A Survey on Visual Saliency Detection and Computational Methods

Ankit Garg[#1], Ashish Negi [*2]

[# 1]Assistant Professor, Department of Computer Science & Engineering,
Amity University, Haryana, India
[1]ankirgitm@gmail.com
[# 2]Associate Professor, Department of Computer Science & Engineering,
GBPC, Pauri, Garhwal, India
[2]ashish.ne@gmail.com

Abstract—In computer vision saliency detection comprises wide range of methods to detect salient object present in the image. These methods focus how important object can be detect from the image. Results of these methods are fully dependent upon the type and quality of input image. In this paper saliency detection methods, Fixation prediction models, saliency map generated through various saliency detection algorithms and its evaluation measures are analyzed.

Keyword - Saliency, Seam carving, Saliency Map, Fixation Prediction, Object Detection

I. Introduction

With the advancement of computer science and information technology information can be transfer with the use of images. Saliency region detection and object segmentation is also known as salient object detection. To preserve structure of important objects present in the image saliency detection algorithm are used. In image resizing accurate saliency map helpful to find out accurate salient object that must be preserve during image resizing in this way improved algorithm can be developed with accurate result where saliency is required to detect. Accurate saliency detection not only improves performance of the algorithm but also reduces computational time. Methods under saliency detection are classified in to two types of space domain based and frequency domain based [1]. Human visual system also has capability to provide information about salient object in the image. In computer vision saliency detection and salient object segmentation are the process through which most salient object can be detected and boundary of salient object can accurately defined [2]. Literature of saliency detection suggests two models: Top-down model and bottom up model.

Bottom up approach pay more attention in color, intensities and edge orientation and top down approach pay more attention on high level features such as face, body and text and their structure. In order to produce better saliency map bottom up approach combines top down approach.

Proposed models of saliency detection have been used to measure the exclusiveness of a location for a salient object which mostly attracts human attention [3]. Researchers are trying to improve the efficiency of salient map which can detect large salient region in image, well defined boundaries of object, uniformly highlight full salient object, disregard noise and other blocking artifacts in the image and also should be capable to output full resolution saliency map. Most of the methods in saliency detection are useful for detection of region of interest (ROI). ROI of an image contains pixels which are similar to each other and are capable to fascinate attention of human eyes.

In [4], Liu et al. proposed region based pixel-wise saliency detection algorithm which can detect prominent object or region from the image and generate full resolution energy map. Saliency detection models and algorithm suggested by the researchers are widely used in various multimedia disciplines like image compression [5], content aware image resizing [2], image adaptation [6], image editing [7].

In [8], Itti, presented a model which was based on saliency detection for rapid scene analysis. The results of this suggested model was very much close to human perception. In [9] Long Mai et al. described how to design features of a good saliency map.

In [10] X. Fan presented a Rapid serial visual presentation (RSVP) based on fixation status and saccade status (shifting process). In RSVP amounts of contents are displayed serially, each for a brief period of time. Visual attention shifting is important psycho physiological activity. Human visual attention can be directed on the most interesting part of the image. Navigation of images on small display device can be improved by simulating the human visual fixation and saccade status in a way similar to RSVP. Due to difficulty in measuring human attention some methods are used on eye movement predictions for estimating an abrupt spasmodic movement of eye. Attention movement can be depicted using fixation status and abrupt spasmodic movement of eye. The combination of both composes the whole simulation process. In image salient object can be straight lines and other regular objects.

In [11] J. Kiess et al. presented an algorithm to improve image retargeting by preserving the shape of objects present in the image. This algorithm is based on seam caving and some additional features i.e. line detection and preservation of object. When a seam intersects the straight line then energy local neighborhood of that pixel is increase to remove the effect of object distortion.

In [12] Y. Ma et al. presented a framework which is fast in information retrieval i.e. attended view, attended areas and attended points based on contrast analysis.

In [13] M. Ma et al. presented image cropping technique to detect ROI regions present in the image. User can save ROI region or block or image to save bandwidth and memory resources. ROI suggestion engine is used to determine ROI region. Selection of ROI region can be based on entropy based ROI region selection and image segmentation.

In [14] A. Santella et al. presented Gaze based photo cropping algorithm to identify ROI in the image. This algorithm uses fixation point data to identify the eye movement of the viewer on ROI of the image. For qualitative evaluation of a crop set of compositional rules are identified and implemented. This approach creates pleasing crops without explicit interaction.

In [15] F. Liu et al. presented a nonlinear fisheye view map to find ROI of the image. Fisheye view mapping uses low level and high level object recognition. Using nonlinear image wrapping less important content in the image are deemphasizes instead of completely removal.

In [16] E. Golub discussed manual selection of point of interest using Rule of Third and golden mean and golden spiral. In [5] proposed a region based computational attention model. This model can detect ROI regions of any size. To predict gaze transmission a ROI ranking method is proposed. Based on ROI ranking method an adaptive image browsing sequence is generated. To predict ROI region in saliency map some sensitive factors of human visual system are considered i.e. area factor, global effect and contextual difference.

