# SALIENT OBJECT DETECTION USING OCTONION WITH BAYESIAN INFERENCE

*Hong-Yun Gao and Kin-Man Lam*

Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hong Kong, China

## ABSTRACT

A novel computational model for detecting salient regions in color images is proposed, based on a two-stage coarse-to-fine framework. Firstly, different early visual feature maps – including the edge intensity; the black-white, red-green, and blue-yellow color opponents; and the Gabor features with four directions – are incorporated into the eight channels of an octonion image. Spectral normalization is achieved with the octonion Fourier transform by preserving the phase information of the octonion image. Then, with mean-shift segmentation, the saliency values in each segment are averaged to form a coarse saliency map. Finally, the coarse saliency map is subject to Bayesian inference to further refine the salient regions. The integration of frequency normalization, spatial segmentation and Bayesian inference exploits the benefits from both the spectral domain and the spatial domain. Experimental results show the superiority of the proposed method compared to several existing methods.

***Index Terms***— Salient object detection, octonion image, spectral normalization, mean-shift segmentation, Bayesian inference

## 1. INTRODUCTION

Visual attention is a remarkable mechanism of the human visual system (HVS) which can extract salient and important information from natural scenes efficiently. According to some previous research [1, 2], the visual-attention mechanism can be divided into two categories, namely bottom-up and top-down. The bottom-up mechanism is fast, simple, and task-independent, in which early features are processed. In contrast, the top-down mechanism involves high-level processing, which may require prior knowledge and training. In the existing bottom-up models, the saliency at a given pixel position is determined by computing the pixel's dissimilarity from its neighbors. "Dissimilarity" can be defined in many ways, e.g. center-surround contrast [3], self-information [4], etc. Many problems in computer vision, such as image compression [5], image categorization [6], and object localization [7], can benefit from saliency detection as a preprocessing step in order to focus on those areas of greater importance.

Over the past decades, many computational models have been proposed for computing visual saliency. Among these models, the most famous one was proposed by Itti *et al*. [3]. The algorithm is based on the center-surround mechanism – which is similar to the visual receptive field – in intensity, color and orientation. Bruce and Tsotsos [4] proposed a bottom-up model based on the principle of maximizing information sampled from a scene. Hou and Zhang [8] presented the first spectral approach for visual saliency detection, based on the difference between the perceived log-spectrum and the characteristic log-spectrum of natural images. Guo *et al*. [9] proposed that phase is the key in obtaining the saliency map, and introduced the use of quaternion image to detect salient regions by preserving phase information only. Achanta *et al*. [10] proposed using the difference of Gaussian (DoG) filtering to eliminate redundant information and generate saliency maps with well-defined object boundary. More recently, Goferman *et al*. [11] combined local low-level cues, global considerations, visual organization rules, and high-level features to highlight salient objects with their contexts. Gao and Lam [12] proposed using octonion image to incorporate more feature channels in generating saliency maps.

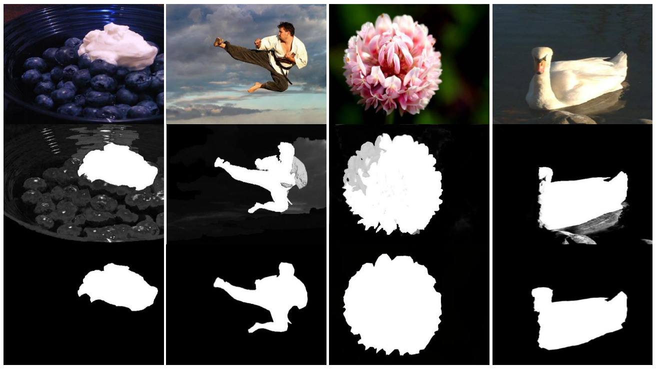

Figure 1 Saliency map. From top to bottom: the original sample images, our saliency map, and the ground-truth images.

The main contributions of our proposed framework are twofold. First, we propose a coarse saliency map method through octonion algebra and mean-shift segmentation, which is fast and can avoid exhaustive searching like the sliding-window method and multi-scale computation. Second, we combine spectral analysis with spatial processing, which can exploit the benefits of detecting saliency in both the spectral domain and the spatial domain. As Figure 1 indicates, our proposed method can detect the salient objects accurately.

