlusion.

C. Beyan et.al [02] presented a fully automated multi object tracking system based on mean shift algorithm. With combination of Gaussian followed by shadow and noise removal foreground is extracted to initialize object tracking. Hitesh A Patel et.al proposed an efficient method for moving object detection and tracking. Here detection of moving object has been done using simple background subtraction method and object tracking is achieved using kalman filter. Researchers in [04], [05], [06], [07], [08], [10] also proposed an efficient method for object detection and tracking. In the proposed work we made use of spatial edge probability Map and super pixel segmentation approach along with super pixel segmentation method for object detection along with spatiotemporal edge probability map and spatiotemporal saliency Map for proper segmentation of the interested object. Tracking of this object is done by using Kalman filter.

II. METHODOLOGY

The proposed methodology is as in the Figure 1. Initially Video is taken and passed to frame generation block. Here the video is divided into number of frames. Every time the generated frame is then passed to pre-processing block for applying few pre-processing steps like image resizing, colour conversion. Processed image is then passed to

Spatial edge probability Map and Super pixel segmentation block to get the spatial edge image and super pixel image. These two images are then combined to get the Spatio temporal edge probability map after averaging. The obtained image is then passed to Spatio temporal saliency mapping for object segmentation. The segmented object is then passed to Kalman filter for object tracking.

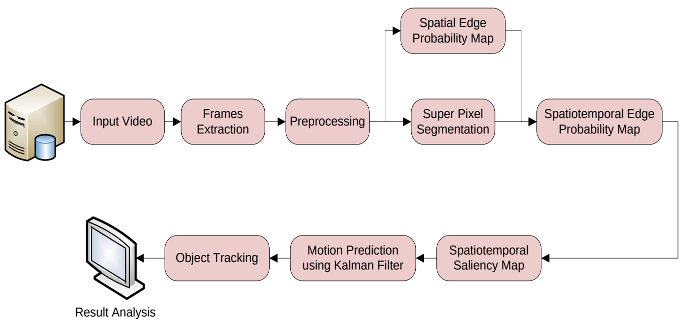

Figure 1: Architecture of Proposed System

A. Spatial Edge Probability Mapping and Super Pixel Segmentation

Edges provide good idea in finding object boundaries in case of video frames. Calculation motion would also acts as a good indicator of object as the pixels in it could change abruptly when compared to neighbours. To the obtained video frame the edge probability map is calculated. The video sequence is given by the equation,

$$F = \{F^1, F^2, \ldots.\} \quad (2)$$

Edge probaility map is given by, $E_c^k(x_i^k)$ corresponding to kth frame F^k at pixel x_i^k. Each frame is segmented into super pixel using SLIC [01]. Let $Y^k = \{ Y_1^k, Y_1^k, \ldots \}$ be the super pixel set of the generated frame F^k. Once the pixel edge map E_c^k is calculated, the edge probability of each of the superpixel Y_n^k is computed as the average value of the pixels with the ten large valued edge probabilities within Y_n^k. Using this super edge map $\hat{E}_c^k$ is generated. When these spatial edge and super pixels are fused together the output spatiotemporal edges are able to determine the location of the foreground object. Results show that the regions either have high values of spatiotemporal edge or these areas are surrounded by high edge probability regions. Based on this we make use of geodestic distance to differentiate the visually salient areas from background and estimate their likelihoods for foreground.

B. SpatioTemporal Saliency Map

The obtained probability map is not able to locate the object in the foreground more precisely. Due to the over segmentation object probabilities in the background regions near the boundaries of the object gets increased. As background and foreground becomes visually different especially between adjacent frames we proposed an efficient method which leverages this information to obtain spatio temporal saliency results between the adjacent pairs of frames.

For each set of sequence F an weighted graph is constructed which is given by,

$$G'^k = \{V'^k, \varepsilon'^k\} \quad (1)$$

The obtained nodes V'^k consists all of the superpixels Y^k of all the frames F^k. For each of this frame a self adopted threshold is used to decompose these frames into compose frame into the background region B^k and object regions U^k using the probability map. Further the spatiotemporal object segmentation as in [] is done to get the segmented object [11]. The segmented object is then passed to kalman filter for object tracking. Detail Flow for this is given as in the Figure2.