In [17] F W M. Stentiford presented a cognitive visual attention model. This CVA model is generalization of visual attention. This approach model is based on dissimilarity between neighborhoods in the image. In [2] S. Avidan suggested an approach to reduce size of image by removing less important pixels from the image. In this process optimal seams are identified horizontally and vertically using dynamic programming. A Seam is a horizontal and vertical path of pixels which are not important in the image. After removal of seam image content are readjusted to minimize the effect of seam removal from the image. Important area detection is also a part of this algorithm which draws human attraction. Importance measure can be computed using saliency map.

In [18] H. Jiang et al. proposed salient object segmentation framework which combines both bottom up salient stimuli and object level shape prior according to three characteristics of a salient object. Saliency defined in this paper is based on super pixels. Super pixels are generated after fragmenting the image. Context based saliency and object level shape prior computation are combined in to an iterative energy minimization framework leading to binary segmentation of salient object. Main difference of proposed method with previous method is the use of object level shape prior in to consideration to define better salient object. This paper presents principle method of saliency detection. Analysis on image database is performed to depict run time comparison of saliency detection algorithms.

II. METHODS UNDER SALIENCY DETECTION

Saliency detection methods can be classified in to two domain i.e. space domain and frequency domain. Saliency detection algorithm based on space domain has high computation cost. These space based algorithm are not commonly used in real time system. Saliency detection algorithm based on frequency domain overcomes this problem and gives better way for parameter selection.

A. Saliency Detection Algorithm Based on Space Domain

In [19] Ming-Ming Cheng et al. proposed regional contrast based salient object extraction algorithm. Two global contrast based saliency computation methods are used, namely Histogram based contrast (HC) and spatial information enhanced region based contrast (RC). HC maps assign pixel-wise saliency value simply on color separation from all other image pixels to produce full resolution saliency maps. We compute color histogram of image and reduce number of color by ignoring less frequently occurring colors and choose more frequently occurring colors. In region based contrast (RC) image is segmented into regions and compute color contrast at region level and finally define saliency for each region as the weighted sum of the region's contrasts to all other regions in the image.

In [20] J. Harel presented a bottom up visual saliency model based on graph computation, GBVS, including a frame work for activation and normalization/combination. A comparison of GBVS is made against existing benchmarks on a data set of gray image of natural environment. The model present in this paper is better than classical algorithm of Itti & Koch.

In [21] S. Geoferman et al. proposed context aware saliency (CA) which detects the salient region of the image. Presented algorithm is based on 4 principles observed in the psychological literature:

1. Area having distinctive color can be considered as salient region. Conversely region having similar color value can be considered as low salient region.

2. Frequently occurring features must be suppressed.

3. Pixels which are salient must be grouped and should not be scattered in the image.

4. High level factor such as human face.

According to principle (1)-(2) pixels is salient when its color is different from all other pixels present in the image. In visual organization rule which is the third principle of context aware saliency algorithm the distance between clusters is also an important aspect. Clusters which are related to background can have many similar clusters. Clusters with less positional distance are considered as salient cluster and conversely clusters with high positional distant are considered as less salient.

B. Saliency Detection algorithm Based on Frequency Domain

In [22] Y. Fang proposed method to detect saliency map in spatial domain. Method presented in paper first extract the spectral residual of an image. The proposed method is general purpose object detection method. Proposed method is based on long spectra representation of image. Based on comparison made in this paper it is proved that proposed method is fast and reliable computational model for early stage video processing.

In [23] Achanta R. et al. proposed frequency tuned approach. Frequency tuned approach uses color, texture, and shape to compute image saliency. Saliency map using this approach is faster and gives full resolution of saliency map. Proposed saliency map is capable for object segmentation. Frequency tuned approach is computationally efficient and can detect salient regions with well defined boundaries. Range of spatial frequency retained by the proposed method is more appropriate than the algorithm used for the comparison.

In [24] C. Guo et al. proposed saliency detection model based on phase Spectrum of Quaternion Fourier Transform which is the key of this approach to detect salient region of scene. This model only consider bottom up information. This method is better than SR method and has two limitations. Firstly, algorithm cannot deal with closure pattern. Secondly, if noise is very similar to salient feature this method is fail to detect salient feature. Algorithm is computationally fast and can be used in engineering application like object recognition and video image processing.

III. ABOUT SALIENT OBJECT DETECTION

Salient object detection and segmentation has two stages:

1. Detection of most salient objects which human choose as most important region in the image.

2. Accurately segmenting boundary of the objects.

A wide range of salient objects detection and fixation prediction model has been developed. The aim of these models is to segment the object and prediction maps is used to find various salient objects in the scene. In salient object detection or salient object segmentation accuracy is major problem. An algorithm which is best suited for one image may not be best suited for the other images.

