Copyright ⓒ 2018 KSII

Visual Saliency Detection Based on color Frequency Features under Bayesian framework

Naeem Ayoub[1], Zhenguo Gao* [1], Danjie Chen[2], Rachida Tobji[1] and Nianmin Yao[1]
[1] Department of Computer science and technology, Dalian University of technology
Dalian, China.
[2] College of software, Beijing institute of technology
Beijing, China.
[E-mail : gzg2012@dlut.edu.cn]
*Corresponding author : Zhenguo Gao

Received July 5, 2017; revised September 2, 2017; accepted September 25, 2017; published February 28, 2018

Abstract

Saliency detection in neurobiology is a vehement research during the last few years, several cognitive and interactive systems are designed to simulate saliency model (an attentional mechanism, which focuses on the worthiest part in the image). In this paper, a bottom up saliency detection model is proposed by taking into account the color and luminance frequency features of RGB, CIE L*a*b* color space of the image. We employ low-level features of image and apply band pass filter to estimate and highlight salient region. We compute the likelihood probability by applying Bayesian framework at pixels. Experiments on two publically available datasets (MSRA and SED2) show that our saliency model performs better as compared to the ten state of the art algorithms by achieving higher precision, better recall and F-Measure.

***Keywords*:** Saliency Detection, image processing, vision system, Bayesian Saliency, Color frequency, Log-Gabor filter

1. Introduction

Visual attention in computer vision is very ardent research over the previous few years, saliency detection is one of the main and important research topic in this area and has importance in most of computer vision systems. Recently, different visual attention models have been introduced by researchers, after the theoretical framework for saliency detection that was introduced in 1985 by Koch et al. [1]. The first persuasive and well known model was introduced by Itti et al. in 1998 [2]. Itti's model imitate the Feature Integration Theory [3], in which A. M. Treisman et al. proposed to generate separate low-level visual features maps by decomposing the visual scene. There are various beneficial use of saliency detection methods and are employed for several applications in computer vision, artificial intelligence systems and transmission of data. Some of the efficacious applications are facial feature detection [4], image retrieval [5] object detection [6], image summarization [7], image segmentation [8, 9], image and video compression [10] and attentional mechanism for interaction in robotic systems [11]. Some existing saliency detection approaches treat images pixel wise or scan windows, which is the same as the way visual system of human operates and some of the existing techniques use local, regional and global features of color images for saliency detection. Recently, in [12], Zhang et al. Proposed to use spatial information of color cues to estimate edges information and uses divisive normalization framework to suppress the surrounding responses and consider the color information as global to quantify the saliency information. In [13], Tie et al. uses the local, regional and global features of the image such as contrast, histogram, and color spatial distribution, for salient region detection. In [14], a Global contrast based saliency detection algorithm was proposed by M. Cheng, which evaluates the global contrast differences and spatial coherence. Achanta et al. in [15] proposed a model, which states that a full resolution saliency map can be obtained by first getting image's band pass filtered responses from CIE L*a*b* color space channels, then computing normalized center-surround difference maps for each features and combining them with a weighting scheme. Mostly, following method provides pleasing results and computationally very efficient. In [16], graph-based visual saliency (GBVS) model proposed by Harel et al. by introducing a novel graph-based normalization scheme. In **[17]**, Jian et al. introduced a new saliency detection scheme, in which an image decomposed into intensity and discriminant color channels and wavelet-based salient patch detector was employed to extract the visually informative. Then orientation feature map computed to estimate directional saliency information and combined these maps to estimate most salient object in the image. In [18], Jian et al. substantiated the image saliency using three different saliency factors of the image. Firstly, edges and outliers are detected using Quaternionic Distance Based Weber Descriptor. Secondly, center prior and than color prior cues are used to estimate the salient object. Saliency can be drawn in two ways bottom up and top down manner.

Bottom-up and Top-down Attention: attention to the salient object can be modeled in two ways bottom-up and top-down factors [19]. Bottom-up factors are those which attract the attention as a more discriminative object in the visual scene [20] and the responsible feature for this attraction must be sufficiently discriminative with respect to surrounding features, this attentional mechanism is also called exogenous (Exogenous orienting is considered to be reflexive and automatic and is caused by a sudden change in the periphery), automatic, reflexive, or peripherally cued [21]. Top-down attention derived by cognitive factors such as

knowledge, expectations and specific goals [22]. Other terms for top-down attention are endogenous (endogenous orienting occurs when attention is oriented according to the current goals or desires, allowing the focus of attention to be manipulated by the demands of a task) [23], voluntary [24], or centrally cued attention. There are many intuitive examples of this process. A taxi driver is more likely to see the petrol stations on road and cyclists notice cycle tracks. If you are looking for a blue highlighter on your book, blue regions will attract the gaze promptly than other regions.