C. Tracking Using Kalman Filter

Kalman filter is an efficient approach for object tracking. It is usually used to estimate the state of a linear system where it is assumed that state is distributed by a Gaussian. Object tracking usually involves predicting the objects position from the previous frame and verifying it's existing. Second step is to observe the likelihood function and learning of motion model is also done using some sample of frame sequences. It provides estimation of present, past and also future states. The kalman filter [13] estimates the feedback control system. The equations of kalman filters fall in two groups for optaining time update equations and measurement update equations. The time update equation is responsible for projecting the in time (forward) the current state and estimation using error covariance is done to obtain the prior estimation for the next time step. The equation sof measurement update are responsible for the feedback operation. Figure 3 depicts the final estimation process giving the predictor corrector algorithm for solving numerical problems.

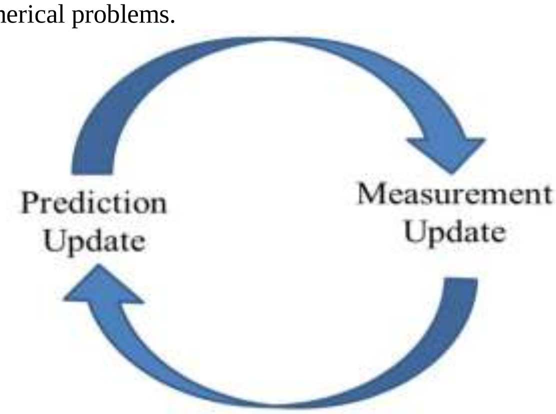

Figure 2: Kalman Filter Cycle

The time update always projects the current state estimate forward in time. The measurement update always adjusts the projected estimate by performing actual measurement [09].

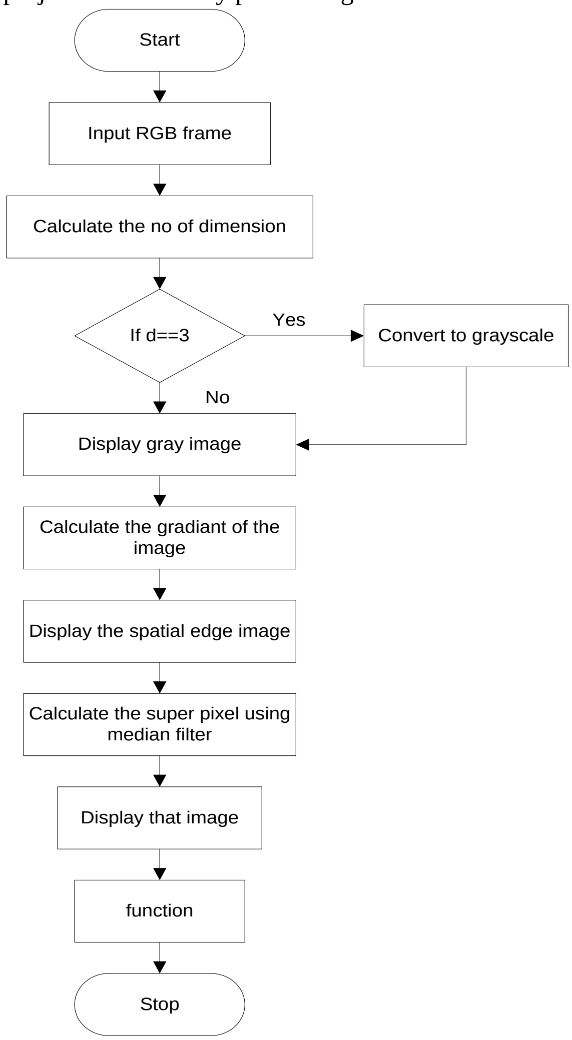

Figure 3: Flow chart for Spatio Temporal Edge Map

III. Experimental Result

This section gives the results obtained at each stage of the proposed system. For results demonstration only few frames outputs are shown. The generated frames as in Figure 4 (a) undergo few pre-processing likes resizing and gray conversion as in (b). Furthermore the Spatial edge image is obtained as in (c).

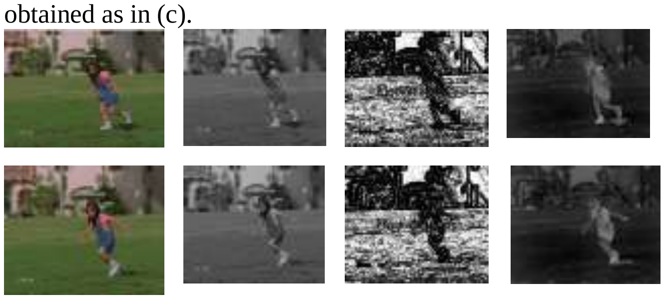

Figure 4: (a) Input Image; (b) Grey Image; (c) Spatial Edge Image; (d) Spatiotemporal Edge Mapped Image.