A good saliency detection algorithm must satisfy three criteria:

1. A saliency detection algorithm should be capable to identify important object from the image which predict high degree of human visual attention and the degree of marking background as important salient region should be less.

2. Saliency map should have high resolution so that exact image information and location of salient object can remain intact.

3. Computation efficiency should be good even in case of complex scene analysis.

Traditional work performed by Itti et al. [8] and Liu et al. [25] on saliency adopts both stages.

IV. A REVIEW ON STATE OF ART METHODS

In this section we review three categories:

1. Models for image saliency detection based on top-down and bottom-up approach.

2. Model for human visual fixation prediction

3. Various applications of salient object detection.

Models between years (1998-2007) are based on fixation prediction and between years (2007-Now) are based on salient object detection and segmentation of most important visual subset of the image.

V.   MODELS FOR SALIENT OBJECT DETECTION

Traditional methods had been proposed for detecting salient object from the image. For identifying visual subset of the image two attributes of traditional approaches can be considered.

A. Block based vs. Region based

Many traditional approaches used for saliency detection has taken Blocks as a visual subset, which represent pixels and patches. With the emergence of superpixel algorithm region based saliency detection approaches are widely used [26].

B. Intrinsic Cues vs. Extrinsic cues

On the basis of two attribute i.e. Extrinsic and intrinsic cue saliency detection algorithm can be categorized in the subgroups. (a) Block based model with intrinsic cue (b) Region based model with intrinsic cue (c) Models with extrinsic cue [27]. Block based model with intrinsic cue is based on blocks i.e. pixels and patches. This model is not efficient as compared to region based model with intrinsic cue because no of region in the image are less than number of blocks present in the image. A potential block based algorithm is required to extract sophisticated features in the image. The key idea in salient object detection is to separate salient objects form distracters. In local saliency (intrinsic) is computed on the basis of patch. Saliency of each patch of the image is computed against remaining patches of the same input image. In global (extrinsic) image saliency is computed on the whole image.

Bottom-up models detect salient object or salient locations from the image which are visually important and attract human attention. Bottom up approach pay more attention in color, intensities and edge orientation and top down approach pay more attention on high level features such as face, body and text and their structure. In order to produce better saliency map bottom up approach combines top down approach.

Cognitive Models, Information Theoretic Models, Graphical Models, Decision Theoretic Models, Spectral Analysis Models, Pattern Classification Models are 7 categories in to which above bottom up saliency detection models can be classified. Top down salient object detection model use cognitive science which includes study of human behavior and artificial intelligence with a focus on how nervous systems represent, process, and transform information.
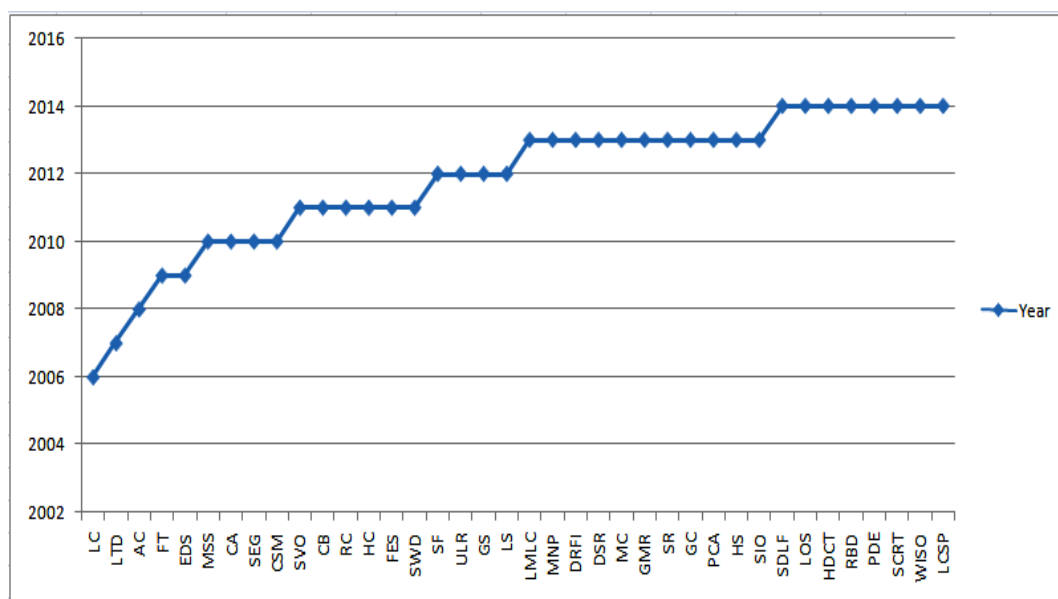

Fig. 1. Salient Object Detection Models

TABLE I Bottom Up Saliency Detection Models with 8 different factors[32]