In this paper, we propose a novel bottom up saliency detection method by taking into account color and luminance features of the image. We propose to use Bayesian framework for saliency detection based on center surround principle by taking into account the low-level visual features. At first, the behavior that the human visual system discerns salient objects in the image, can be modeled by applying band-pass filters [15]. We construct the prior map based frequency features of each color channel in CIE L*a*b* color space to get saliency information, by apply Log-Gabor band-pass filter [25] to acquire the information about edges of salient region. Secondly, the gradient information of grayscale and color spaces are very beneficial to extract the saliency information from the image. In [26] Weijer et al. proposed a method for saliency detection, in which color information was computed based on frequency of the image and color saliency is boosted by transformation function. We use color-boosted saliency to detect salient points and compute convex hull based on the salient points to estimate salient region. We use the information of low-level visual feature obtained from inside and outside the convex hull to compute the observation likelihoods at every pixel and apply Bayesian interface to compute the final saliency map. Our proposed approach achieves the better saliency results, having low computational cost than other saliency models. The performance of saliency maps derived from our method and other saliency detection methods is examined on the benchmark MSRA, and SED2 images datasets, our saliency model achieves outstanding results.

2. Formulation and Related Work

In this section, we present our saliency detection Algorithm and other related previous frameworks essential to calculate proposed saliency map. We adopt two methods to get saliency information from an image according to the frequencies of color channels. First, we extract the salient region from RGB image by taking into account the color features of image. Weijer et al. [26] proposed a model in which the gradient information of both grayscale and color space channels are used to estimate salient regions. The frequency based color information of the image derivative is computed and transformation function is developed to boost up the color saliency level. The color Harris detector [27] is operated to detect feature points based on the frequency of color features on image. As Compared with the other feature detectors on intensity level, the boosted color saliency points are more stable and informative [26]. In this paper, for the purpose to detect contour points of salient regions in a color image, we use color boosted Harris point operator [26]. These salient points furnish useful spatial information according to the object of interest in the image, by enclosing the salient points with convex hull, we get the salient region. As the selected interest points usually surround the salient object, this method provides a useful information for estimating the salient region. Secondly, converting the RGB image to CIE L*a*b* and apply band-pass filters, Band-pass filtering can be used to enhance edges (suppressing low frequencies). Achanta et al. in [15] indicates that the saliency detection mechanism can be modeled by integrating band-pass filter responses to the color space channels. Achanta et al. adopted the Difference of Gaussian (DoG)

filter to highlight the salient region. Achanta's work gives the initiative to use band-pass filtering method to get saliency maps according to salient region of the image. We apply Log-Gabor filter [25] instead of DoG filter because the Log-Gabor is useful for finding edges and blobs. Gabor filter can be constructed with arbitrary bandwidth and the bandwidth can be optimized to produce a filter with minimal spatial measure, having no DC component. Furthermore, the transfer function of the Log-Gabor filter can approximate the salient points with high frequency, which makes it more adept for encoding the images than other band-pass filters [25, 28].

2.1 Log-Gabor Filter

The transfer function of Log-Gabor filter g(x) (x=(x, y) $\in R^2$ in the frequency domain can be formulated as

$$G(\mathbf{r}) = \exp\left(-\left(\log \frac{\| \mathbf{I} \|_2}{\lambda_f}\right)^2 / 2\sigma_f^2\right) \tag{1}$$

Where $I = (u_1, u_2) \in R^2$ is the coordinate under the frequency domain of image, $\lambda_f = 3$ is the center frequency and σ_f^2 (σ_f = 0.01) controls bandwidth of filter. As due to the singularity in the log function at the origin g($\boldsymbol{x}$) is unable to express analytically. By applying inverse Fourier transform to G($\mathbf{r}$), we obtain approximate value of g(x). By the given image in RGB domain {R(x): $\mid$ $x \in \Omega$, where Ω denotes the image in spatial domain}, where R(x) represents a vector that contains color channels values, representing R (red), G (green), and B (Blue) intensities at the position x, its saliency map $\boldsymbol{S_R}(x)$ can be modeled by convolution of two prior maps obtained by band-pass filtering, which is similar to [15]. At first, R(x) should be converted from RGB to CIE L*a*b* space, these three color channels are denoted by and$f_L(x)$, $f_a(x)$ and $f_b(x)$. Then, the saliency by band pass filtering to each color space can be formulated as.

$$S_{FC} = \prod_{c \in \{L,a,b\}} \sqrt{(f_{\|c\|_2} * LG)} \tag{2}$$

In addition, saliency for each color channel to its mean difference by band-pass filtering can be formulated as

$$S_{FC} = \prod_{c \in \{L,a,b\}} \sqrt{(f_{(c - \mu c)} * LG)} \tag{3}$$

Where $*$ is convolution operator, $c \in \{L, a, b\}$ denotes color channels in CIE L*a*b* color space and μ is the mean, hence μc = mean (c$\in$ {L, a, b}). LG is Log-Gabor filter. $\| c \|_2$ is L_2 norm. The prior saliency map $S_R(x)$ can be calculated as

$$S_R(x) = S_{FCM} \cdot S_{FC} \tag{4}$$

Where S_{FCM} and S_{FC} are highlighted salient region maps obtained by convolution with band pass filter and $S_R(x)$ is prior saliency map.