The rest of the paper is organized as follows. Section 2 first introduces the concept of octonion algebra, octonion image construction, and octonion Fourier transform. Then, the generation of a coarse saliency map by our octonion saliency map and mean-shift segmentation is presented. In Section 3, we refine the coarse saliency map using Bayesian inference to produce the final saliency map. Experimental results and the evaluation of the proposed model are given in Section 4. Finally, the conclusion is provided in Section 5.

## 2. OCTONION FOURIER TRANSFORM AND MEAN-SHIFT SEGMENTATION FOR SALIENCY DETECTION

978-1-4799-5751-4/14/$31.00 ©2014 IEEE

Guo *et al*. [9] were the first to propose using quaternion algebra for saliency detection. A quaternion image is constructed based on four feature maps, including one intensity map, two color opponent maps, and a motion map. Due to the limited channel capacity, we cannot incorporate more feature maps into the quaternion image. Hence, in order to accommodate more feature maps, we propose extending the quaternion image into the octonion image, i.e. a more advanced algebra structure is employed.

### 2.1. Basic Octonion Properties

Among the normed algebra in mathematics, the octonions $O$ are the largest such group, with the other three being the real numbers $R$, the complex numbers $C$, and the quaternions $H$. Therefore, we propose to use octonion image as the container to accommodate more feature maps. Octonions, which were discovered by Graves and Caylay [13] independently, have eight dimensions, thus doubling the number of the quaternions, and can be defined as follows:

$$o = x_0e_0 + x_1e_1 + x_2e_2 + x_3e_3 + x_4e_4 + x_5e_5 + x_6e_6 + x_7e_7, \quad (1)$$

where $\{e_0, e_1, e_2, \mathrm{e}_3, e_4, e_5, e_6, e_7\}$ are unit octonions that can be treated as perpendicular axes in the eight-dimensional space. $\{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ are real numbers.

The addition and subtraction of octonions are done by adding and subtracting corresponding terms. However, multiplication is more complex, since it is neither commutative nor associative, i.e. $e_ie_j = -e_je_i \neq e_je_i$, if $i, j \neq 0$, and $(e_ie_j)e_k = -e_i(e_je_k) \neq e_i(e_je_k)$, if $i, j, k$ are distinct and non-zero. The multiplication rules is shown is Table 1.

| × | $\boldsymbol{e_0}$ | $\boldsymbol{e_1}$ | $\boldsymbol{e_2}$ | $\boldsymbol{e_3}$ | $\boldsymbol{e_4}$ | $\boldsymbol{e_5}$ | $\boldsymbol{e_6}$ | $\boldsymbol{e_7}$ |
|---|---|---|---|---|---|---|---|---|
| $\boldsymbol{e_0}$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
| $\boldsymbol{e_1}$ | $e_1$ | $-e_0$ | $e_3$ | $-e_2$ | $e_5$ | $-e_4$ | $-e_7$ | $e_6$ |
| $\boldsymbol{e_2}$ | $e_2$ | $-e_3$ | $-e_0$ | $e_1$ | $e_6$ | $e_7$ | $-e_4$ | $-e_5$ |
| $\boldsymbol{e_3}$ | $e_3$ | $e_2$ | $-e_1$ | $-e_0$ | $e_7$ | $-e_6$ | $e_5$ | $-e_4$ |
| $\boldsymbol{e_4}$ | $e_4$ | $-e_5$ | $-e_6$ | $-e_7$ | $-e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $\mathbf{e_5}$ | $e_5$ | $e_4$ | $-e_7$ | $e_6$ | $-e_1$ | $-e_0$ | $-e_3$ | $e_2$ |
| $\boldsymbol{e_6}$ | $e_6$ | $e_7$ | $e_4$ | $-e_5$ | $-e_2$ | $e_3$ | $-e_0$ | $-e_1$ |
| $\boldsymbol{e_7}$ | $e_7$ | $-e_6$ | $e_5$ | $e_4$ | $-e_3$ | $-e_2$ | $e_1$ | $-e_0$ |

Table 1. Multiplication rules for the unit octonions.