S. No.	Saliency Detection Models	Year	BU	TD	Sp($\sqrt{}$) /SpT($\times$)	Static	Dynamic	Sp-B ($\sqrt{}$) /Ob-B ($\times$)	Features
1	Itti et al.	1998	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	Cr, In, Or
2	Privitera & Stark	2000	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	-
3	Salah et al.	2002	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	Or
4	Sun & Fisher	2003	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\times$	Cr,In,Or
5	Gao & Vasconcelos	2004	$\times$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	$\times$	DCT
6	Ma et al.	2005	$\sqrt{}$	$\times$	$\times$	$\sqrt{}$	$\times$	$\sqrt{}$	M*
7	Zhai & Shah	2006	$\sqrt{}$	$\times$	$\times$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	SIFT
8	Harel et al.	2006	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	In, Or
9	Cerf et al.	2007	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	Cr, In, Or,F
10	Hou & Zhang	2008	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	ICA
11	Zhang et al.	2008	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	DoG, ICA
12	Kienzle et al.	2009	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	P
13	Geoferman et al.	2010	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	Cr, F
14	Murray et al.	2011	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	Cr, In, Or
15	Wang et al.	2011	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	ICA

TABLE II TOP-DOWN SALIENCY DETECTION MODELS WITH 8 DIFFERENT FACTORS[32]

S. No.	Saliency Detection Models	Year	BU	TD	Sp($\sqrt{}$) /SpT($\times$)	Static	Dynamic	Sp-B ($\sqrt{}$) /Ob-B ($\times$)	Features
1	McCallum	1995	$\times$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	-
2	Ramstorm & Christiansen	2002	$\times$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	$\sqrt{}$	Cr, In
3	Sprague & Ballard	2003	$\times$	$\sqrt{}$	$\times$	$\times$	$\sqrt{}$	$\times$	S*
4	Paletta et al.	2005	$\times$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	$\times$	SIFT
5	Jodogne & Piater	2007	$\times$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	$\times$	SIFT
6	Verma & McOwan	2009	$\sqrt{}$	$\times$	$\sqrt{}$	$\sqrt{}$	$\times$	$\times$	Cr, In, Or
7	Borji et al.	2010	$\times$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	$\times$	Cr, In, Or
8	Borji et al.	2012	$\times$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\times$	Cr, In, Or

BU- Bottom-Up Saliency Detection Model, TD- Top-Down Saliency Detection Model, Sp- Spatial, SpT- Spatial Temporal, Sp-B- Space Based, Ob-B- Object Based, Cr, In, Or: Color, Intensity, Orientation, DCT: Discrete Cosine Transformation, M*:Motion, Static and auditory Saliency, DoG: Gaussian DoG Filtering, ICA: Independent Component Analysis. SIFT: Scale Invariant Feature Transform, s*: Color Matching, Depth and Line, F: Face.

VI. FIXATION PREDICTION MODELS

A number of fixation prediction models has been developed which predict a set of fixation points which mostly fascinate human attention while looking at some image. Most of the fixation prediction models are based on low level features of the image i.e. color, orientation and intensity. If image is complex having a large number of objects and complex background, it will be difficult for algorithm to determine which object in the image is salient object and will attract more human attention.

Researchers through their research has proven that human attention is not only depend upon low level factors of the image but also depend upon high level factors of the image such as action, event etc. Some algorithm combines saliency detection and fixation prediction models to scale up existing models. Fixation prediction models have been applied on image dataset to show that human attention usually directed towards high level factors and the fixation time of these factors are more than fixation time of low level factors.

TABLE III Fixation Prediction Models [26]

S.No.	Fixation Prediction Model	Publication	References
1	IT	Pattern Analysis and Machine Intelligence	Itti's et al.
2	Attention based on Information Maximization (AIM)	Journal of Vision	Bruce & Tsotsos
3	Graph Based (GB)	Nural Information processing System	Harel et al.
4	Spectral Residual (SR)	Computer Vision and Pattern Recognition	Hou & Zhang
5	Saliency using natural Statistics (SUN)	Journal of Vision	Zhang et al.
6	SeR	Journal of Vision	Seo & Milanfar
7	SIM	Computer Vision and Pattern Recognition	Murray et al.
8	SS	Pattern Analysis and Machine Intelligence	Hou et al.
9	Covariance Saliency (COV)	Journal of Vision	Erdem & Erdem
10	Boolean Map based Saliency (BMS)	International Conference on Computer Vision	Zhang et al.
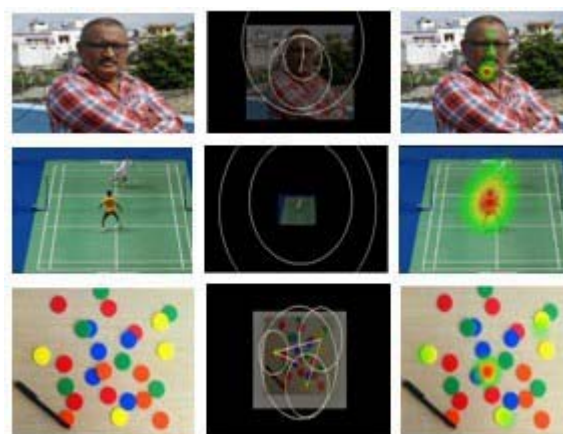
Fig. 2. (a) Original Image (b) Set of fixation prediction (c) Attention Map

