Volume 4, Issue 4, April 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Wavelet Domain Based Saliency Map Computation and Evaluation Using Novel Quadrant Based Approach

Manish Mahajan
Associate Professor &Ph.D. Research Scholar
Department of IT
Chandigarh Engineering College, Landran
PTU, Kapurthala, India

Dr. Navdeep Kaur
Associate Professor & Head
Department of CSE
SGGSWU
Fatehgarh Sahib, India

Abstract: Saliency map is an important and very useful representation of visually important parts of the image. It highlights the salient parts of the image by limiting the visual information coming to the human visual system (HVS). Many models have been proposed in the literature for computing the saliency maps. All of them are having their own pros and cons. No model is perfect. But it has been analyzed that instead of looking for perfect saliency model one should use the concept of multiple models each having its own advantages. From the entire available models one should choose the saliency model which is most appropriate for the application under scan. Further rather than going for high contrast or full resolution saliency maps one should go for saliency maps with acceptable resolution, giving well defined boundaries and highlighting bigger as well as smaller objects of importance. This is because saliency maps generated are generally not the final output to evaluate, rather the application, for which computed saliency map has to be used, is being evaluated. So increasing the performance of application should be more emphasized rather improving the visual quality of saliency map. Keeping this concept in mind we have proposed wavelet domain based saliency using all four components of wavelet decomposing. All four components have been used to ensure that all visually important objects, whether small or large, are highlighted with acceptable resolution. Finally this work has been evaluated on various parameters and it has been found that proposed method is good for saliency detection as compared to other state-of-art methods. An evaluation method for measuring effectiveness of the method for locating true salient region has also been proposed in this work.

Keywords: Saliency may, Wavelet domain, Receiver operating characteristics, Precision, Recall

I. INTRODUCTION

Humans do not process the entire area of an input visual image uniformly, but focus their attention on a limited area (attended area) in the field of view and then shift their attention from one place to another, depending on the situation and task. Research on human visual characteristics shows that people only perceive clearly a small region of 2-5^o of visual angle. The human retina possesses a non-uniform spatial resolution of photoreceptors, with highest density on that part of the retina aligned with the visual axis (the fovea), and the resolution around the fovea decreases logarithmically with eccentricity [1]. What's more, research results show that observers' scan paths are similar, and predictable to some extent [1]. These research results provide a new pathway to encode images based on human visual characteristics which will allow us to only encode a small number of well selected interesting regions (attention regions) with high priority to keep a high subjective quality, while treating less interesting regions with low priority to save bits. Recently, many subjective-quality-based image coding methods have been developed. According to the way of obtaining attention regions, they can be coarsely classified into four categories, as follows: (1) In a first approach, considering that human attention prediction is still an open problem, human machine Interaction methods are adopted to obtain the attention regions. (2) A second class of approaches uses machine vision algorithms to automatically detect interesting regions. For instance, due to the importance of human faces while people perceive the world [5], [6], it is reasonable to consider that human faces may likely constitute interesting regions. (3) A third class of approaches uses knowledge about human psychophysics to guide the encoding process. (4) The fourth class of approaches exploits recent computational neuroscience models to predict which regions in image are more likely to attract human attention and to be gazed at.

II. LITERATURE REVIEW

Saliency estimation methods can broadly be classified as biologically based, purely computational, or a combination. In general, all methods employ a low-level approach by determining contrast of image regions relative to their surroundings, using one or more features of intensity, color, and orientation. In general, bottom-up VS models extract feature information by means of contrast representation be it for color, intensity, orientation, or any other low-level feature. This contrast representation is performed by obtaining the difference of a region in the image relative to their surroundings. Theoretically, this method of obtaining the contrast is akin to the center surround process in the human eye. In the biological category one of the reputable works is by Itti and Koch [7] whose method was based upon biologically plausible architecture proposed by Koch and Ullman [10]. They determine center-surround contrast using a Difference of Gaussians (DoG) approach. Frintrop et al. [9] present a method inspired by Itti's method, but they compute center surround differences with square filters and use integral images to speed up the calculations.

In computational VS models, low-level features and the contrast approach are still used but the model is not constructed based on any biological mechanism. The contrast is mainly obtained through the use of Euclidian distance in different sized window filters [8, 11, 12]. The contrast images obtained in the works of [8, 11, 12] are summed to form the final saliency map.

The third category of methods is those that incorporate ideas that are partly based on biological models and partly on computational ones. For instance, Harel et al. [13] create feature maps using Itti's method but perform their normalization using a graph based approach.

Recently, there is a trend to model VS computationally in the frequency domain [2, 3]. In the works of Hou and Zhang [2], a spectral residual approach was used to generate the saliency map. The saliency map is the inverse of the spectral residual. The saliency map of this method is rather accurate in providing the locations of important regions in a given visual scene but is terribly low in resolution. In [3], Achanta et al. debated that the saliency map should have well-defined borders, uniformly highlighting the object if it is salient, and most of all; the saliency map should be in high resolution. In the authors' opinion, without conforming to the points mentioned [3], the saliency would have limited usefulness in certain applications. Therefore, Achanta et al. proposed a method which generates the saliency map solely by contrast representation.