The conjugate of an octonion is defined as follows:

$$o^* = x_0e_0 - x_1e_1 - x_2e_2 - x_3e_3 - x_4e_4 - x_5e_5 - x_6e_6 - x_7e_7. \quad (2)$$

The product of an octonion with its conjugate is always a non-negative number as follows:

$$o^*o = x_0^2 + x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 + x_7^2. \quad (3)$$

Using the definition above, the norm is consistent with the standard Euclidean norm, and is defined as follows:

$$\|o\| = \sqrt{o^*o}\ . \quad (4)$$

### 2.2. Feature Map and Octonion Image Construction

In general, color images have 3 channels, such as *RGB*, *YUV* or *CIE Lab*. Instead of using the *RGB* color space, our algorithm transforms an image into the *CIE Lab* color space to form the feature maps. This transformation de-couples the luminance channel from the two color-carrying channels, i.e. the red-green (RG) and the blue-yellow (BY) color opponents, which is consistent with [14]. We retain the two color-opponent maps as the color feature maps. The luminance channel of the *YUV* color space is used as the intensity feature map or black-white (KW) color-opponent feature map. To better discriminate between different objects in an image, we use edge intensity to form another feature map. Besides intensity and color feature maps, Itti *et al*. [3] also incorporated the orientation feature maps in their framework, to approximate the receptive-field sensitivity profile of orientation-selective neurons in the primary visual cortex [15]. The orientation information is extracted by Gabor filters, which have a simplified form as follows:

$$g(x, y; \sigma, \theta, \omega, \varphi) = \exp(-\frac{x'^2 + y'^2}{2\sigma^2})\exp(i(\omega x' + \varphi)), \quad (5)$$

where $x' = x\cos\theta + y\sin\theta$ and $y' = -x\sin\theta + y\cos\theta$. In (5), $\sigma$ is the standard deviation of the Gaussian envelope, $\theta$ is the orientation selectivity of the Gabor filter, $\omega$ is the frequency of the sinusoidal factor, and $\varphi$ is the phase offset. For simplicity, we select only four orientations, namely the horizontal, vertical, and two diagonal orientations. Hence, the orientations for the Gabor filter are: $\theta \in \{0°, 45°, 90°, 135°\}$.

Once selected, these eight early feature maps can be integrated to form an octonion image. The constructed octonion image is given as follows:

$$o(x, y)=e(x, y)e_0 + kw(x, y)e_1 + rg(x, y)e_2 + by(x, y)e_3$$
$$+ o_0(x, y)e_4 + o_{90}(x, y)e_5 + o_{45}(x, y)e_6 + o_{135}(x, y)e_7, \quad (6)$$

where $e(x, y)$ is the edge map; $kw(x, y)$, $rg(x, y)$ and $by(x, y)$ are the black-white, red-green, and blue-yellow color maps, respectively; and $o_0(x, y)$, $o_{90}(x, y)$, $o_{45}(x, y)$, and $o_{135}(x, y)$ are the four orientation maps for 0°, 90°, 45°, 135°, respectively.

### 2.3. Octonion Fourier Transform and Spectral Normalization

Since there is no existing clearly defined Fourier transform for octonion images, we propose extending the existing quaternion Fourier transform on color images proposed by Ell and Sangwine [16], to implement the octonion Fourier transform. According to [16], the general quaternion Fourier transform is defined as follows:

$$F[u,v] = \frac{1}{\sqrt{MN}}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} e^{-\mu 2\pi((mv/M)+(nu/N))} f(n,m), \quad (7)$$

where $(n, m)$ and $(u, v)$ are locations in the spatial and the frequency domain, respectively; $M$ and $N$ are the height and width of the image $f(n, m)$, respectively; $\mu$ is any unit pure quaternion.

In order to implement the octonion Fourier transform with the quaternion Fourier transform, we can express an octonion image in the symplectic form as follows:

$$o(x, y) = q_1(x, y) + q_2(x, y)e_4, \quad (8)$$

where $q_1(x, y)$ and $q_2(x, y)$ are the simplex part and the perplex part of the octonion image $o(x, y)$ respectively, defined as follows:

$$q_1(x, y)=e(x, y)e_0+kw(x, y)e_1+rg(x, y)e_2+by(x, y)e_3, \quad (9)$$

$$q_2(x, y)=o_0(x, y)e_0+o_{90}(x, y)e_1+o_{45}(x, y)e_2+o_{135}(x, y)e_3. \quad (10)$$

Taking the quaternion Fourier transform of (8), we obtain:

$$O(u, v) = Q_1(u, v) + Q_2(u, v)e_4, \quad (11)$$

where $Q_1(u, v)$ and $Q_2(u, v)$ are the quaternion Fourier transform that can be calculated using (7).