Finally, Bayesian visual saliency map for the image can be calculated by using center-surround principle. In [29], Rahtu et al. presented a model in which saliency is measured with contrast between scanned windows and surrounding region within Bayesian inference, in which the saliency of the image is computed by applying a sliding windows to the image and comparing pixels inside and outside a kernel. Xie et al. proposed a well-modeled saliency detection algorithm by using low and mid-level visual cues of an image under the Bayesian framework [30]. This algorithm performed well by defining the pixel inside and outside the salient region in the image. Zhang et al. in [31] also proposed a saliency model SUN, constructed under the Bayesian framework using natural statistics. Bayesian formula can define well the saliency of a pixel or region [29, 30, 31]. We use the same way as [30] to define the salient region and use a Bayesian framework to conduct our final saliency map.

2.2 Bayesian framework for saliency

Let ω_1 is a pixel to be salient and ω_0 is the pixel related to the background. Let W denote the whole region, W_1 is the foreground (salient region) and W_0 is background area that is outside the convex hull. The probability of a pixel x to be salient can be defined as

$$S_{\omega}(x) = \frac{P\left(x/W_1\right) . P(W_1)}{(P(x))} \tag{5}$$

Where

$$P(x) = P\left(x/W_1\right) \cdot P(W_1) + P(W_0) \cdot P\left(x/W_0\right) \tag{6}$$

And

$$P(W_0) = 1 - P(W_1) \tag{7}$$

The convex hull segments the image into two disjoint regions, the region inside the convex hull and the region, which should be treated as background. We estimate saliency values from both convex hull and prior saliency map obtained by eq. (4) to estimate region of interest or salient object. Thus, likelihood probability of salient region can be calculated as

$$P(W_1) = \frac{|\ S_R(x) \cap Convex\ Hull\ |}{|\ S_R(x)\ |} \tag{8}$$

Where $S_R(x)$ denotes Prior map, Convex Hull is the region of salient points, |.| represents the set of total numbers of elements.

It can be assume that all the salient points assort in the region W_1, the pixels in W_1 tend to be salient while those in W_0 are more likely to be not salient. This method is successfully utilized in [29] and [30]. The saliency of every pixel is approximated on the level of similarity to pixels of W_1, likewise its difference from those of W_0 by using color histograms. The CIE L*a*b* color space is employed as it preserves luminance as well as color information of the image. In CIE L*a*b* color channels each pixel x is represented by [L*(x), a(x), b(x)*], we calculate color histograms of both regions W_1 and W_0. Let NW_1 be the number of pixels in W_1, and $NW_{1(c(x))}$ be the count that region W_1 contains $c(x)$. In the same way, we calculate color histograms of pixels in the region W_0 with NW_0, and $NW_{0(c(x))}$. We assume the CIE L*a*b* color channels as independent to simplify computations. The likelihood probability at each pixel x is computed in the same way as [30]

$$\omega_1 = \prod_{c \in \{L,a,b\}} \frac{NW_{1(c(x))}}{NW_1} \tag{9}$$

$$\omega_0 = \prod_{c \in \{L,a,b\}} \frac{NW_{0(c(x))}}{NW_0} \tag{10}$$

Where $\omega_1 = P\left({}^{x}/_{W_1}\right)$ is probability of pixel being salient and $\omega_0 = P\left({}^{x}/_{W_0}\right)$ is probability of pixel in the background. By eq. (6, 7, 8, 9 and 10), the eq. (5) can be rewritten as

$$S_\omega(x) = \frac{\omega_1 \cdot P(W_1)}{\omega_1 \cdot P(W_1) + \omega_0 \cdot (1-(P(W_1))} \tag{11}$$

Fig. 1 shows a comparison between saliency maps obtain from the methods constructed under Bayesian framework and our proposed method.

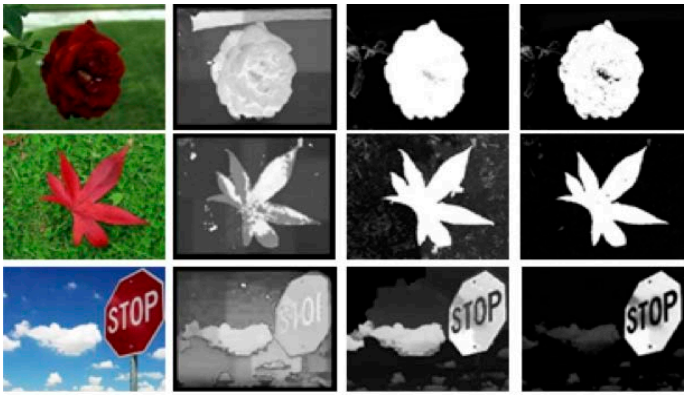


Fig. 1. From Left side to right first column shows RGB images, second, third and fourth columns show the saliency maps by SIV [29], BLM [30] and our proposed method, respectively.

3. Proposed Algorithm

Fig. 8 shows Illustration of our proposed method.
Our proposed Algorithm based on the following steps.
Step 1: Input Image (RGB Image).

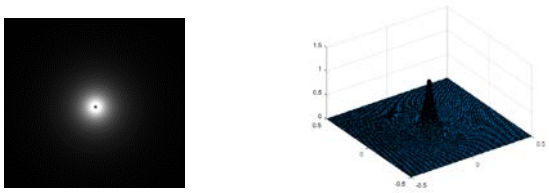
Step 2: Detect salient region by exploiting low-level cues by defining the region of interest for saliency values to be extracted, by employing color features detection method. We use color boosted Harris point detector to define salient region [26].

