Video Salient Object Detection using Spatiotemporal Deep Features

P.Sivaprasad[1], B.Malleswari[2]

[1]*(PG DECS), Department of ECE QIS College of Engineering &Technology (Autonomous), Ongole, A.P.India,*

[2]*AssociateProfessor, DepartmentofECE,QISCollegeofEngineering&Technology (Autonomous), Ongole, A.P.India,*

Abstract- This paper presents a method for detecting salient objects in videos where temporal information in addition to spatial information is fully taken into account. Following recent reports on the advantage of deep features we propose a new set of Spatiotemporal Deep (STD) features that utilize local and global contexts over frames to the accurate detection of salient objects' boundaries and noise reduction during detection. Our proposed method first segments an input video into multiple scales and then computes a saliency map at each scale level using STD features The final saliency map is computed by fusing saliency maps at different scale levels. Our experiments, using publicly available benchmark datasets, confirm that the proposed method significantly outperforms state-of-the-art methods. We also applied our saliency computation to the video object segmentation task, showing that our method outperforms existing video object segmentation

Keywords- videosaliency, salient objects detection, Spatiotemporal Deep Features, video object detection

I. INTRODUCTION

Saliency detection have utilized deep learning to obtain high-level features to detect salient regions. Spatio temporal deep feature using Detection of moving object in an image sequences is crucial issue of moving video. A very common difficulty in the detection of moving object is the presence of ego motion. We solve it by computing the dominant motion. Visual surveillance is a very active research area in computer vision. The use of this concept in surveillance for security, fight against terrorism and crime. The main task includes motion detection, object classification, tracking activity. The detection of moving objects in video streams is the first relevant step of information extraction in many computer vision applications. Many organizations and institutions needs to secure their facilities thus need to use security and surveillance systems that are equipped with latest technology. Intelligent video sensors were developed to support security systems to detect unexpected movement without human intervention. The important information of move,location, speed and any desired information of target from the captured frames can be taken from the camera and can be transferred to the analysis part of the system. Movement detection is one of these intelligent systems to which detect and tracks moving targets. There are different methods to detect moving objects but these methods having some limitations for real time application. Due to this reason, in this paper we use background subtraction method, which is more suitable for real time application which gives accurate result.

II. PROPOSED SYSTEM

Our proposed technique first sections an information video into various scales and after that processes a saliency delineate each scale level utilizing STD highlights with . The last saliency outline figured by intertwining saliency maps at various scale levels. Our investigations, utilizing freely accessible benchmark datasets, affirm that the proposed strategy fundamentally outflanks the best in class strategies. We additionally connected our saliency calculation to the video protest division errand, demonstrating that our technique beats existing video question division strategies.Object detection in video streams is the first step of many computer vision applications. Background modelling and subtraction for moving detection is the most common technique for detecting, while how to detect moving objects correctly is still a challenge. Some methods initialize the background model at each pixel in the first N frames. However, it cannot perform well in dynamic background scenes since the background model only contains temporal features. Herein, a novel pixel wise and nonparametric moving object detection method is proposed, which contains both spatial and temporal features. The proposed method can accurately detect the dynamic background. Additionally, several new mechanisms are also proposed to maintain and update the background model.

Spatial temporal deep feature:

For each region (segment) at each frame, our proposed STD feature is computed by concatenating a local feature and a global feature. The local feature is extracted using a region based background followed by aggregation over frames, while the global feature is computed using a block-based background whose input is a sequence of frames of the video. The STD feature extraction for a region is illustrated in Fig. 3. 1) Local Feature Extraction: A region at each frame, which is defined from a temporal segment at a frame, is fed into a region-based back ground to extract its region-based feature which is with a dimension of. As our region-based background, we use the publicly available The region-based feature contains the local context of the region but does not contain temporal information because it is computed frame-by-frame. In order to incorporate temporal information, for a region, we aggregate its region-based features over a sequence of frames, resulting in the consistent

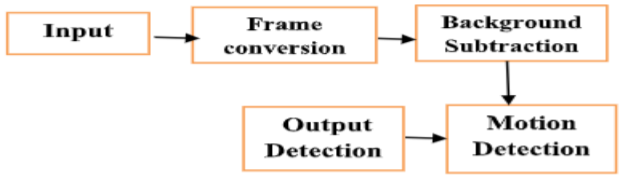
Object Detection:
In order to allow high-resolution images of the people in the scene to be acquired it is reasonable to assume that such people move about in the scene. To monitor thescene reliably it is essential that the processing time per frame be as low as possible. Hence it is important that the techniques which are employed are as simple and as efficient as possible. For that reason the well-known technique of background subtraction was selected for this application.