Although the approach used by Achanta et al. gives high resolution maps which has its usefulness in some applications but in many other applications such as content based image retrieval (CBIR) all that matters is the detection of salient objection with acceptable resolution. In fact, the approach used in [3] will eliminate many small detailed objects and textures which could be of importance when the smoothing is applied to the spatial domain. Furthermore, as long as the saliency map provides the correct location of important objects and is of reasonable resolution (object can be visually identified), the map can be considered acceptable. In another approach Discrete wavelet transform was used by Christopher et al. [4] to compute the saliency map. This approach is successful in detecting salient regions in an image with acceptable resolution but it considers only the contrast of LL band, it totally ignores the other three detail bands. This act will eliminate the smaller & finer details. But sometimes some important information may be present in these detail components. To solve this problem we have proposed a saliency model based on wavelet transform domain by processing approximation (LL) as well as all the three detail coefficients (LH, HL, HH).

III. PROPOSED WORK

The input colored image is first converted to lab color lab so as to make it device independent. Also the L component distinguishes the intensity or luminance component from the color information. Then taking the all the l, a, b components individually we performed the single level DWT decomposition. After this we got the four individual components named LL1, HL1, LH1, HH1 for all the three l, a, b images. Then for each component individually we calculate the contrast image using the Euclidean distance with the help of following formula:

$$C(x,y) = \sqrt{(l_\mu(x,y) - l_\mu)}^2 \ \ldots\ldots\{1\}$$

Where l_μ is the mean of LL component of L image of L, a, b component images & l_μ (x, y) is the mean intensity of all pixels of LL component of L sub-image similarly 12 contrast images will be calculated for four sub-bands of each component image of Lab color space. Then inverse DWT operation will be performed taking four processed sub-bands of each sub-image to get processed l, a, b sub-image L_p, a_p, b_p. then these processed components will be normalized to the range [0,255].These processed sub-images will be combined to get the saliency map using the following formula:

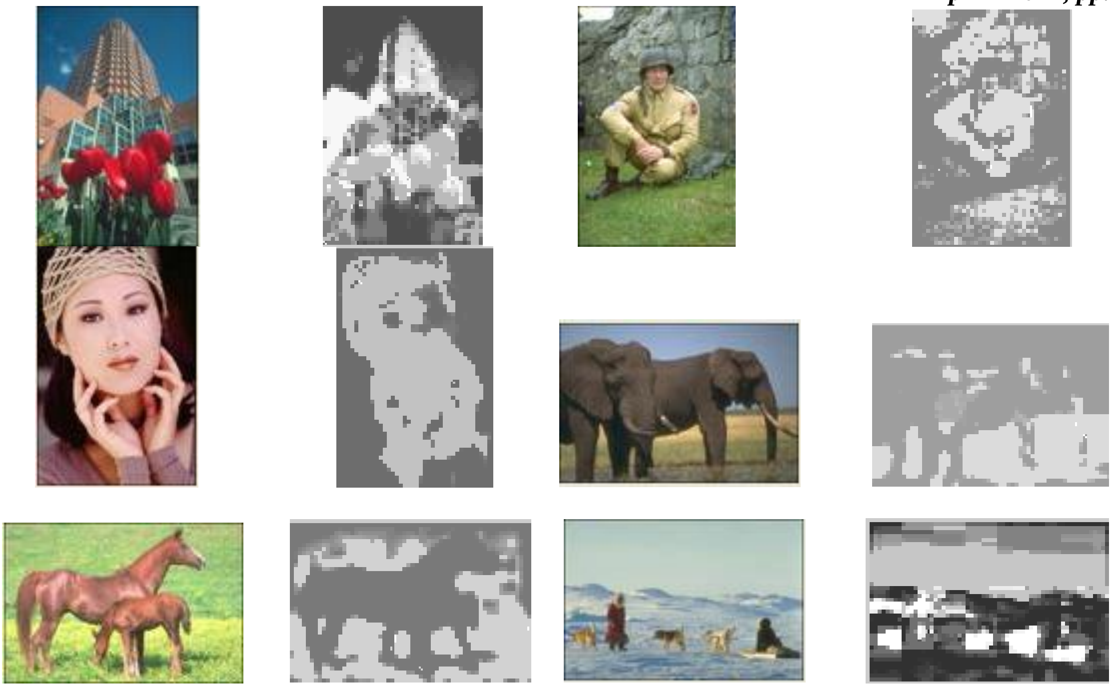
$$Sm = L_p(x,y) + a_p(x,y) + b_p(x,y) \ldots\ldots\{2\}$$

Where Sm is the saliency map and L_p a_p, b_p are processed sub-images of Lab Color lab.

Again the saliency map will be renormalized to the range [0,255] to get the final saliency map. Finally we can equalize the histogram of saliency map to get sharper saliency maps.

