Region Contrast based Salient Object Detection

Sayali Doifode, Prof.M.B.Zalte

***Abstract*—Salient Object Detection methods are used to detect salient object from images. It solve segmentation problems. The method is based on Hyper Complex representation, in this phase spectrum of Fourier transform (PFT) model used. This method segment only the salient foreground object from the background. It find out which portion is object from entire region, based on Spatiotemporal Saliency map . We show the performances of different methods on the basis of ROI values. The experiment has been conducted on various test images and compared the performance in terms of ROI values.**

***Index Terms*— Saliency Detection, Hyper Complex Form Algorithm, Phase Spectrum of Fourier transform (PFT), Background Subtraction, Region of Interest.**

I. INTRODUCTION

One of the main tasks of image processing is to segment only the salient foreground object from the background. Humans pay more attention to some part of objects in images. They are experts at quickly and accurately identifying the most visually noticeable foreground object in the scene, known as salient objects [1]. While essentially solving a segmentation problem, salient object detection approaches segment only the salient foreground object from the background, rather than partition an image into regions [2].Background separation from the images is the one of the challenging task in the machine application. Saliency detection plays important roles in many image processing applications, such as regions of interest extraction and image resizing, content-aware image editing [4],[5] object recognition In this paper, hyper complex from algorithm is used to represent salient object which is composed of intensity, color and motion features. Existing salience detection algorithms [3] are built in the amplitude spectrum domain. Propose the method is based on phase spectrum of Fourier transform (PFT) model.

Most traditional object detection method is based on human's visual attention . In 1890, James [3] suggested the saliency model based on visual attention. A set of basic visual features, such as color, motion and edges are calculated. And then ,saliency map[6] is computed to indicate the locations of salient areas.

In recent years , X. Hou and L. Zhang[8] proposed a spectral residual (SR) approach based on Fourier Transform. This approach does not rely on the parameters and can detect salient objects rapidly.

M. Z. Aziz and B. Mertsching [9] proposed the region-based approaches to calculate the feature maps for their saliency models. At first it calculate clustering and to reduce computational complexity, by these clusters compute the feature maps. However, their models need to set many parameters to obtain useful results, and still cannot work in real time (only a few frame per second).

To perform some applications, need to add motion feature to these models. however, the additional motion feature will increase the computational cost of the model. To develop motion into a saliency model without influencing its computational cost is a challenging task that motivates to generate saliency map.

The propose methods based on Hyper complex representation of an image yield very simple algorithms that are often more powerful and easy to work with than traditional methods of function estimation. It consists of quaternion representation of an image and then calculate spatiotemporal saliency detection map. In quaternion representation of images ,each pixel of the image consists of color, intensity and motion feature.

There are number of methods are available for salient object detection like spectral residual, region-based, graph-based etc. In this paper we deal a new method.

The paper is organized as follows: Section II deals with the introduce the quaternion representation of an image and calculate spatiotemporal saliency map of the image. Section III discusses the results and analysis of salient object detection techniques. This is followed by conclusions.

II. HYPER COMPLEX FORM

Hyper Complex Form Algorithm is used to represent an image which is composed of intensity, color, and motion features. Many image processing techniques that are dependent on Fourier transforms It use phase spectrum of Fourier transform algorithm.

A. Quaternion Representation of an Image

Pixels of color images have three components red, green and blue. And it can be represented in quaternion form. The input image captured at time t as F(t) , $t = 1,2, \ldots . T$ where T is the total frame number.

A pixel at image coordinates (n, m) in RGB image and it can be represented as

$$f(n, m) = r(n, m)i + g(n, m)j + b(n, m)k \quad (1)$$

where $r(n, m), g(n, m), b(n, m)$ is the red, green and blue components of pixels respectively.

B. Spatiotemporal Saliency Map

Sayali Doifode, Department of Electronics & Telecommunication, K. J. Somaiya College of Engineering, Mumbai, India, 9765326223.
Prof.M.B.Zalte, Department of Electronics & Telecommunication, K. J. Somaiya College of Engineering, Mumbai, India, 9821132516.