TABLE IV Fixation Calculation applied on different images a

ID	Subject Name	Fixation Count	Start Time	Length	Position X	Position Y
123	Image1	1	0	1150	683	383.70834350585938
124	Image1	2	1650	700	704.33331298828125	289.39999389648438
125	Image1	3	2400	2250	702.86956787109375	197.60868835449219

TABLE V Fixation Calculation applied on different images b

ID	Subject Name	Fixation Count	Start Time	Length	Position X	Position Y
126	Image2	1	0	2900	688.0169677734375	379.62710571289062
127	Image2	2	2950	1900	711.4102783203125	333.61538696289062

TABLE VI Fixation Calculation applied on different images c

ID	Subject Name	Fixation Count	Start Time	Length	Position X	Position Y
128	Image3	1	0	1400	677.44830322265625	395.10345458984375
129	Image3	2	1550	550	489.33334350585938	345
130	Image3	3	2600	800	838.64703369140625	279.35293579101562
131	Image3	4	4050	800	737.64703369140625	511.11764526367188

VII.   APPLICATIONS OF SALIENT OBJECT DETECTION

Salient object detection is a technology which detects most interesting or important objects from a digital image or video. Among a wide range application automatic image cropping [28], Content aware retargeting [2], Feature Extraction [29], Image segmentation [30], Image/video compression [31] are most common applications.

VIII.   FEATURES TO MEASURE QUALITY OF SALIENT MAP

A range of features are available to measure salient object detection result. A good saliency map is one that satisfies following range of features. These features are useful to measure quality of saliency detection results.

Fig 3. (a) Original mage, (b) Image after cropping, (c) Image after compression, (d) Image after applying content aware image retargeting algorithm, (e) Image after applying feature extraction, (f) Image after segmentation, (g) Image stitching.

Saliency Coverage- Quality of saliency map is depending upon on the area covered by salient pixels. A saliency which covers abnormally large or small area of salient object is dubious to be a good saliency map.

1. Compactness of saliency map- In this feature we compute density of salient pixel distribution in salient object present in the image.

2. Saliency using Histogram- Quality of saliency map is depending upon the distribution of saliency value in the saliency map. A good salient map can separate salient object from the background.

3. Color Separation- A saliency map that is capable to distinguish color of background and salient object can be considered as good saliency map.

4. Segmentation- A saliency map should generate good segmentation results.

5. Boundary Quality- A saliency map should generate precise and well defined object boundaries. A parameter can be defined to measure boundary feature of saliency map. High value of parameter will give accurate defined saliency map with strong edges.
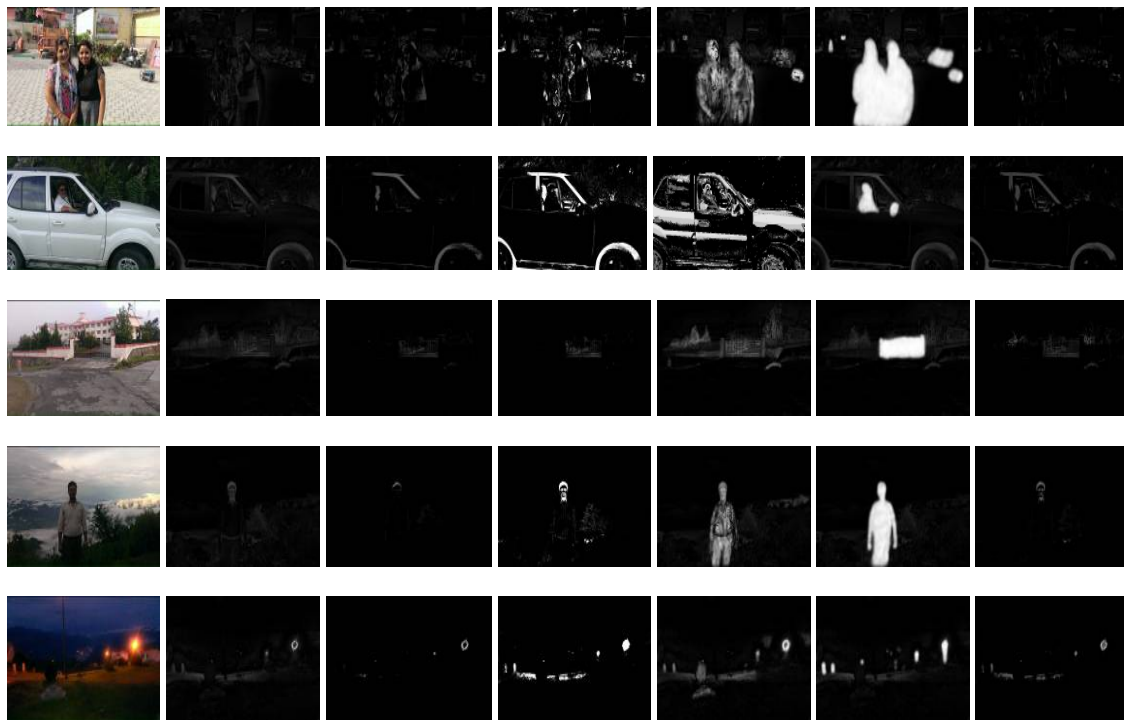
Fig. 4. Example of saliency Map (a) Original Image, (b) FT, (c) AC, (d) CA, (e) GB, (f) SR, (g) MR
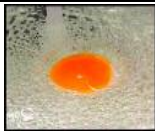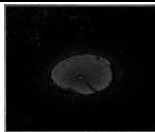
IX.   DATASET AND EVALUATION MEASURES