Step 3: Enclose the salient region by using convex hull to get pixel's information inside and outside the convex hull. **Fig.2-(b)** shows the salient points and **Fig.2-(c)** shows the convex hull enclosing the salient region.

(a) RGB Image **(b) Salient Points** **(c) Convex Hull**

Fig. 2. From left to right (a) RGB Original Image, (b) Salient Points and (c) Convex Hull

Step 4: Convert the image to CIE L*a*b* space and uniformly highlight the salient object by integrating band-pass filter to color channels, as band-pass filtering method has been successfully adopted in [2, 15, 16, 27] to highlight edges and boundaries of the salient object in the image. We use Log-Gabor filter to highlight the salient object and getting saliency values by using eq. (2) and eq. (3). **Fig. 3** shows the visualization of 2D Log-Gabor filter. **Fig. 4-(a)** and **Fig.4-(b)** show the highlighted salient region.



(a) **(b)**

Fig. 3. Visualization of the log-Gabor filter in the frequency domain, with $\lambda_f = 3$ and $\sigma_f = 0.01$ shown in 2D image format (a) and 3D surface format (b), respectively.

Step 5: By fusion of two saliency maps we got a prior saliency map, which highlights the salient region in the image by eq. (4), **Fig.4-(c)** shows the prior saliency map.

(a) S_{FC} (b) S_{FCM} (c) $S_R(x)$

Fig. 4. From left to right Highlighted salient region (a) S_{FC} by eq. (2), (b) S_{FCM} by eq.(3) and (c) Prior Saliency map

Step 6: we address the pixels inside and outside the salient region with respect to the frequency of the salient object highlighted by prior Map and convex hull. We extract useful information about saliency of a pixel with respect to highlighted salient region of prior map that the pixel is salient or related to background.

Step 7: Compute the final saliency by applying Bayesian framework at each pixel by eq. (11).

Step 8: Final Saliency Map

Fig. 5. Final Saliency map

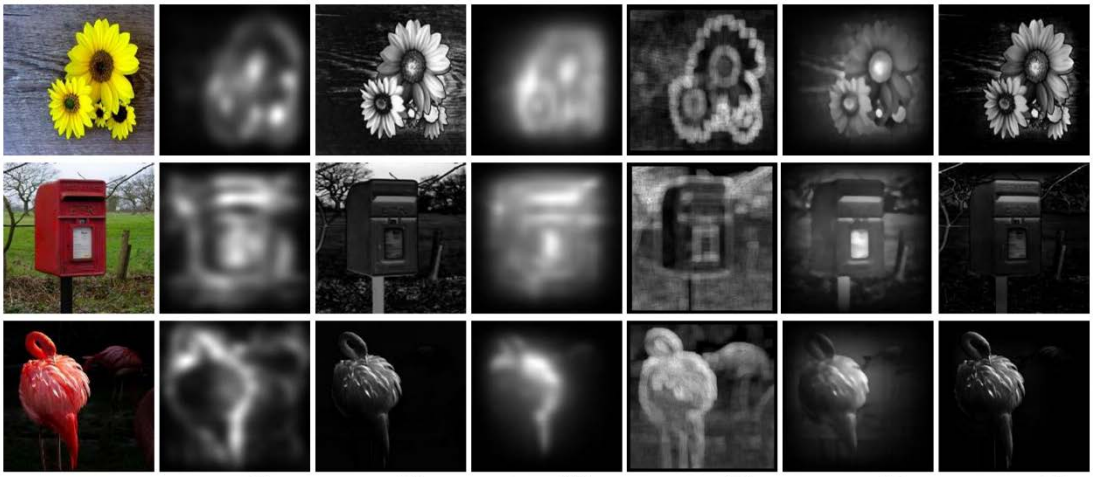


Fig. 6. Saliency maps obtained from different methods from left to right (a) original image, (b) IT [2], (c) FT [15], (d) GB [16], (e) AIM [18], (f) RCS [32], (g) MSS [33] respectively.