IV. RESULTS & COMPARISON

Proposed saliency map has been evaluated visually as well as objectively. Various visual results are shown below in figure 1.



(a) (b) (c) (d)

Fig. 1 Column (a) & (c) are original images and column (b) & (d) are their corresponding saliency maps

Also proposed work is evaluated on various parameters like precision, recall, f-measure, TPR and FPR. Here precision is calculated as ratio of total saliency, that is sum of intensities in the saliency map captured inside the user annotated rectangle to the total saliency computed for the image. Recall is calculated as the ratio of the total saliency captured inside the user annotated rectangle to the area of rectangle. F-measure is the overall performance measurement as the weighted harmonic mean between the precision and recall values. σ is real and positive constant which decides the importance of precision over recall. In our work σ is taken as 0.3 because precision is more important than recall. Precision has been taken with importance of 70% and recall with importance of 30%. True positive rate (TPR) is percentage of intensity assigned to true salient region or user annotated rectangle to total intensity assigned by proposed method. Similarly False positive rate (FPR) is percentage of intensity assigned to false salient region or outside user annotated rectangle to total intensity assigned by proposed method. The equations for the parameters are as follows:

$$\text{Precision} = \frac{\sum_x \sum_y (t(x,y) \times s(x,y))}{\sum_x \sum_y s(x,y)} \quad \text{-------------------- \{Equation 3\}}$$

$$\text{Recall} = \frac{\sum_x \sum_y (t(x,y) \times s(x,y))}{\sum_x \sum_y t(x,y)} \quad \text{-------------------- \{Equation 4\}}$$

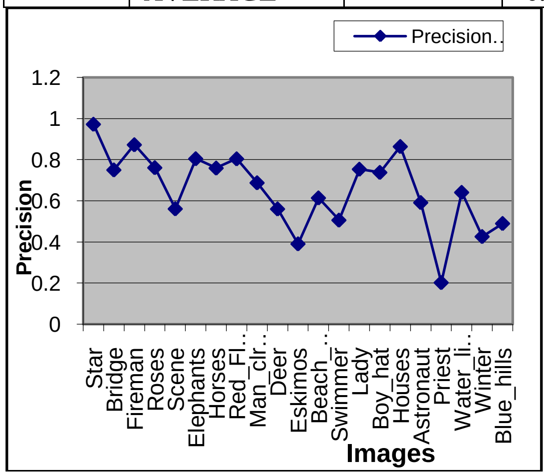
$$\text{F- Measure} = \frac{(1+\sigma) \text{ X P X R}}{(\sigma \text{ X P}) + \text{R}} \quad \text{-------------------- \{Equation 5\}}$$

All the results for above parameters have been compiled in the following table 1. It may be seen that values for all the parameters for almost all the images lies above the expected values. The strength of any proposed method may be truly analyzed only in comparison of values for state-of-art methods. For this proposed method has been compared with waveieee [4] and frequency-tuned method [3]. The comparisons have been shown inn figure 2. (f-l), figure 3, 4 and 5 below.

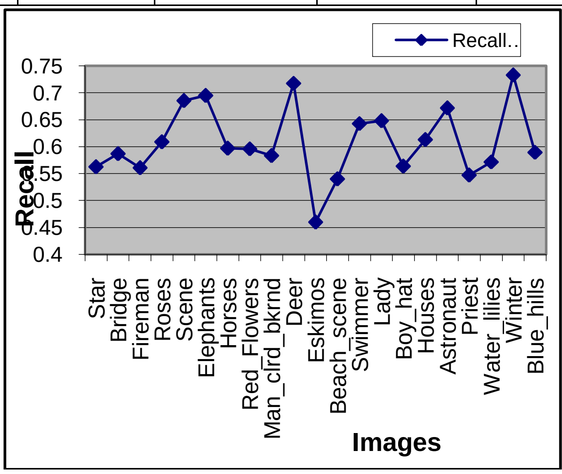
TABLE 1:
RESULTS OF PROPOSED METHOD ON VARIOIUS PARAMETERS

Sr. No	**Image**	**Size**	**Precision**	**Recall**	**F_Measure**	**TPR**	**FPR**
1	Star	120 X 80	0.9726	0.5627	0.8326	0.5627	0.4109
2	Bridge	80 X 120	0.7503	0.5869	0.7050	0.5869	0.4556
3	Fireman	120 X 80	0.8728	0.5606	0.7734	0.5606	0.4944