Background subtraction allows moving objects to be detected by taking the
Point-by-point absolute difference of the current image and a background
image which must be acquired when there are no moving objects in the scene (See equation
Moving (i; j) = |Imaget (i; j) - Background (i; j)| Such a mechanism is impractical for the surveillance system described in this paper as it may not be possible to obtain a background image with no moving objects, and more importantly the background of the scene may change due to lighting conditions or `stationary' objects being moved (e.g. a gate being opened and then left open). As Bartolome et al. point out, algorithms based on direct
Grey-level comparison are not robust enough against sudden lighting changes". It is Possible though to overcome these problems by using a dynamic background together with normalized cross correlation to evaluate any changes. The dynamic background is initialized with the first image acquired (whether or not that image contains any moving objects), and is updated if a point changes and remains changed for a number of frames

Background subtraction:
The main problem of the background subtraction approach to moving object detection is its extreme sensitivity to dynamic scene changes due to lighting and extraneous events. Although these are usually detected, they leave behind “holes” where the newly exposed background imagery differs from the known background model (ghosts). While the background model eventually adapts to these “holes,” they generate false alarms for a short period of time. Therefore, it is highly desirable to construct an approach to motion detection based on a background model that automatically adapts to changes in a self-organizing manner and without *a* priori knowledge.
We propose to adopt a biologically inspired problem-solving method based on visual attention mechanisms. The aim is to obtain the objects that keep the user attention in accordance with a set of pre defined features, including gray level, motion and shape features. Our approach defines a method for the generation of an active attention focus to monitor dynamic scenes for surveillance purposes. The idea is to build the background model by learning in a self-organizing manner many background variations, i.e., background motion cycles, seen as trajectories of pixels in time. Based on the learnt background model through a map of motion and stationary patterns, our algorithm can detect motion and selectively update the background model

ADVANTAGES:
- It is high speed in motion detection
- Better accuracy in segmentation under various illuminations
- Less time consuming process
- Flexibility in background updating model
- It is less sensitive to background noise

SYSTEM ARCHITECTURE:



APPLICATION:
- Video surveillance
- Machine Vision systems
- Object Recognition

III. CONCLUSION

A real computer vision system able to model a stationary object background or a movement object background in cluttered environment has been presented. The proposed system is based on the modelling of the structure of the scene. The quality of the detection is improved when the background is highly texture .Therefore in our future works we will use this modelling method in our object tracking system for tracking rigid and non –rigid movement object

IV. RESULT

Fig.1: Input video

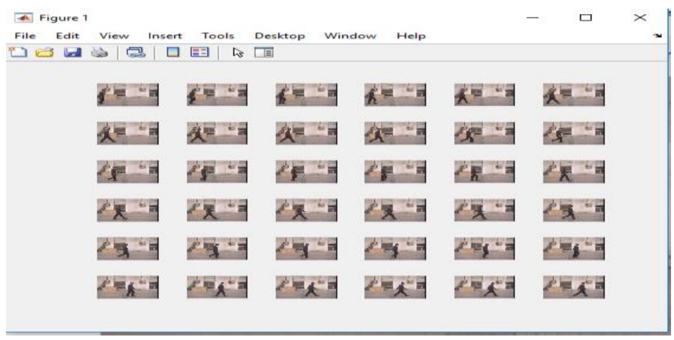


Fig.2: converted frame separations

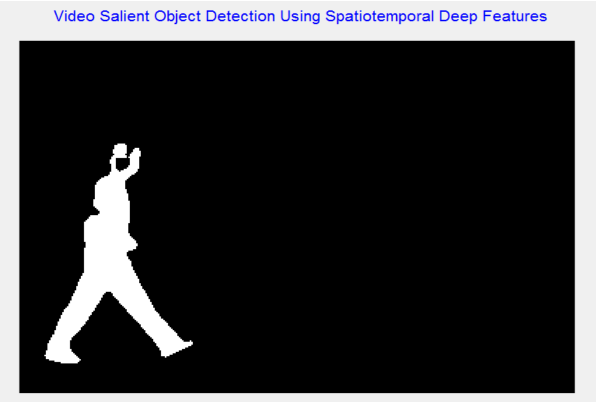


Fig.3: Background subtraction

Fig.4: object detection

V. REFERENCES

[1]. C. P. Yu, H. Le, G. Zelinsky, and D. Samaras, "Efficient video segmentation using parametric graph partitioning," in ICCV, Dec 2015, pp. 3155–3163.

[2]. T.-N. Le, K.-T. Nguyen, M.-H. Nguyen-Phan, T.-V. Ton, T.-A. Nguyen, X.-S. Trinh, Q.-H. Dinh, V.-T. Nguyen, A.-D. Duong, A. Sugimoto, T. V. Nguyen, and M.-T. Tran, "Instance re-identification flow for video object segmentation," The 2017 DAVIS Challenge on Video Object Segmentation - CVPR Workshops, 2017.

[3]. X. Li, L. Zhao, L. Wei, M.-H. Yang, F. Wu, Y. Zhuang, H. Ling, and J. Wang, "Deepsaliency: Multi-task deep neural network model for salient object detection," IEEE TIP, vol. 25, no. 8, pp. 3919–3930, Aug 2016.

[4]. G. W. Taylor, R. Fergus, Y. LeCun, and C. Bregler, "learning of spatio-temporal features," in ECCV, 2010, pp. 140–153.

[5]. W. Wang, J. Shen, and L. Shao, "Consistent video saliency using local gradient flow optimization and global refinement," IEEE TIP, vol. 24, no. 11, pp. 4185–4196, Nov 2015.

[6]. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in CVPR, June 2014, pp. 580–587.

[7]. K. Fukuchi, K. Miyazato, A. Kimura, S. Takagi, and J. Yamato, "Saliency-based video segmentation with graph cuts and sequentially updated priors," in ICME, June 2009, pp. 638–641.