Ell and Sangwine [10] provide a way to transform a quaternion image f(n, m) to the frequency domain, It is used to calculate the spatiotemporal saliency map.
From the equation (1), quaternion representation of an image is calculated and each pixel is denoted as f(n, m), where (n, m) is represent each pixel of image.
Generalized quaternion Fourier transform representation as follows:

$$F[u,v] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{m-1} \sum_{n=0}^{n-1} e^{-u2\pi\left(\left(\frac{mv}{M}\right)+\left(\frac{nv}{N}\right)\right)} f(n,m) \qquad (2)$$

Where (n, m) and (u, v) are locations of each pixels in spatial and frequency domain, respectively. M , N are the image's height and width.
The inverse form of (2) is obtained by changing the sign of exponential and summing over $u$and v instead of m and n.
Spatiotemporal saliency map is computed as follow:

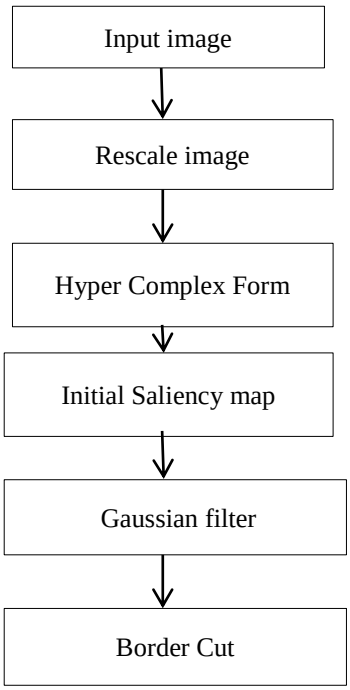
$$sM(t) = g * \sum_{i=0}^{3} \left(w_i p^2{}_i(t)\right) \qquad (3)$$

where g is a 2-D Gaussian filter with variance as σ.

Here Gaussian Filter is used to smoothen the image by reducing the time delay of the image frame. Gaussian filter have the properties to minimizing the size and fall time. It considers the ideal frequency domain filter.
In two dimensions, it is the product of Gaussians, one per direction:

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (4)$$

Where x and y are the distance from the origin in the horizontal and vertical axis respectively, σ is standard deviation of the Gaussian distribution.
Complete diagrammatical representation of work flow as follow:



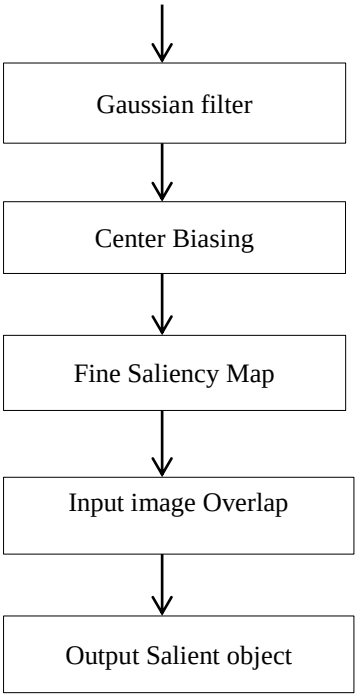
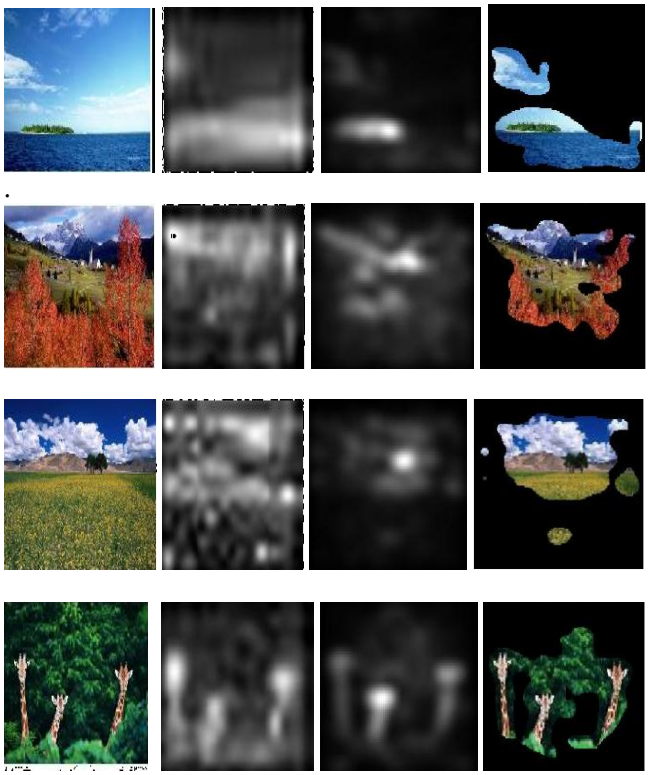


Fig. 1: Diagrammatical representation of work flow

III. RESULTS AND ANALYSIS

Fig. 2 show the experimental results of Spectral Residual method and Hyper Complex form method. We used same dataset THUS10000 for both methods and compare the result on the basis of ROI of image and time factor. Spectral Residual and Suggested Method for Salient Object Detection are applied to various test images and the results were obtained. Table I Results for performance of two different method.