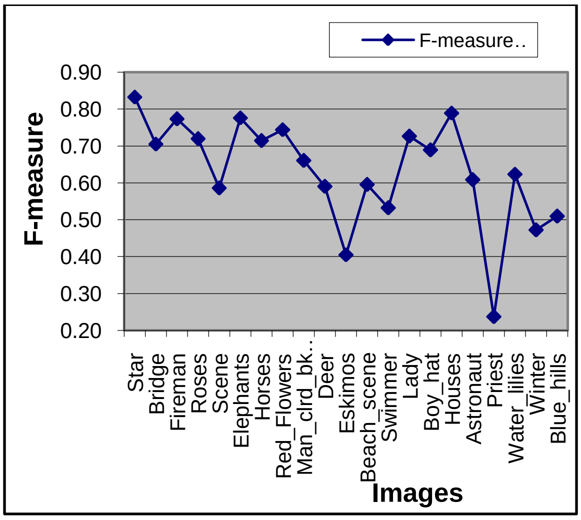
4	Roses	80 X 120	0.7613	0.6087	0.7197	0.6087	0.4109
5	Scene	120 X 80	0.5612	0.6853	0.5857	0.6853	0.53
6	Elephants	120 X 80	0.8048	0.6947	0.7764	0.6947	0.6828
7	Horses	120 X 80	0.7593	0.5971	0.7145	0.5971	0.6634
8	Red_Flowers	120 X 80	0.8043	0.5961	0.7443	0.5961	0.4377
9	Man_clrd_bkrnd	120 X 80	0.6878	0.5833	0.6605	0.5833	0.6314
10	Deer	120 X 80	0.5606	0.7176	0.5904	0.7176	0.6402
11	Eskimos	120 X 80	0.3908	0.46	0.4049	0.46	0.5712
12	Beach_scene	120 X 80	0.6142	0.5401	0.5954	0.5401	0.5599
13	Swimmer	120 X 80	0.5061	0.6429	0.5322	0.6429	0.5499
14	Lady	80 X 120	0.7538	0.648	0.7264	0.648	0.473
15	Boy_hat	120 X 80	0.7382	0.5638	0.6890	0.5638	0.6141
16	Houses	120 X 80	0.8641	0.6129	0.7894	0.6129	0.5007
17	Astronaut	120 X 80	0.5915	0.6718	0.6083	0.6718	0.7305
18	Priest	120 X 80	0.2024	0.5472	0.2368	0.5472	0.6084
19	Water_lilies	800 X 600	0.6411	0.5717	0.6236	0.5717	0.4207
20	Winter	800 X 600	0.4263	0.733	0.4719	0.733	0.4079
21	Blue_hills	800 X 600	0.4898	0.5895	0.5097	0.5895	0.6019
	AVERAGE		**0.585523**	**0.624965**	**0.5756124**	**0.6249654**	**0.566608**



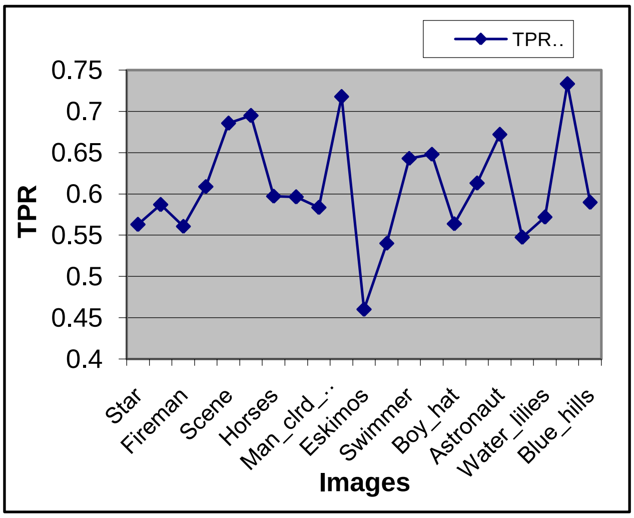
(a)



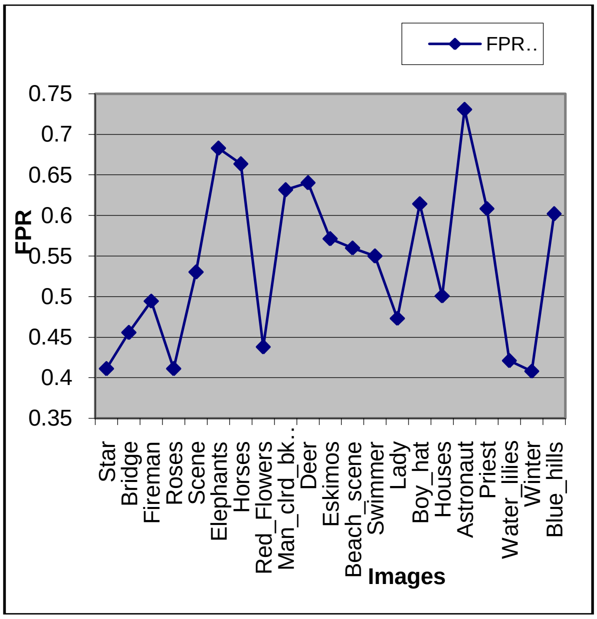
(b)