As indicated in [16], a quaternion image can be expressed in polar form. Hence, $Q_i(u, v)$, where $i$ = 1, 2, can be expressed as follows:

$$Q'(u, v) = |Q'(u, v)|\, e^{\mu\, \varphi(u,v)} \tag{12}$$

where $\mu$ and $\varphi$ are referred to as the eigenaxis and eigenangle of the quaternion, respectively.

We follow the phase-preserving idea in [9] by neglecting the magnitude and keeping the phase information unchanged. By transforming back to the spatial domain, we can obtain the spectral-normalized quaternion image $q'(x, y)$ to form the saliency map as follows:

$$Q'(u, v) = e^{\mu\varphi(u,v)}, \tag{13}$$

$$q'(x,y) = \frac{1}{\sqrt{MN}} \sum_{v=0}^{M-1} \sum_{u=0}^{N-1} e^{-\mu_1 2\pi((mv/M)+(nu/N))} Q'(u,v), \tag{14}$$

where $Q'(u, v)$ is the normalized frequency representation.

The saliency map can be calculated by squaring the eight normalized channels and summing them together, as follows:

$$SM= A_0^2+A_1^2+A_2^2+A_3^2 +A_4^2+A_5^2+A_6^2+A_7^2, \tag{15}$$

where $SM$ is the saliency map; and $A_0$, $A_1$, $A_2$, $A_3$, $A_4$ $A_5$, $A_6$, and $A_7$ are the eight channels of the spectral normalized octonion image $q'(x, y)$, respectively.

### 2.4. Coarse Saliency Map Generation

Saliency detection performance can be improved by high-quality segmentation algorithms, e.g. mean-shift segmentation [17], since the segmentation process can provide accurate boundaries of the objects in an image. The coarse saliency map can be generated by averaging the saliency values in each segment. Assume that there are $N$ segments in an image after mean-shift segmentation, and each of the segments is denoted as $R_i$ ($i$ = 1, 2,…$N$). $N$ segmentation masks $M_i$ ($i$ = 1, 2, …, $N$) can thus be generated, with the values within $R_i$ being 1 and other regions being 0. The saliency value $s_i$ in the segment $R_i$ can be calculated as follows:

$$s_i = [\sum_x \sum_y M_i(x,y) S(x,y)] / |R_i|, \quad i=1, 2,\ldots N, \tag{16}$$

where $S(x, y)$ is the saliency map computed using (15), and $|R_i|$ is the number of pixels in the segment $R_i$. As (16) shows, the average saliency value in each segment is considered, thus producing clear object boundaries. After this operation, the false-alarm rate outside salient objects can be greatly reduced, since the saliency leakage is shared by all the pixels in that segment, which can be discarded after thresholding. In order to avoid small segments, we set a minimum segment size in the mean-shift segmentation algorithm.

## 3. SALIENCY REFINEMENT VIA BAYESIAN INFERENCE

In this section, we refine the coarse saliency map generated by our octonion saliency map and mean-shift segmentation based on Bayesian inference. In [18, 19], Bayes formula has been adopted in saliency detection with success. As indicated, the Bayesian inference can be formulated in saliency detection as:

$$p(sal \mid x) = \frac{p(sal)p(x \mid sal)}{p(sal)p(x \mid sal) + p(bg)p(x \mid bg)}, \tag{17}$$

where $p(sal|x)$ is the posterior saliency probability; $p(sal)$ is the prior saliency probability; and $p(x|sal)$ is the likelihood. We use the Bayes formula to represent the saliency of each pixel with the posterior probability. In the following two sub-sections, the procedures to derive the prior probability $p(sal)$, and the likelihood $p(x|sal)$ are demonstrated.