Various Image dataset are available to measure saliency detection algorithms. A good saliency detection algorithm is one which can detect small as well as large salient regions from the image. MSRA is the largest saliency dataset which contain 25000 images.

Most of the saliency detection algorithms test their efficiency with the use of this dataset. Saliency databases can contain images with following categories.

1.   Images having small salient objects.
2.   Images which contains large salient objects.
3.   Images contains intermediate salient object.
4.   Images contains complex background
5.   Image having repetitive background or foreground patterns
6.   Images which comprises both small and large salient objects.

To detect visual saliency in video datasets are also available.

ABLE VII Overview of popular Saliency Image datasets

S.No.	Data Set	Year	Images	Objects	Salient Object Annotation	Image Resolution	Original Image	Saliency Detection
1	MSRA-A	2007	20K	~1	Bounding Box	400 X 300		
2	MSRA-B	2007	5K	~1	Bounding Box	400 X 300		
3	ASD	2009	1000	~1	Pixel-wise	400 X 300		
4	iCoSeg	2010	643	~1	Pixel-wise	~500 X 400		
5	MSRA5K	2011	5K	~1	Pixel-wise	400 X 300		
6	ImgSal	2013	235	~2	Pixel-wise	640 X 480		
7	Judd-A	2014	900	~5	Pixel-wise	1024 X 768		

In order to compute various saliency detection methods some standard measures are available. First Two methods are (1) Precision-recall (PR) (2) ROC (Receiver operating characteristics)

With the use of two above saliency measures F-measure is calculated which use recall and precision and Area under Curve (AUC). Binarization is required in saliency measure evaluation.

To calculate precision –Recall (PR) we convert saliency map S into binary mask (M). After getting binary mask we compare it with the ground truth G.

To calculate precision

$$\Pr ecision = \frac{\left| BinaryMask(M) \cap GroundTruth(G) \right|}{\left| BinaryMask \right|} \qquad (1)$$

$$\mathrm{Re} call = \frac{\left| BinaryMask(M) \cap Groundtruth(G) \right|}{\left| GroundTruth \right|} \qquad (2)$$

Binary mask of image can be find out after applying some image processing operations i.e thresholding and image segmentation. This binary mask has two possible values for each pixel in the image 0 or 1.

Algorithm for binarization can be divided into two categories.

1. Local binarization- In this threshold is calculated pixel by pixel
2. Global binarization- In this one threshold is set for full image.

Various methods are available for binarization based on local and global binarization.

Ankit Garg et al. / International Journal of Engineering and Technology (IJET)

Methods for Local Binarization:
1. Nilback method
2. Adaptive method
3. Sauvola method
4. Bernsen method

Method for Global Binarization:
1. Fixel Threshold
2. Otsu method
3. Kitter method

F-measure- F-score or F-measure in terms of saliency detection is used to measure the accuracy of the result. Precision (P) and recall (R) is used to calculate the test score.

F-measure is the Harmonic mean of the precision (P) and recall (R).

$$F1 = 2.\cfrac{1}{1/P + \cfrac{1}{R}} = 2.\frac{\Pr ecision.\operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \qquad (3)$$

$$P = \frac{\text{No. of positive result}}{\text{No. of all positive result}} \qquad (4)$$

$$R = \frac{No\ of\ correct\ Positive\ \operatorname{Re} sult}{No\ of\ Positive\ \operatorname{Re} sult\ that\ must\ be\ return} \qquad (5)$$

To give more weight to recall or to precision use weighted F measure

$$F\beta = \frac{(1+\beta^2).\operatorname{Re} call.\Pr ecision}{\operatorname{Re} call.\beta^2.\Pr ecision} \qquad (6)$$

If $0 < \beta < 1 \rightarrow$ Give more weight to precision

If $\beta > 1 \rightarrow$ Give more weight to recall.