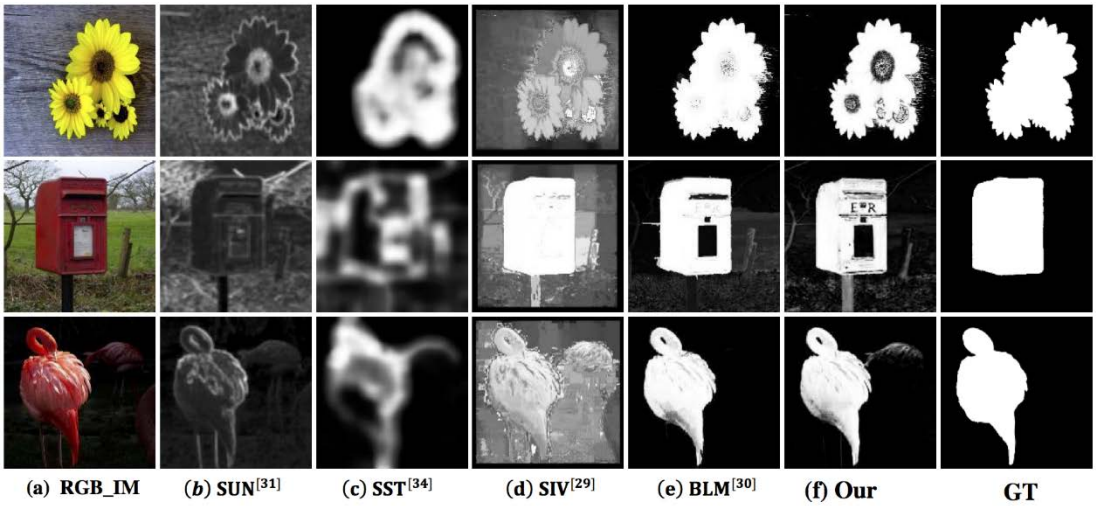


Fig. 7. Saliency maps obtained from different methods and our saliency method, from left to right (a) original image, (b) SUN [31], (c) SST [34], (d) SIV [29], (e) BLM [30], (f) Our and GT (ground truth) respectively.

4. Experimental Classification Results and Analysis

4.1. Dataset and parameter settings

We compared our proposed approach with other ten different saliency detection methods on MSRA dataset in which the salient objects of the images are segmented manually for the ground truth data with pixel-wise accuracy [15] and second is the SED2 dataset [37], which contains a set of 100 images of two objects and have accurate human-labelled ground truth. The methods used for comparison are IT [2], FT [15], GB [16], AIM [18], SIV [29], BLM [30], SUN [31], RCS [32]. MSS [33] and SST [34].

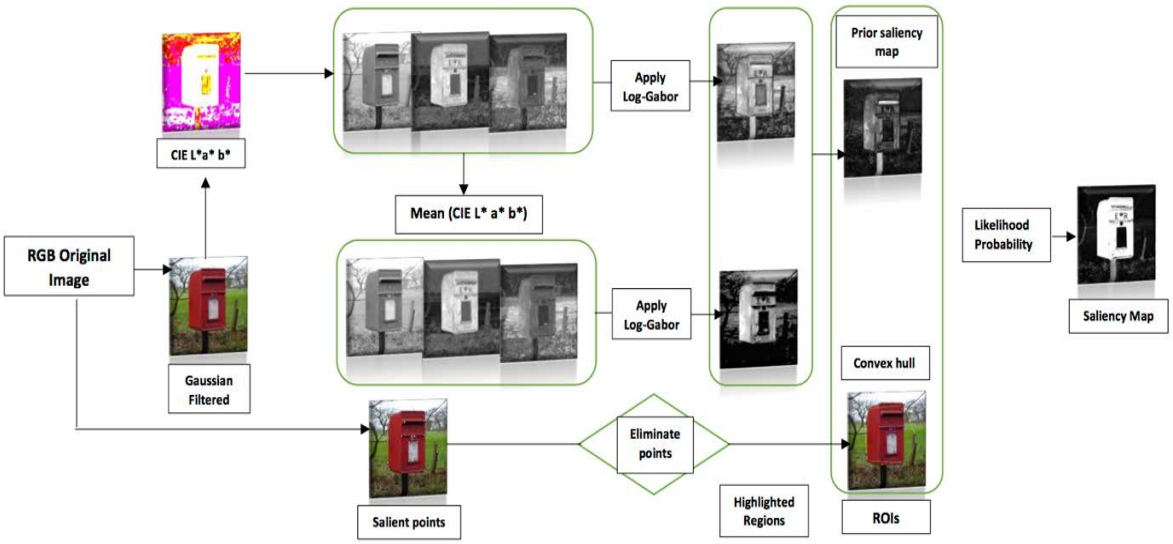


Fig. 8. Illustration of our saliency detection method

4.2 Graphical Representation:

4.2.1 Precision Recall Curves: We evaluate the performance in contrast to the task of best estimating the salient regions, we employ the method recommended and used in [15, 30, 32, 33, 35], where the saliency maps are binarized and compared with the ground truth fixation masks. The saliency maps are thresholded within domain of [0, 255] to obtain a binary mask and is compared with the ground truth masks. The thresholds are varied from 0 to 255 and recall-precision metrics are computed at each binarizing threshold. The recall-precision metric is deemed more suitable to evaluate the salient region detection performance [15, 32]. The recall and precision metrics are computed as in [32].