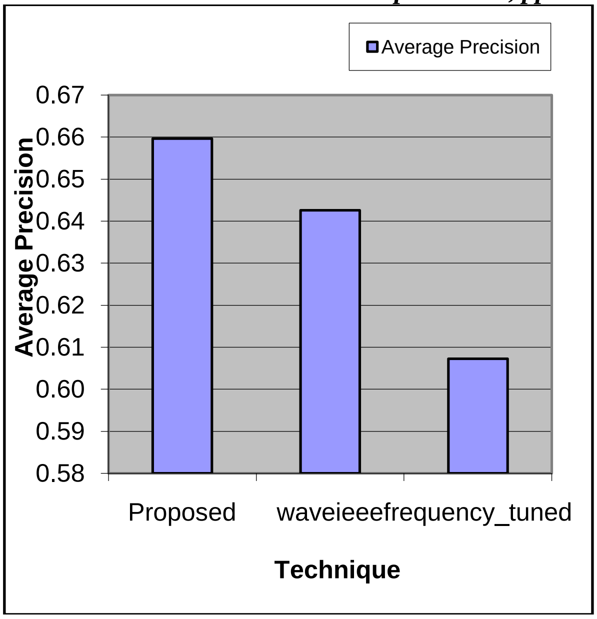
(c)



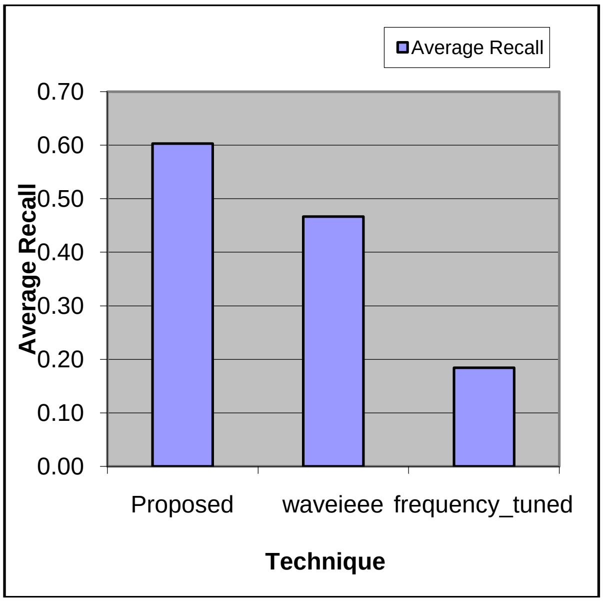
(d)


FPR..
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
FPR
Star
Bridge
Fireman
Roses
Scene
Elephants
Horses
Red_Flowers
Man_clrd_bk..
Deer
Eskimos
Beach_scene
Swimmer
Lady
Boy_hat
Houses
Astronaut
Priest
Water_lilies
Winter
Blue_hills
Images

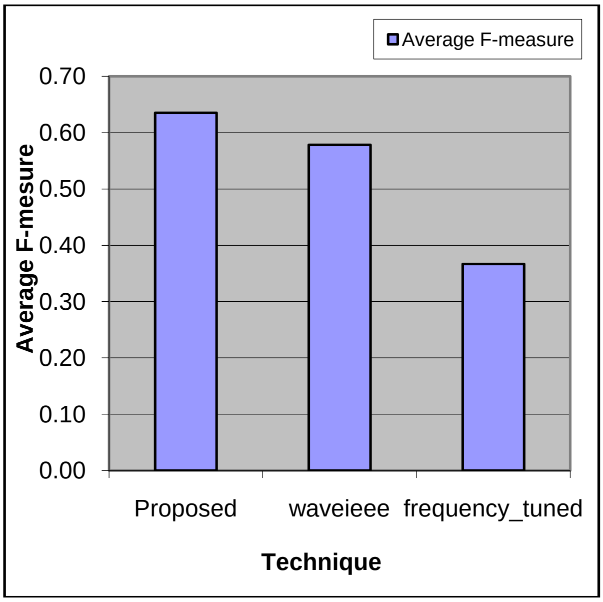
(e)


Average Precision
0.67
0.66
0.65
0.64
0.63
0.62
0.61
0.60
0.59
0.58
Average Precision
Proposed
waveieeefrequency_tuned
Technique

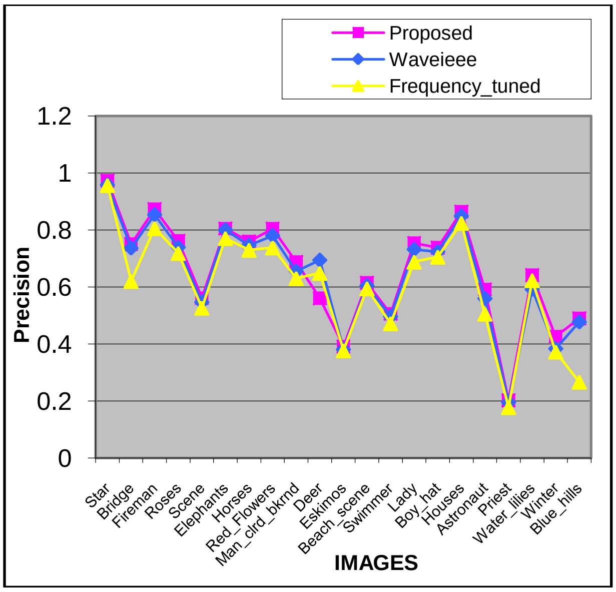
(f)