### 3.1. The Prior Probability

In the previous work [18] and [19], Rahtu *et al*. [18] simply use a constant for the prior probability; Xie and Lu [19] obtain the prior probability using complicated processes, including superpixel generation, $k$-means clustering, and exhaustive searching. As a compromise, we use the coarse saliency map obtained previously as the prior probability map, since the saliency measure of each segment after the segmentation reflects the importance of each segment.

With the coarse saliency map providing some prior saliency estimation, the prior probability $p(sal)$ can be modeled as the saliency value of each pixel scaled in the range of (0,1) in the coarse saliency map. $p(bg)$ is thus defined as 1 - $p(sal)$.

Compared with the constant prior probability in [18], our method is more accurate, since the spectral normalization and the mean-shift segmentation in the first stage can locate the general salient region and eliminate the background region to some degree. Compared with the complex processes used in [19], our method is comparable in terms of precision, but our processing time is less than 5% of [19]. This is because our algorithm is primarily based on FFT-like operations, while [19] requires lengthy iterations.

### 3.2. The Likelihood

The coarse saliency map can separate an image into two disjoint sets: the salient pixels $S$ and the background pixels $B$. The pixels in $S$ tend to be salient, while the pixels in $B$ tend to be non-salient. In order to threshold the coarse saliency map into $S$ and $B$, an adaptive threshold is empirically determined as the mean value of the coarse saliency map. In this scheme, the pixels which are larger than the threshold are classified as the background pixels $S$, and the pixels which are smaller than the threshold are classified as the salient pixels $B$.

We compute the likelihood $p(x|sal)$ and $p(x|bg)$ in a similar way as [19] by representing the feature at each pixel $x$ in the *CIE Lab* color space as $F(x) = (l(x), a(x), b(x))$. The color histogram is applied to describe the features in the $S$ and $B$ regions. The colors in each channel are uniformly quantized into 16 bins. By assuming the independence of the three color channels, the likelihood at a pixel $x$ can be computed as follows:

$$p(x \mid sal) = \prod_{f\in\{l,a,b\}} \frac{N_{S(f(x))}}{N_S}, \text{ and } p(x \mid bg) = \prod_{f\in\{l,a,b\}} \frac{N_{B(f(x))}}{N_B}, \tag{18}$$

where $N_S$ and $N_B$ are the number of pixels in $S$ and $B$; $N_{S(f(x))}$ and $N_{B(f(x))}$ represent the number of pixels in $S$ and $B$ that have the same feature value in one of the three feature channels specified by $f$. Having determined the prior probability and likelihood, the final saliency value can be calculated with equation (17).

## 4. EXPERIMENT RESULTS

In this section, we evaluate our algorithm on the MSRA dataset [19] with 5,000 color images. Besides, 1,000 ground-truth images from MSRA dataset are available in [10]. To evaluate our method, as mentioned in Section 1, we compare the proposed method with six existing well-accepted saliency models: the classical center-surround model IT [3]; the information maximization model AIM [4]; the spectral analysis model SR [8]; the difference of Gaussian filtering model FTS [10]; the image patch model CA [11]; and the Bayesian-based model RA [18].