The Final Spatiotemporal edge mapped image is obtained as in (d) after combing the super pixel image and spatial edge image. The object detection as in Figure 5 (a) is done before object segmentation. Once the object segmentation is done it is passed to Kalman filter for object tracking as in (b).

Figure 5: (a) Object Detection; (b) Object Tracking

IV. Conclusion

Moving object detection and tracking becomes an attractive and challenging research topic now days. Hence designing an efficient tracking system is a difficult task. An efficient method for video surveillance is proposed in this paper to get good recognition and tracking. The system efficiently detects and tracks object using spatio temporal edge map for indicating the location of background and foreground. The approach combined spatio temporal edge map and geodesic distance to get accurate spatiotemporal saliency results for accurate object segmentation. The obtained segmented image is passed to kalman filter to track the interested object in the every sequence. The proposed system could give good result compared to the existing system.

References

[1] Radhakrishna Achanta, Appu Shaji, Kevin Smith,Aurelien Lucchi, Pascal Fua, and Sabine Susstrunk, "SLIC Superpixels Compared to State-of-the-art Super pixel

Methods", Journal of latex class files, Vol. 6, Issue 1, **2011**.

[2] C. Beyan A and Temizel, "Adaptive Mean-Shift for Automated Multi Object Tracking", The Institution of Engineering and Technology, **2011**.

[3] Houari Sabirin and Munchurl Kim, "Moving Object Detection and Tracking Using a Spatio-Temporal Graph in H.264/AVC Bit streams for Video Surveillance", IEEE transactions on multimedia, Vol. 14, Issue 3, **2012**.

[4] Kinjal A Joshi and Darshak G. Thakore, "A Survey on Moving Object Detection and Tracking in Video Surveillance System", International Journal of Soft Computing and Engineering, Vol. 2, Issue 3, **2012**.

[5] Bahadir Karasulum and Serdar Korukoglu, "Moving object detection and tracking by using Annealed background subtraction method in Videos: Performance Optimization", Elsevier, pp**. 33–43, 2012**.

[6] Alexander Schick, Martin Baum and Rainer Stiefelhagen, "Improving Foreground Segmentations with Probabilistic Superpixel Markov Random Fields", IEEE, pp**. 27 – 31, 2012**.

[7] Xiaofang Wang, Huibin Li, Simon Masnou, Liming Chen, "Sparse Coding and Mid-Level Superpixel-Feature for l0-Graph Based Unsupervised Image Segmentation", HAL, pp**. 160-168, 2013.**

[8] Sayed Hossein Khatoonabadi and Ivan V. Baji´c, Senior Member, IEEE, "Video Object Tracking in the Compressed Domain Using Spatio-Temporal Markov Random Fields", IEEE, Vol. 22, Issue 1, **2013**.

[9] Hitesh A Patel and Darshak G Thakore, "Moving Object Tracking Using Kalman Filter", IJCSMC, Vol. 2, Issue 4, pp**. 326 – 332, 2013**.

[10] Dan Oneata, Erome Revaud, Jakob Verbeek and Cordelia Schmid T, "Spatio-Temporal Object Detection Proposals", HAL, pp**.737-752, 2014**.

[11] Wenguan Wang, Jianbing Shen and Fatih Porikli, "Saliency-Aware Geodesic Video Object Segmentation", IEEE, 3395 – 3402, **2015**.

[12] Divyani Prajapati and Hiren J Galiyawala, "A Review on Moving Object Detection and Tracking", International Journal of Computer Application, Vol 5, Issue 3, **2015.**

[13] Sanjivani Shantaiya, Kesari Verma and Kamal Mehta, "Multiple Object Tracking using Kalman Filter and Optical Flow", European Journal of Advances in Engineering and Technology, Vol. 2, Issue 2, pp**.34-39, 2015**.

AUTHORS PROFILE

Rakshitha N has received B.E degree in Computer Science and Engineering from VKIT, Bengaluru, Visvesvaraya Technological University in 2014 and is currently pursuing her MTECH degree in Computer Science and Engineering at EWIT, Bengaluru, Visvesvaraya Technological University from 2014-2016. Her area of intrest includes Image Processing, Cloud Computing Datamining and Big data.

Mangala C N received B.E degree in Computer science and engineering from NCET, Bengaluru, Visvesvaraya Technological University in 2006 and got MTECH degree in Computer Science and Engineering from RVCE, Bengaluru, India. She is currently a Asst. Prof in the facult of CSE, EWIT – Bengaluru, India. Her area of intrest includes Image Processing, Datamining and Big data.