Average Recall
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Average Recall
Proposed
waveieee
frequency_tuned
Technique

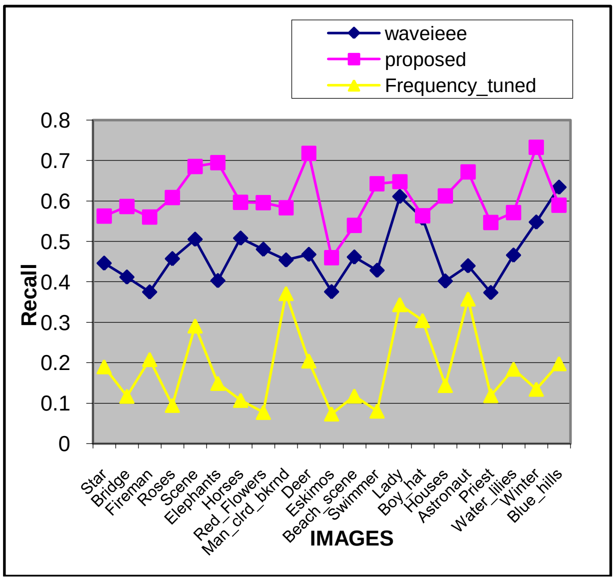
(g)


Average F-measure
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Average F-mesure
Proposed
waveieee
frequency_tuned
Technique

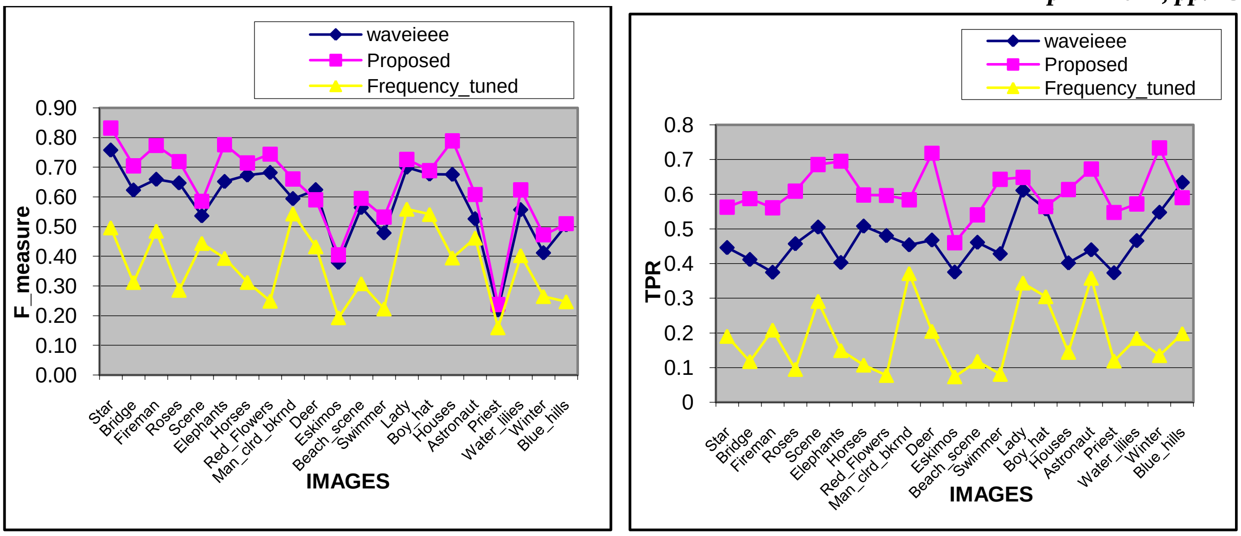
(h)


Proposed
Waveieee
Frequency_tuned
1.2
1
0.8
0.6
0.4
0.2
0
Precision
Star
Bridge
Fireman
Roses
Scene
Elephants
Horses
Red_Flowers
Man_clrd_bkrnd
Deer
Eskimos
Beach_scene
Swimmer
Lady
Boy_hat
Houses
Astronaut
Priest
Water_lilies
Winter
Blue_hills
IMAGES

(i)


waveieee
proposed
Frequency_tuned
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Recall
Star
Bridge
Fireman
Roses
Scene
Elephants
Horses
Red_Flowers
Man_clrd_bkrnd
Deer
Eskimos
Beach_scene
Swimmer
Lady
Boy_hat
Houses
Astronaut
Priest
Water_lilies
Winter
Blue_hills
IMAGES

(j)



(k) (l)

Fig. 2. (a)-(e) Results for Precision, Recall, F-measure, TPR, FPR respectively for proposed method; (f)-(l) comparison of proposed method with [3] & [4] for various parameters.

In figure 2 (a-e) above results for proposed method may be seen for various parameters like Precision, Recall, F-measure, TPR, FPR in (a)-(e) respectively. Also the comparison of proposed method with waveieee [4] and frequency-tuned [3] method has been shown graphically for the same parameters. It may be seen that graph for proposed method always rides upon the graph of other two methods in comparison. These graphs witness the strength of proposed method.