ROC Curve- It is used to plot true positive rate (TPR) and False positive rate (FPR) at various threshold setting.

TPR is also called sensitivity or probability of detection.

FPR is also called Fallout or Probability of false alarm.

$$TPR = \frac{\sum True positive}{\sum False positive} \ , \ FPR = \frac{\sum False positive}{\sum Condition Negative} \qquad (7)$$

When binarization saliency map M with a set of some threshold FPR and TPR will be

$$TPR = \frac{\left| Binary Map \cap Ground Truth \right|}{\left| Ground Truth \right|} \qquad (8)$$

$$FPR = \frac{\left| Binary Mask \cap Ground Truth \right|}{\left| Binary Map \cap Ground Truth \right| + \left| Opposite\ of\ Binary\ Mask \cap Opposite\ of\ ground\ truth \right|} \qquad (9)$$

Area Under ROC Curve: Accuracy of the test can be measured using Area Under Curve (AUC).

Area 1 represents test is perfect and Area 0.5 represents worthless test.

Mean Absolute Error: MAE is used to measure how are we close to the predictions which we made earlier. MAE is computed between salinecy map (S) and Ground truth (G). The formula for measuring mean absolute error is

$$MAE = \frac{1}{W\ X\ H}.\sum_{x=1}^{w}\sum_{y=1}^{H}\left\| S(x-y) - G(x,y) \right\| \qquad (10)$$

X.   CONCLUSION

This paper presents various methods for saliency detection in the image. Salient object of image fascinate human visual attention. Identification of salient object or region is field of research. Salient map are generated through different saliency detection algorithm. Saliency detection algorithm can be based on space domain and frequency domain. A review on saliency detection models based on top down and bottom up saliency is also presented. There are various features which are useful for evaluating the quality of saliency map of the image.

Some fixation prediction models are also reviewed in this paper and fixation calculation is applied on various images. Various datasets are available to measure the saliency detection algorithm.

In this paper overview of popular saliency image dataset are given. Finally in the last section evaluation measures (ROC Curve) is described.

XI.   FUTURE WORK ON SALIENCY DETECTION

1.   Existing saliency detection algorithm takes more time for execution. Run time of complex saliency algorithm must be minimized in order to work on real time projects.

2.   Various saliency detection algorithms had now been developed which are based on bottom up saliency detection models. New methods for top down saliency detection models should be developed.

3.   Range of applications of existing saliency detection algorithm is limited it should be expended to machine learning, signal processing and data mining.

4.   Existing algorithm can work efficiently on image which contains small objects and less complex background. Algorithm should be improved to detect object having large size and complex background.