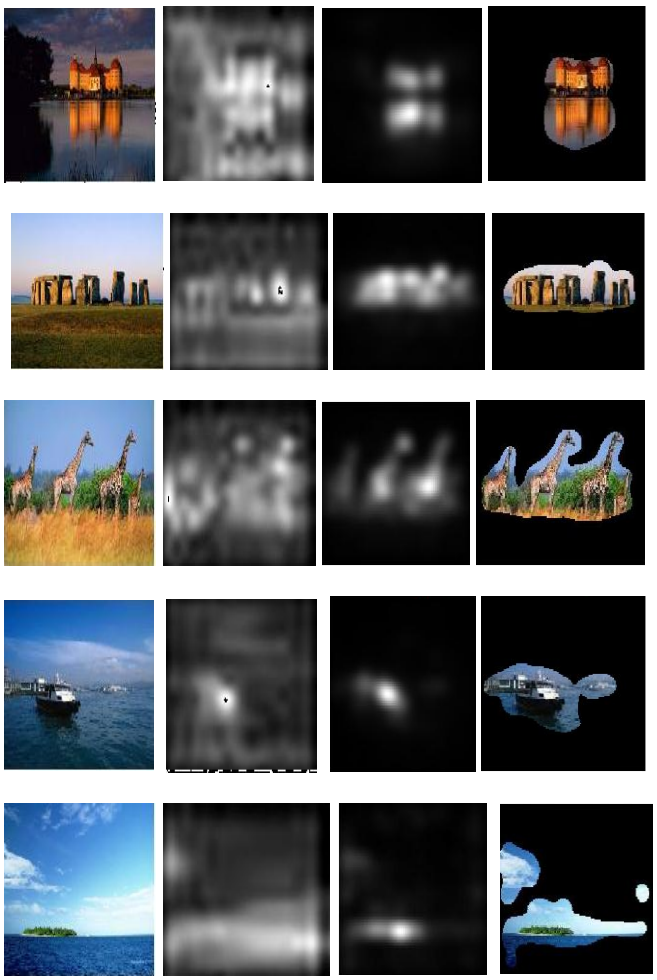


Fig. 2: The result of Spectral residual and propose method (1)Original image, (2) saliency map generated by spectral residual method ,(3)and Hyper complex form method,(4)Detected Salient object..

Table I: Performance of Two methods

Method	Spectral Residual method	Hyper Complex Form method
Time(s)	61.621	1.89
ROI	0.57	0.91

As maximum is the ROI value detected Salient Object is more accurate and minimum is the time method is fast to compute results. From Table I we observe that the ROI and time for Hyper Complex Form method is higher than Spectral Residual method.

IV. CONCLUSION

This paper presented a comparison between two Salient Object Detection methods based on the ROI and time factor. We observe that Suggested Method gives high quality saliency maps at the cost of reduced computational efficiency than Spectral Residual method. ROI factor is also higher for Suggested Method noise and the main content of the image are accurately detected.

REFERENCES

[1] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, T. X., and S. H.Y. ,"Learning to detect a salient object," IEEE TPAMI, no. 2, 2011
[2] G A. Borji, D. N. Sihite, and L. Itti, "Salient object detection: Abenchmark," in ECCV, 2012 .
[3] T. Liu, J. Sun, N. Zheng, X. Tang, and H. Y. Shum, "Learning to detect a salient object," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*,Jul. 2007.
[4] G.-X. Zhang, M.-M.Cheng, S.-M. Hu, and R. R. Martin,"A shape preserving approach to image resizing," Comput. Graph. Forum, vol. 28, no. 7, pp. 1897–1906, 2009.
[5] M.-M. Cheng, F.-L. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu,"Repfinder: Finding approximately repeated scene elements for image editing," ACM TOG, vol. 29, no. 4, pp. 83:1–8, 2010.
[6] [3] H. Zhou, H. Friedman, and R. von der Heydt, "Coding of Border Ownership in Monkey Visual Cortex," J. Neuroscience, vol. 20, no. 17 pp. 6594-6611, 2000.
[7] [6] C. Koch and S. Ullman, "Shifts in selection in visual attention: Toward the underlying neural circuitry," *Human Neurobiol.*, vol. 4, no. 4, pp.219–227, 1985. .
[8] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach,"presented at the CVPR, 2007..
[9] M. Z. Aziz and B. Mertsching, "Fast and robust generation of feature maps for region-based visual attention," *IEEE Trans. Image Process.*,vol. 17, no. 5, pp. 5, 633–644, May 2008.
[10] T. Ell and S. Sangwine, Hyper complex Fourier transforms of color images, IEEE Transactions on Image Processing 16 (2007), 22-35.

Sayali Doifode,
Post-Graduate Student
Electronics & Telecommunications Engineering,
K. J. Somaiya College of Engineering, Mumbai.

Prof.M.B.Zalte,
Professor
Electronics & Telecommunications Engineering,
K. J. Somaiya College of Engineering, Mumbai.