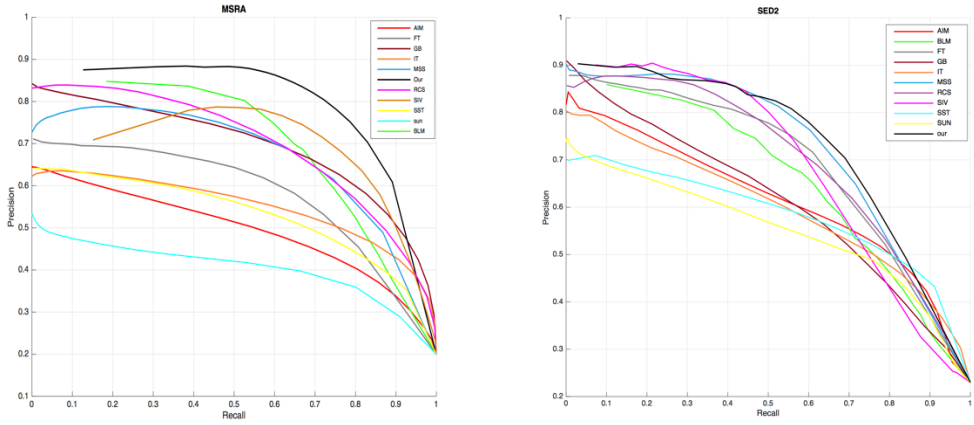
$$Precision = \frac{t_p}{sm} \tag{11}$$

t_p is the number of the pixels inside the salient regions of the saliency map and ground truth, sm is the number of the pixels inside the salient region of saliency map.

$$Recall \quad = \frac{t_p}{gt} \tag{12}$$

gt is the number of the pixels inside the salient regions of the ground truth map.

Fig. 9-(a) shows the Precision Recall results on MSRA dataset, and **Fig. 9-(b)** shows Precision Recall graph on SED2 Dataset, by our method against ten other saliency detection methods.



(a) MSRA (b) SED2

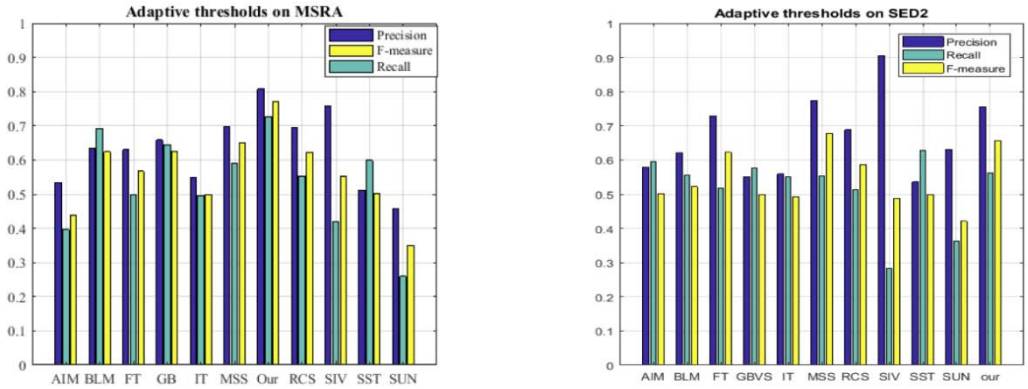
Fig. 9. Precision-recall curves, our method perform well and shows higher precision value against 10 state of art algorithms.

4.2.2 F-Measure: The F-Measure is defined as

$$F_\beta = \frac{(1+\beta^2) \cdot Precision \cdot Recall}{(\beta^2 \cdot Precision + Recall)} \tag{14}$$

We set β^2 to 0.3 to weigh precision more than recall as this evaluation method is suggested and employed in [15, 30]. The F-measure is the overall performance measurement computed by the weighted harmonic of precision and recall. Among all the methods, our algorithm achieves the best accuracy with higher Precision, recall and best F-Measure values on the large (MSRA) dataset of 1000 images.

Fig. 10-(a) and **Fig. 10-(b)** shows the F-Measure results on MSRA dataset and SED2 dataset by our method against 10 other saliency detection methods.



(a) MSRA (b) SED2

Fig. 10. from left to right F-Measure results on MSRA dataset and SED2, respectively.

Table 1 shows the quantitative results of precision, recall and f-measure values.

Table 1. Precision, Recall, and F-measure performance of saliency maps from various methods and our proposed algrithm.

Saliency map	MSRA			SED2		
	Precision	**Recall**	**F-measure**	**Precision**	**Recall**	**F-measure**
IT [2]	0.547	0.496	0.499	0.559	0.550	0.492
FT [15]	0.629	0.498	0.566	0.729	0.517	0.622
GB [16]	0.658	0.644	0.623	0.549	**0.577**	0.498
AIM[18]	0.534	0.396	0.437	0.578	**0.593**	0.502
SIV [29]	0.758	0.420	0.552	**0.904**	0.283	0.486
BLM [30]	0.635	0.691	0.623	0.621	0.555	0.521
SUN [31]	0.455	0.259	0.350	0.630	0.363	0.420
RCS [32]	0.695	0.553	0.622	0.688	0.513	0.586
MSS [33]	0.698	0.591	0.649	**0.773**	0.553	**0.677**
SST [34]	0.512	0.599	0.501	0.535	**0.629**	0.498
Proposed	**0.807**	**0.726**	**0.771**	**0.755**	**0.561**	**0.656**

From the above **Table 1**, it can be clearly seen that our method performs well as compared to the ten state of the art algorithms and achieves higher precision, recall and f-measure values on MSRA dataset. On SED2 dataset, SIV [29] and MSS [33] have better precision values but our saliency map achieves better recall and f-measure values as compared to SIV. GB [16], AIM [18] and SST [34] achieves better recall values but our method has advantage on them by achieving better Precision and recall values.