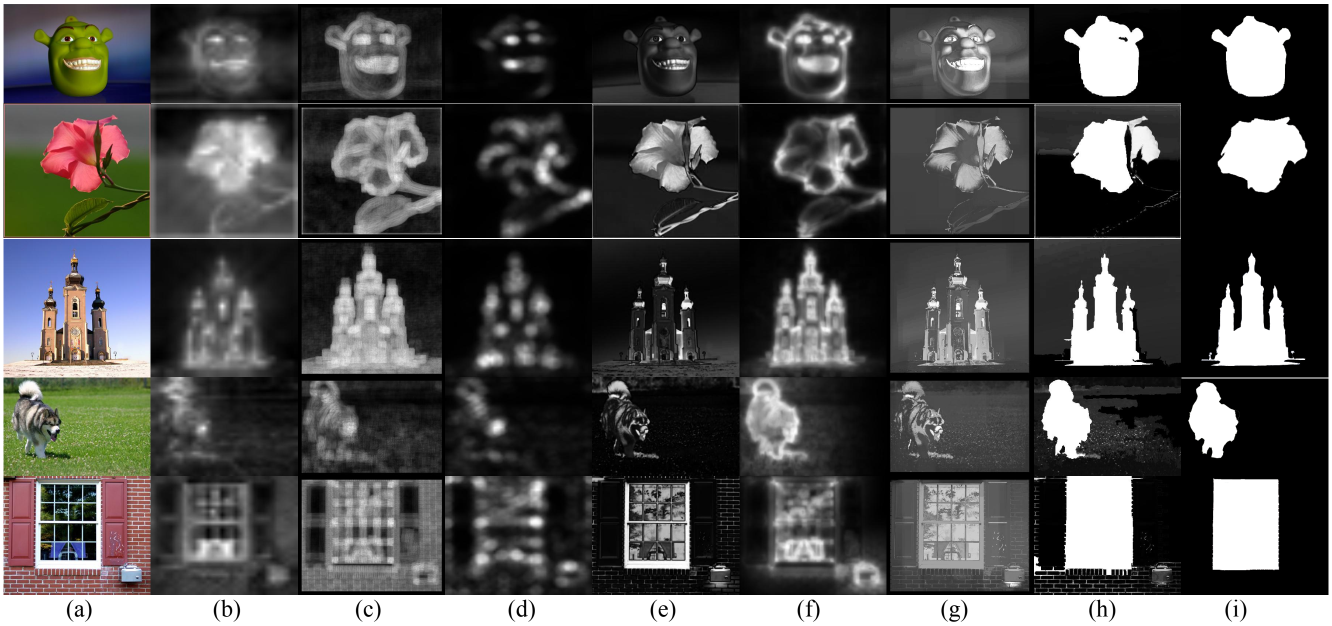

Figure 2. Comparison of our algorithm with six existing methods. From left to right: (a) original images, (b) IT [3], (c) AIM [4], (c) SR [8], (e) FTS [10], (f) CA [11], (g) RA [18], (h) our proposed algorithm, (i) ground-truth images

Figure 2 shows the saliency maps generated by the different methods. As presented, our method can detect salient objects accurately, with the boundaries being much better defined than with the other methods. We also evaluate the performance of our method quantitatively with the Receiver Operating Characteristic (ROC) curve [21]. The saliency map can be divided into two parts, namely the salient points and the non-salient points; and the binary ground-truth map can be divided into object points and background points. The percentage of the object points that fall into the salient points in the saliency map is the True-Positive Rate. The percentage of background points that fall into the salient points in the saliency map is the False-Positive Rate. The overall performance can be reflected by the area under the ROC curve (AUC), where a larger AUC score indicates a better performance.

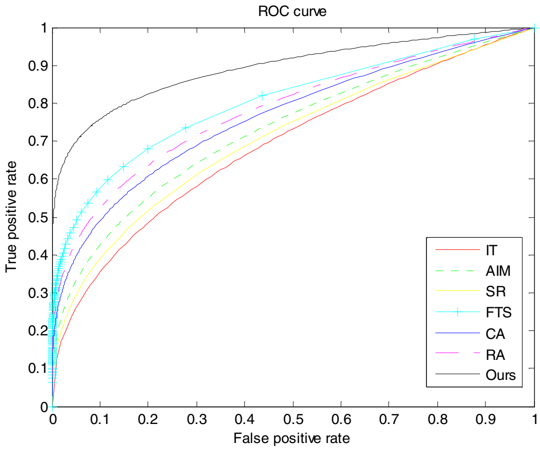

Figure 3. ROC curves for the different saliency detection methods.

The ROC curves of the different methods and our proposed method are shown in Figure 3, and the AUC scores of these methods are shown in Figure 4. From Figure 3 and Figure 4, it can be seen that our method has the largest AUC score, and thus achieves the best performance among the different methods.

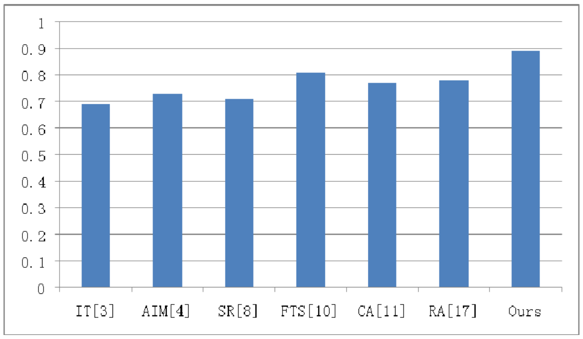

Figure 4. AUC scores for the different saliency detection methods.