Another important parameter is receiver operating characteristics (ROC) curve. This curve is representative of the effectiveness or accuracy of the method taking TPR and FPR as base for measurement. In this graph TPR values are taken at y-axis and FPR values are taken at x-axis. Then the points where TPR value coincides with FPR value of respective images are marked in space. It is considered that if most of the values are above the diagonal line (x=y, 45^o) then the method is effective in marking the true salient region. On the other hand if most of the values are below diagonal line then method is considered not to be effective in marking the right salient region. ROC for proposed method & two other state-of-art techniques is shown below:

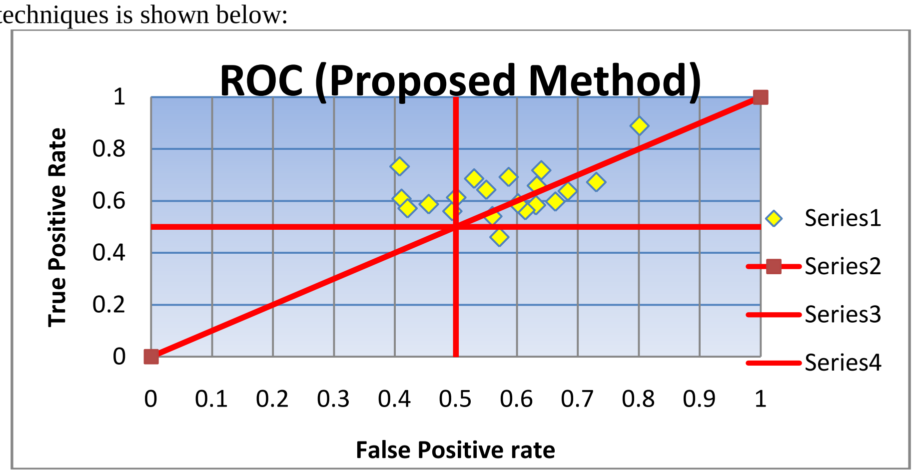


Fig 3: ROC for proposed method

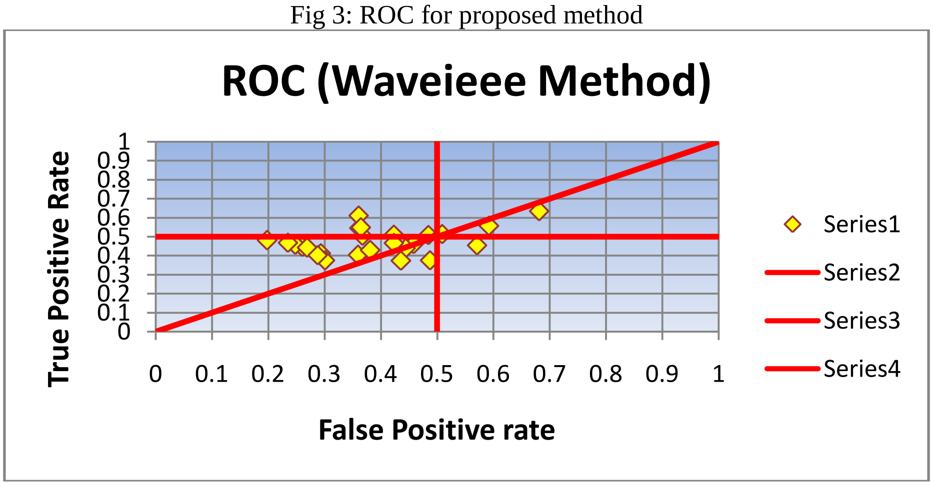


Fig 4: ROC for waveieee method [4]

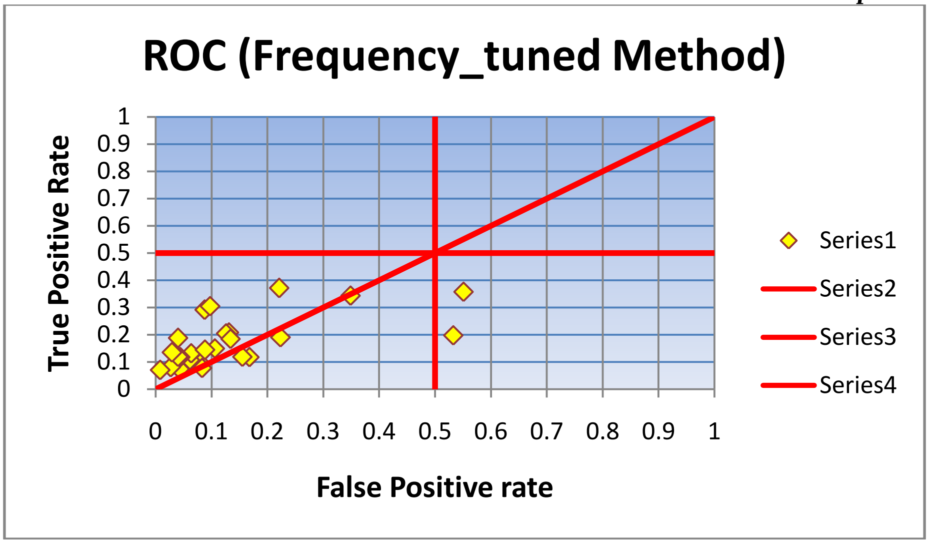


Fig 5: ROC for Frequency-tuned method [3]