REFERENCES

[1]   Runxin MA, Yang YU, Xiaomin YUE, "Survey on Image Saliency Detection Methods", In: Proc. of International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, pp. 329-338, 2015.
[2]   S. Avidan and A. Shamir," Seam carving for content-aware image resizing", ACM Transactions on Graphics, vol. 26, no. 3, 2007.
[3]   Xu, L., Zeng, L., Duan, H., & Sowah, N., Saliency detection in complex scenes. EURASIP Journal on Image and Video Processing, 2014(1), 31.
[4]   Manipoonchelvi P., Muneeswaran K., " Region –based saliency detection", In IET image processing ,Vol. 8, Iss. 9, pp. 519–527,2014.
[5]   C. Christopoulos, A. Skodras and T. Ebrahimi, "The JPEG2000 still image coding system: an overview," J. Consumer Electronics, IEEE Transactions on, 2000, vol. 46, no. 4, pp. 1103-1127.
[6]   J. Sun, J. Xie, J. Liu, S. Member, T. Sikora, and S. Member, "Image Adaptation and Dynamic Browsing Based on Two-Layer Saliency Combination," vol. 59, no. 4, pp. 602–613, 2013.
[7]   M. Ding and R. F. Tong, "Content-aware copying and pasting in images," The Visual Computer, 2010, vol. 26, pp. 721-729.
[8]   L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Transactions on pattern
[9]   analysis and machine intelligence, vol. 20, no. 11, pp. 1254-1259,1998.
[10]   Mai L, Liu F, " Comparing salient object detection results without ground truth", In: Proc. of European Conference on Computer Vision, pp. 76–91, 2014.
[11]   Fan X, Xie X, Ma WY, Zhang HJ, Zhou H, "Visual attention based image browsing on mobile devices", In: Proceedings of the IEEE international conference on multimedia and expo (ICME 03), Baltimore, July 2003.
[12]   J. Kiess, S. Kopf, B. Guthier, and W. Effelsberg. Seam carving with improved edge preservation. In Proceedings of IS&T/SPIE Electronic Imaging (EI) on Multimedia on Mobile Devices, vol. 7542, no. 1, pp. 1 – 11, 2010.
[13]   Y.-F. Ma and H.-J. Zhang, "Contrast-based image attention analysis by using fuzzy growing", In ACM International Conference on Multimedia, 2003.
[14]   Ma, M., Guo, J.K. , "Automatic Image Cropping for Mobile Devices with Built-in Camera", Proc. Consumer Communication & Networking Conf. pp. 710-711, 2004.
[15]   A. Santella, M. Agrawala, D. DeCarlo, D. Salesin, and M. Cohen, "Gaze-based interaction for semi-automatic photo cropping", In CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems, pp. 771–780, New York, NY, USA, 2006. ACM.
[16]   F. Liu and M. Gleicher, " Automatic image retargeting with fisheye-view warping", In Proc. of the 18th annual ACM symposium on User interface software and technology, pp. 153–162, 2005.
[17]   Golub, E., "Photocropr: A first step towards computer-supported automatic generation of photographically interesting cropping suggestions," tech. rep., University of Maryland HCI Lab, 2007.
[18]   Stentiford, F. W. M, "An attention based similarity measure with application to content based information retrieval", In Storage and Retrieval for Media Databases, M. M. Yeung, R. W. Lienhart, C-S Li, Editors, Proc. SPIE Vol. 5021, pp. 20-24, 2003.
[19]   H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, and S. Li. "Automatic salient object segmentation based on context and shape prior", In British Machine Vision Conference (BMVC), vol. 6, no. 6, 2011.
[20]   Cheng, M.M., Mitra, N.J., Huang, X., Torr, P.H.S., Hu, S.M.," Salient object detection and segmentation", Tech. rep., Tsinghua University. http://mmcheng.net/SalObj/. Submission NO. TPAMI-2011-10-0753,2011.
[21]   J. Harel, C. Koch, and P. Perona," Graph-based visual saliency. Advances", In Neural Information Processing Systems, vol. 19, pp. 545–552, 2007.
[22]   Goferman S.,Zelnik-Manor L.,Tal A., ""Context-Aware Saliency Detection", IEEE Transactions on Pattern Analysis and Machine Intellegence, vol. 34, no. 10, pp. 1915-1926, 2012.
[23]   Hou X., Zhang L.,"Saliency Detection: A Spectral Residual Approach", In IEEE Conference on Computer Vision and Pattern Recognition,pp. 1-8, 2007.
[24]   Achanta R., Hemami S., Estrada F., Susstrunk S.,"Frequency-tuned salient region detection" , In IEEE Conference on Computer Vision and Pattern Recognition, pp. 1597-1604, 2009.
[25]   C. Guo, Q. Ma, and L. Zhang, "Spatio-Temporal Saliency Detection Using Phase Spectrum of Quaternion Fourier Transform," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1-8, 2008.

[26] Liu T.,Yuan Z.;Sun J., Wang J.,Zheng N., Tang X.,Shum H., " Learning to Detect a Salient Object", In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 2, pp. 353-367, 2011.
[27] A. Borji, M.-M. Cheng, H. Jiang, and J. Li., "Salient object detection: A survey." [Online]. Available: http://arxiv.org/abs/1411.
[28] 5878, 2014.
[29] Wang M., Konrad J., Ishwar P., Jing K., Rowley H.,"Image saliency: From intrinsic to extrinsic context", In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 417-424, 2011.
[30] Cao, C., Liu, J., & Zuo, Y. Automatic image cropping via the novel saliency detection algorithm. Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS, pp. 955–958, 2013.
[31] Zhang, L., He, B., Song, Y., & Yan, T., Underwater image feature extraction and matching based on visual saliency detection. OCEANS 2016 - Shanghai, pp. 1–4, 2016..
[32] Li, Q., Zhou, Y., & Yang, J., Saliency based Image Segmentation, In International Conference in Multimedia Technology, 2011, pp. 5068–5071.
[33] Ho-Phuoc, T., Dupret, A., & Alacoque, L. , Saliency-based data compression for image sensors. 2012 IEEE Sensors, pp. 1–4, 2012.
[34] Ali Borji, State-of-the-Art in Visual Attention Modeling, In IEEE Transaction On Pattern Analysis and Machine Intellegence, Vol. 35, No. 1,pp. 185-207, 2013

AUTHOR PROFILE

Ankit Garg is pursuing Ph.D from Uttarakhand Technical University, Dehradun. He is working as assistant professor in Amity University, Haryana. He has completed MCA and M.Tech. He has published a number of good research articles in International journals and conferences.

Dr. Ashish Negi has done his Ph.D and working as associate professor in Govind Ballabh Pant Engineering College, Pauri Garhwal. He has published a number of good research articles in International journals and conferences.