## 5. CONCLUSION

In this paper, we have proposed a two-stage computational model to locate salient objects. In the first stage, an octonion image is constructed to incorporate eight feature maps, and is subject to the Fourier transform and spectral normalization. Then, mean-shift segmentation is applied to the spectral normalized octonion image to obtain accurate objects' boundaries. In the second stage, Bayesian inference further improves the detection accuracy. We have formulated the saliency value as the posterior probability in the Bayes formula, and it is determined by the prior probability and likelihood. Experimental results have shown that our method can detect salient objects more accurately than six existing salient object detection methods.

## 6. ACKNOWLEDGEMENT

This work is supported by an internal grant from the Hong Kong Polytechnic University (Grant No. G-YN21).

## 7. REFERENCES

[1] A. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology,* vol. 12, no. 1, pp. 97-136, 1980.

[2] C. Koch and S. Ullman, "Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry," *Human Neurobiology*, vol. 4, no. 4, pp. 219-227, 1985.

[3] L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 20, no. 11, pp. 1254-1259, 1998.

[4] N. Bruce and J. Tsotsos, "Saliency based on information maximization," in *Advances in Neural Information Processing Systems,* 2005.

[5] L. Itti, "Automatic foveation for video compression using a neurobiological model of visual attention," *IEEE Trans. on Image Processing*, vol. 13, no. 10, pp. 1304-1318, 2004.

[6] C. Siagian and L. Itti, "Rapid biologically-inspired scene classification using features shared with visual attention," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 29, no. 2, pp. 300-312, 2007.

[7] D. Gao, S. Han and N. Vasconcelos, "Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 31, no. 6, pp. 989-1005, 2009.

[8] X. Hou and L. Zhang, "Saliency Detection: A Spectral Residual Approach," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[9] C. Guo, L. Ma and L. Zhang, "Spatio-temporal Saliency Detection Using Phase Spectrum of Quaternion Fourier Transform," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.

[10] R. Achanta, S. Hemami, F. Estrada and S. Susstrunk, "Frequency-tuned salient region detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[11] S. Goferman, L. Zelnik-Manor and A. Tal, "Context-aware saliency detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 34, no. 10, pp. 1915-1926, 2012.

[12] H.-Y. Gao and K.-M. Lam, "From quaternion to octonion: feature-based image saliency detection," in *International Conference on Acoustics, Speech and Signal Processing*, 2014.

[13] Arthur Carlay, "On Jacobi's elliptic functions, in reply to the Rev. B. Bronwin; and on quaternions", *Philosophical Magazine*, vol. 26, pp. 208-211, 1845.

[14] S. Engel, X. Zhang, and B. Wandell, "Colour Tuning in Human Visual Cortex Measured With Functional Magnetic Resonance Imaging," *Nature*, vol. 388, no. 6637, pp. 68-71, 1997.

[15] A. Leventhal, The Neural Basis of Visual Function: Vision and Visual Dysfunction, vol.4, *CRC Press*, 1991.

[16] T. Ell and S. Sangwine, "Hypercomplex Fourier Transform of color images," *IEEE Trans. on Image Processing*, vol. 16, no. 1, pp. 22-35, 2007.

[17] D. Comaniciu and P. Meer, "Mean Shift: a robust approach toward feature space analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619, 2002.

[18] E. Rahtu, J. Kannala, M. Salo and J. Heikkila, "Segmenting salient objects from images and videos," in *European Conference on Computer Vision*, 2010.

[19] Y. Xie and H. Lu, "Visual saliency detection based on Bayesian model," *IEEE International Conference on Image Processing*, 2011.

[20] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang and H-Y. Shum, "Learning to Detect a Salient Object," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 353-367, 2011.

[21] B.W. Tatler, R.J. Baddeley and I.D. Gilchrist, "Visual Correlates of Fixation Selection: Effects of Scale and Time," *Vision Research*, vol. 45, no. 5, pp. 643-659, 2005.

978-1-4799-5751-4/14/$31.00 ©2014 IEEE