4.3 Computational cost

We also evaluate the performance of our saliency Algorithm in contrast to the computational cost to the saliency prediction accuracy with reference to the other methods in consideration. By taking into consideration the systems' different specifications, all experiments of different saliency methods were executed on standard HP Pavilion dv6 workstation with a 2.6GHZ Intel core i3 having 4GB of RAM, With the software MATLAB R2015b. The time costs to detect the salient object by each method, for processing the three images of different sizes are listed in the **Table 2** below.

Itti's method IT [2] calculate saliency map in less time (**Table 2**), However, our method has advantage over itti's model, it calculates full resolution saliency map and performs better in segmentation process, having better precision and recall values can be clearly seen in **Table 1**.

Table 2. The computational run-time in seconds of different saliency methods under consideration.

Saliency Methods	Code Type	Run time (in S) w.r.t image Size		
		(300 x 400)	(397 x 400)	(400 x 299)
IT [2]	MATLAB	**0.3421**	**0.3987**	**0.3422**
FT [15]	MATLAB	0.5853	0.8516	0.8512
GB [16]	MATLAB	1.9834	1.1802	0.7027
AIM[18]	MATLAB	11.9285	11.0794	15.3795
SIV [29]	MATLAB	15.5145	35.5805	25.1987
BLM [30]	MATLAB	199.4880	223.7677	214.2082
SUN [31]	MATLAB	17.4417	10.5108	7.6667
RCS [32]	MATLAB	1.2892	1.7511	0.7470
MSS [33]	MATLAB	1.7248	1.0499	1.3601
SST [34]	MATLAB	3.0981	3.1041	2.9967
Proposed	MATLAB	**0.5564**	**0.6000**	**0.4630**

4.4 Segmentation by adaptive thresholding

We also present the segmentation performance of the various approaches by considering sample images from the MSRA dataset [15], as shown in **Fig. 11** below. For the illustration purpose, we overlaid the original images with saliency maps after binarizing the respective saliency maps by Tsai's algorithm of thresholding [36]. Vikram et al. also adopted this method in literature [32].

It can be observed from **Fig. 11**, Bayesian method SUN [31] saliency approach fails to localize the salient object in the image effectively when segmentation performed using Tsai's algorithm. FT [15] also uses color features in the CIE L*a*b* space and uses band-pass filter but it is less effective in some cases, but unable to concentrate on the salient object in the image, while detecting the red box. GB [16] model is able to locate the salient objects, but generates low-resolution saliency map and highlights the boundaries and dispense low probabilities to pixels inside the objects. As IT [2] and GB [16] method calculate the saliency of each pixel by its color contrast to the average of the entire image, this method has low probability of finding the salient region when the background and foreground have similar color values and also when the salient region is large.

The SIV [29] model locates the salient objects more accurately than the IT [2] and GB [16] methods, but it also shows the background pixels as salient due to the use of scanning windows. The AIM [18] and MSS [33] algorithms compute saliency according to the color as well as spatial contrasts and it performs well in highlighting small salient objects but fails to focus on the salient object. RCS [32] and SST [34] also cannot locate the salient objects efficiently in some cases. Overcoming the drawbacks, our saliency algorithm is more efficient and remain stable while segmenting the saliency object and the computational cost is also very low (**Table 2**).

By **Fig. 11**, it can be clearly seen that our proposed algorithm can perform favorable against state of art of algorithms while segmenting the image and highlight the full resolution saliency map.

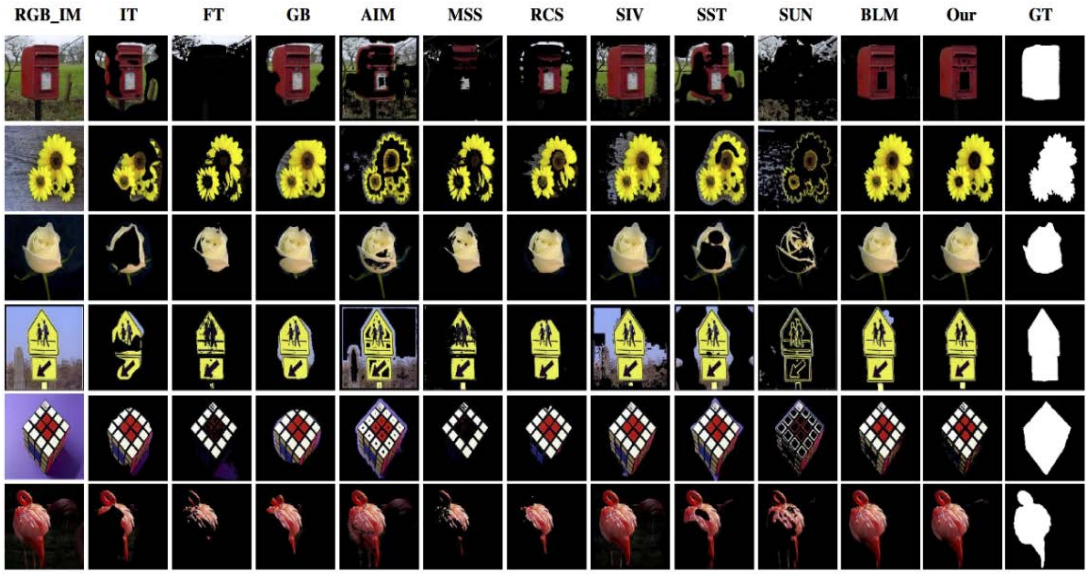


Fig.11. Segmentation results, from left to right first column shows Original images, IT [2], FT [15], GB [16], AIM [18], MSS [33], RCS [32], SIV [29], SST [34], SUN [31], BLM [30] and Our respectively.