It may be seen from figure 3, 4 and 5 above that about 66% values in all the three methods are above the diagonal line. It indicates that the methods are effective in locating the right salient regions in 2/3 of the cases which is acceptable. But using only the diagonal line does not concludes about the fact that which method is more effective. To deal with the problem another alternative has been proposed. In this work ROC has been divided into four quadrants. The quadrants have been numbered I, II, III and IV starting from top left in clock-wise fashion. Now it is proposed that the ranking may be given to different quadrants. Quadrant I has been given rank 1, as it is having high values for TPR and low values for FPR which is most desirable. Quadrant III has been given rank 4 as it is having low values for TPR and high values for FPR which is most undesirable. Quadrant II has been given rank 2 as although it is having high values of FPR but also having high values for TPR. Quadrant IV has been given rank 3as it is having low values for FPR as well as for TPR. Low TPR values means it is not performing whereas high values for both TPR and FPR means the method is performing good but with high error rate which is the case with quadrant II.

According to above theory it may be seen in figure 3 above that for waveieee method [4] most of the values lies in the quadrant IV (rank 3) and very few values lies in (quadrant I, II and IV) which is not good sign for the performance of the method. Similarly in figure 4 above for frequency-tuned method [3] it may be seen that all of the values lies in the quadrant IV (rank 3) and quadrant II (rank 4) which clearly indicates the algorithm is not effective. On the other hand, if one takes a look upon the figure 2 above for proposed method it may be noticed that most of the values lies in quadrant I and II (rank 1 and 2 respectively). This clearly indicates that proposed method is much more effective than state-of-art methods for locating the correct salient regions.

V. CONCLUSIONS

In this paper an effective method for computing saliency map has been proposed. The method has been evaluated on parameters like precision, recall, F-measure, TPR, FPR and ROC. It has been found that proposed method is performing well for almost all the parameters. Further a ROC based evaluation technique has been proposed which evaluates the technique for its effectives based upon the quadrant in which most of the ROC values lies. In this evaluation technique proposed method clearly outshined the state-of-art techniques. In future the proposed saliency may be used for multiple applications like image segmentation, image summarization, image registration etc.

ACKNOWLEDGEMENT

We would like to thank Punjab Technical University, Kapurthala for their kind support in this research work. This is because of the research facilities available at PTU and its affiliated college Chandigarh Engineering College, Landran that made this research work possible.

REFERENCES

[1] B. Wandell, Foundations of Vision, Sunderland, MA: Sin auer, 1995.
[2] Hou and L. Zhang, "Saliency Detection: A Spectral Residual Approach", IEEE Conference on Computer Vision and Recognition 2007, pp. 1-8, 2007.
[3] R. Achanta, S. Hemami, F. Estrada, and S. Sllsstrunk, "Frequency tuned Salient Region Detection", IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2009.
[4] C. Ngau, Li-Minn and K. Seng "Bottom up Visual Saliency Ma using Wavelet Transform Domain" 978-1-4244-5540-9/10/$26.00 ©2010 IEEE.
[5] O. Hershler and S. Hochstein, "At first sight: a high-level pop out effects for faces," in Vision Research, vol. 45, no. 13, pp.1707-1724, 2005.
[6] M. Cerf, J. Harel, W. Einhauser and C. Koch, "Predicting human gaze using low-level saliency combined with face detection," in Advances in neural information processing systems, vol. 20, pp.241-248, 2008.

[7] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, pp. 1254-1259, Nov. 1998.
[8] Y.-F. Ma and H.-J. Zhang, "Contrast-based Image Attention Analysis by Using Fuzzy Growing", Proceedings of the 11th ACM International Conference on Multimedia, pp. 374-381, 2003.
[9] S. Frintrop, M. Klodt, and E. Rome. A real-time visual attention system using integral images. International Conference on Computer Vision Systems, 2007
[10] C. Koch and S. Ullman. Shifts in selective visual attention: Towards the underlying neural circuitry. Human Neurobiology, 4(4):219–227, 1985
[11] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk, "Salient Region Detection and Segmentation", LNCS, Computer Vision Systems, Springer Berlin Heidelberg, pp. 66-75, 2008.
[12] W. H. Ngau, L.-M. Ang, K. P. Seng, "Salient Region Detection Using Contrast-Based Saliency and Watershed Segmentation", Proceedings of the 2nd International Conference on Computing & Informatics 2009 (lCOCI09), pp. 475-479, 2009.
[13] J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. Advances in Neural Information Processing Systems, 19:545–552, 2007.

