International Journal of
Electro Mechanics and Mechanical Behaviour

Vol. 1: Issue 1
www.journalspub.com

An Overview of Saliency Method of Object Detection in Image Processing

Anju Vikraman V.J.[1,*], Unni Krishnan[2], Rajkumar E.[3]
[1]P.G. Student, Department of Computer Science, Rajiv Gandhi Institute of Technology, Kottayam, Kerala, India
[2]Assistant Professor, Department of Computer Science, Rajiv Gandhi Institute of Technology, Kottayam, Kerala, India
[3]Associate Professor, Department of Design and Automation, Vellore Institute of Technology, Vellore, Tamil Nadu, India

ABSTRACT

Object detection is one of the emerging approaches in the field of visual analysis. Saliency methods are applicable in object detection and attained auspicious results in the case of ordinary images. This method focuses on the salient region to diminish the computational intricacy by evaluating the contingency of the foreground regions. The objective of this paper is to study various saliency detection methods. The discussed methods are region-based saliency detection, spatiotemporal saliency detection, low rank and sparse decomposition in motion saliency, background priors-based saliency, and video saliency detection model.

Keywords: *motion saliency, object recognition, saliency detection, sparse decomposition*

****Corresponding Author***
E-mail: anjuvikramanrit17@gmail.com

INTRODUCTION

Human eyes are capable of detecting an object from images as well as in complex aggressive videos. The main aim of visual analysis is to consummate human perception so as to isolate an object from a dynamic video. With the advancement in video object extraction and saliency detection, the gap between computer vision and human eye capacity will be lowered. Saliency methods are useful in many areas like target tracking, boundary detection, video or image compression, and behavior detection. Emphasizing the most attractive features of a video will help for video summarization, compression and surveillance.

There are two mechanisms available for getting visual attention: top-down and bottom-up approaches. In top-down methods, prior knowledge gathered from the supervised learning is used for detecting saliency. Bottom-up methods are based on some predefined postulates. Color intensity and orientation maps are calculated and merged to get final feature maps based on center-surrounded differences of various Gaussian pyramids. Convolutional neural networks also obtained high accuracy in visual saliency predictions.

Video object extraction methods will detect and extract significant objects from the videos. Saliency methods play a vital

role in the detection of salient objects from the background. A video sequence is divided into a set of frames and saliency methods are applied separately into the individual frames. Then saliency map of each individual frame is merged to generate the final saliency map of a video sequence. Moving objects will acquire more attention than that of stationary objects. This capability makes saliency methods more precise in videos so that contrast-based method can be applied directly.

DIFFERENT METHODS FOR SALIENCY DETECTION

Saliency detection methods are classified into different types as follows: biologically inspired algorithms, purely calculation-based and hybrid methods. Based on visual attention, the methods can be classified as top-down and bottom-up methods. Based on processing areas, methods are classified into space-domain and frequency-domain methods. This paper discusses about region-based saliency detection, spatiotemporal saliency detection, sparse and low-rank decomposition in motion saliency, background priors-based saliency, video saliency detection model, and wavelet transform using low-level features.

Region-Based Saliency Detection

In [1], Ren et al. discussed about region-based saliency detection in object recognition. First, they introduced an effective region-based calculation for saliency detection. For solving the object recognition task, image features are encoded using the obtained saliency map. Superpixel based adaptive mean shift algorithm is used to extract semantically and perceptually meaningful salient regions form the input images. For each superpixel, saliency is computed using spatial briefness; it is calculated from the Gaussian mixer model based clustering. Modified page rank algorithm is used to propagate the salient features from one cluster to another (Figure 1).

They also applied saliency map for object recognition to improve the detection accuracy. Weighted sparse coding is introduced to learn about the discriminative features. This method can detect large salient region detection as well as it provides noise tolerance in messy backgrounds.

Spatiotemporal Saliency Detection

Trung-Nghia Le et al. [2] proposed saliency-based object detection using spatiotemporal deep features in video. These deep features are used for utilizing the global and local contexts over different frames. They introduced a spatiotemporal conditional random field (STCRF) to calculate the saliency from spatiotemporal deep features. Temporal information generated from the STCRF describes the relationships between the nearest regions. STCRF concentrates on the temporal consistent saliency maps, so that the method can provide more accurate detection of salient regions (Figure 2).

The proposed method works as follows: input video is segmented into multiple scales. For each scale, saliency map is calculated from STCRF algorithm and scale-level spatiotemporal features. Saliency calculations from different scale levels are merged to form final saliency map. Experimental analysis shows that the method works better than other state-of-art methods.

Low–Rank and Sparse Decomposition Method

Xue et al. [3] studied about motion saliency which has more impact in video-processing operations like image and video segmentation, classification, object recognition, and adaptive compression. This method utilizes a low-rank and sparse

decomposition on video, which splits along X–T and Y–T planes. To preserve the information about the motion, object spatial details are also adopted (Figure 3).

Input video is sliced into X–T and Y–T plane slices. Saliency maps are calculated from each plane slices. Final map is obtained through merging the plane saliency maps. Normally, the low-rank constituents always represent the background information and the sparse elements provide the idea of mobile objects in the foreground. This method can efficiently characterize foreground motion objects from background without any additional background processing. Additionally, adaptive threshold selection and noise abolition methods provide better accuracy, and the method outperforms temporal spectrum residual.

Background Priors-Based Saliency

Liu et al. [4] discussed about saliency detection based on background priors. In this model, border set is obtained by taking image border superpixels. To obtain true background superpixels, superpixels with strong image edges are removed. The above process will reduce the background noises. Then background saliency maps and centered anisotropic Gaussian distribution were calculated. Initial saliency map is computed by merging background saliency maps and anisotropic Gaussian distributions (Figure 4).

Initial saliency map is refined through smoothness constraints based on some assumptions like neighbors have similar saliency information. This is achieved through minimizing the continuous quadratic saliency energy function on a graph. It can easily detect the salient regions accurately and suppress the background clutters effectively.

Video Saliency Detection Model

In [5], Fang and Lin investigated a new saliency detection model for videos, which is based on feature contrast in compressed domain. Using discrete cosine transform, color, luminance and texture features for unpredicted frames are extracted. Motion features are extracted from the video using motion vectors for predicted frames. Unpredicted frames consist of three features and a corresponding static saliency map is computed. Finally, motion saliency map and static saliency map are combined (Figure 5).

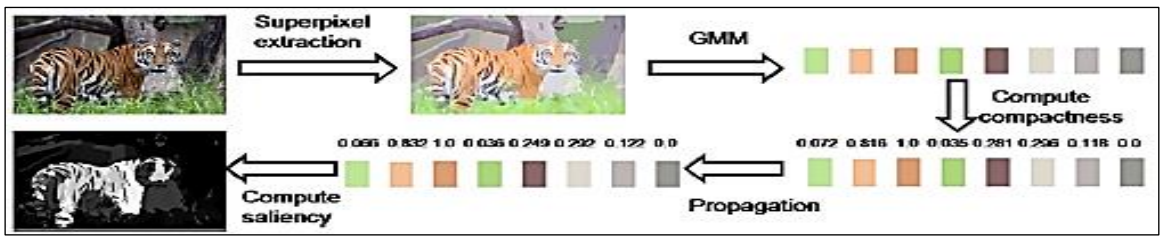

Fig. 1. *Framework of proposed saliency detection. Color rods represent the GMM components.*

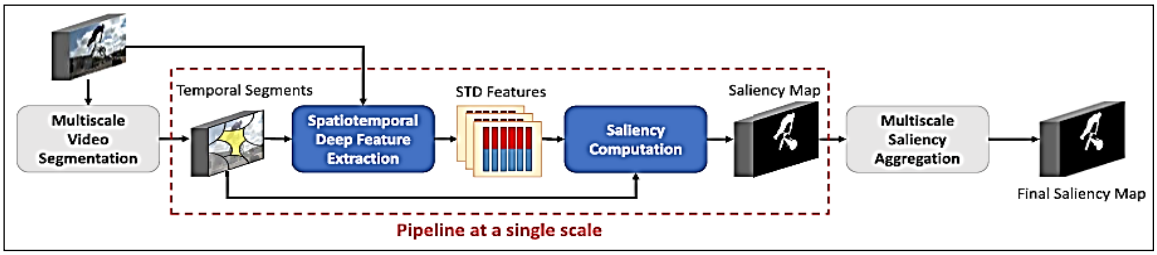

Fig. 2. *Framework of proposed method.*

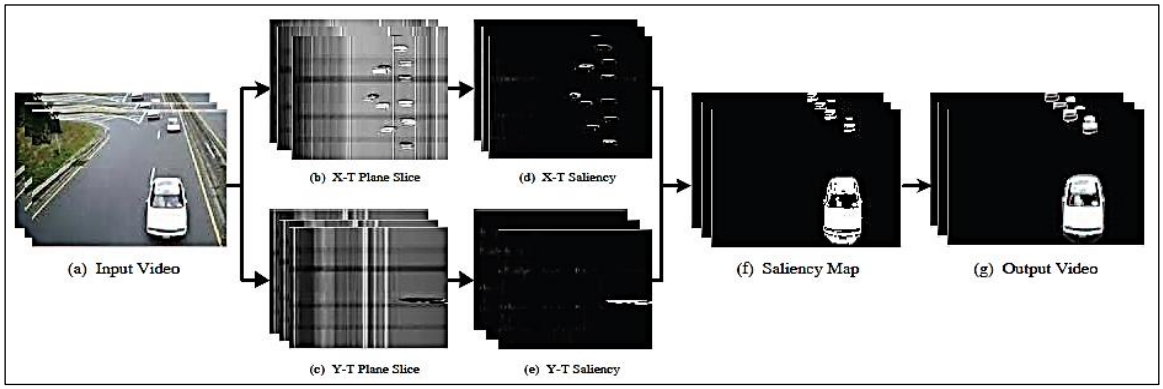

Fig. 3. *Illustration of proposed method.*

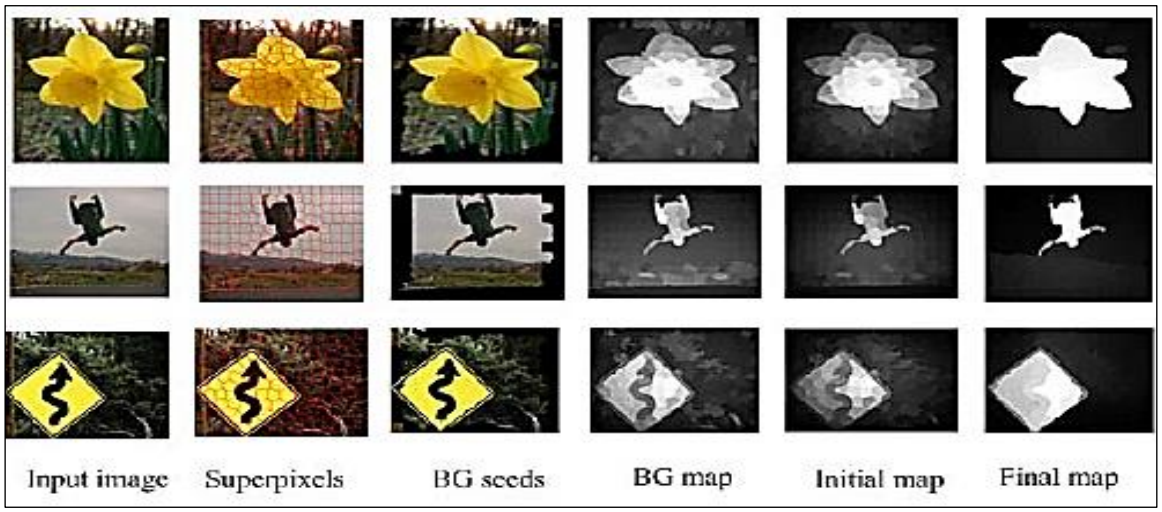

Fig. 4. *Proposed saliency detection.*

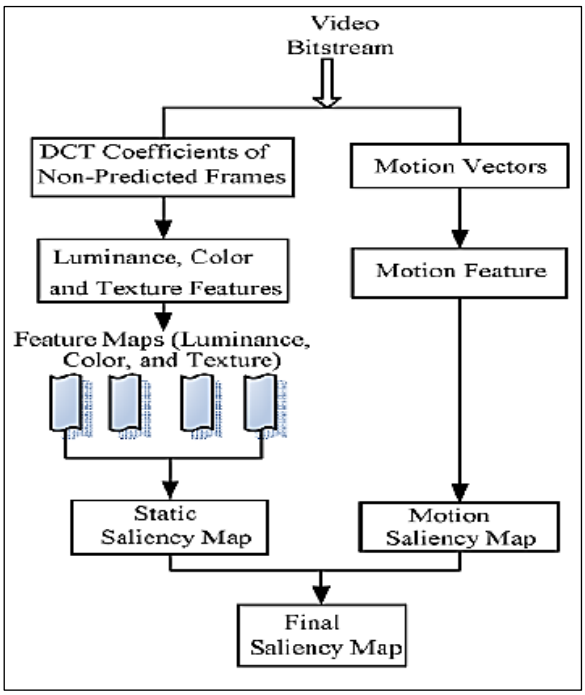

Fig. 5. *Video saliency detection model.*

CONCLUSION

Saliency detection is a very popular topic in recent years. These methods provide so many advantages over video segmentation, classification and object detection. This paper presents a brief review about some of the saliency detection methods. Discussed methods are region-based saliency detection, spatiotemporal saliency detection, low-rank and sparse decomposition in motion saliency, background priors-based saliency, and video saliency detection model. This survey can give some ideas about saliency methods and advantages of each method.

REFERENCES

[1] Ren Z, Gao S, Chia L-T, Wai-Hung Tsang I. Region-based saliency detection and its application in object recognition. *IEEE Trans Circuits Syst Video Technol*. 2014; 24(5).

[2] Le T-N, Sugimoto A. Video salient object detection using spatiotemporal deep features. *IEEE Trans Image Proc*. 2018; 27(10).

[3] Xue Y, Guo X, Cao X. Motion saliency detection using low-rank and sparse decomposition. *IEEE Trans Circuits Syst Video Technol*. 2014; 24(1).

[4] Liu Z, Gu G, Chen C, Dong C, Lin C. Background priors based saliency object detection. *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA).* 2016.

[5] Fang Y, Lin W, Chen Z, Tsai C-M, Lin C-W. A video saliency detection model in compressed domain. *IEEE Trans Circuits Syst Video Technol*. 2014; 24(1).

Cite this Article: Anju Vikraman V.J., Unni Krishnan, Rajkumar E. An Overview of Saliency Method of Object Detection in Image Processing. *International Journal of Electro Mechanics and Mechanical Behaviour*. 2019; 1(1): 1–5p.