4.5 Failure Cases

Our proposed saliency detection method performs well against state of the art algorithms by obtaining higher precision, recall and fmeasure values. We use only low level cues to get salient region information, by using band pass filter and also observation likelihood. In some cases, our method fails to detect the approximate salient region if the interest points do not appear around the saliency region due to similarity of color between foreground and background in the image. However, the other saliency detection methods also do not perform well in such cases.

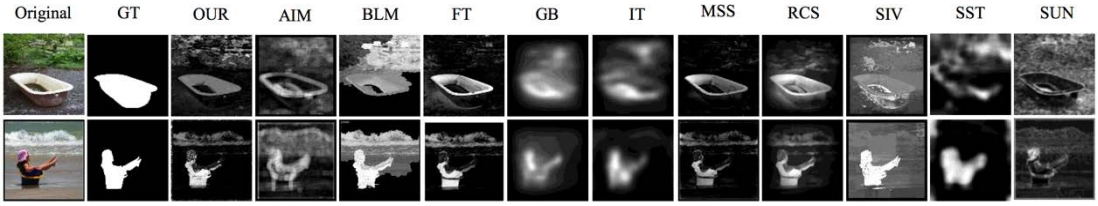


Fig.12. Failure cases of our method and other saliency algorithms, from left to right first column shows Original images, Ground Truth, Our, AIM [18], BLM [30], FT [15], GB [16], IT [2], MSS [33], RCS [32], SIV [29], SST [34], SUN [31], and respectively.

5. Discussion

Our method solely concentrates on low level visual cues of the image by taking into account the frequency features. **Fig. 9**, **Fig. 10** and **Table 1** show that our saliency maps achieve better results against 10 state of art algorithms. In contrast to computational time cost, our method also performs well having low computational cost shown in **Table 2**. As our method is built under the Bayesian framework like SIV [29], BLM [30] and SUN [31], we compare our results with other methods in contrast to Precision, recall and Fmeasure. Our method has advantages on them having better precision, recall and F-Measure values on MSRA dataset. SIV and MSS method, on SED2 dataset have larger precision value. These methods fail to detect full resolution salient map. Our method only uses low-level visual features of the images and SIV uses scanned windows, which makes the later most expensive (**Table 2**) and our saliency algorithm attains higher recall and f-measure values. We also analyze results in contrast to segmentation, our method segments the image well against ten state of art of algorithms. **Fig. 11** shows performance of our method against the other methods in contrast to the segmentation.

6. Conclusion

We propose a bottom-up saliency model within the Bayesian framework utilizing the low-level visual features of the image. Based on the informative saliency points, our method is constructed using color and luminance frequency features of the image. As our method get saliency values from both color spaces (RGB and CIE L*a*b*), it is fast and provides the full resolution saliency maps. We incorporate our saliency map within the Bayesian inference framework and evaluate the performance using ground truth evaluation, time cost and also in contrast to segmenting the salient object, on a MSRA and SED2 data sets with labeled ground truth. Experimental results show the effectiveness of our saliency model. Our algorithm generates more discriminative and full resolution saliency maps with higher precision, recall

and F-Measure values than ten algorithms. Furthermore, our algorithm can be easily incorporated with other applications as a vital initialization step such as image segmentation.

7. Acknowledgement

The work presented in this paper was supported by the National Natural Science Foundation of China under Grant no. 61671169.

8. References

[1] 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. Article (CrossRef Link)
[2] L. Itti, C. Koch and E., Niebur, "A model of saliency based visual attention for rapid scene analysis," *IEEE Transactions on Pattern analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254-1259, November, 1998. Article (CrossRef Link)
[3] A. M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology*, vol. 12, no. 1, pp. 97-136, January, 1980. Article (CrossRef Link)
[4] M. Jian, K. M. Lam and J. Dong, "Facial-feature detection and localization based on a hierarchical scheme," *Information Sciences*, vol. 262, pp. 1 – 14, March, 2014. Article (CrossRef Link)
[5] M. W. Jian, J. Y. Dong and J. Ma, "Image retrieval using wavelet-based salient regions," *The Imaging Science Journal*, vol. 59, no. 4, pp. 219-231, 2011. Article (CrossRef Link)
[6] P. Khuwuthyakorn, A Robles-Kelly and J. Zhou, "Object of interest detection by saliency learning," in *Proc. of Proceedings of the European Conference on Computer Vision*, Heraklion, Crete, Greece, pp. 636-649, September, 2010. Article (CrossRef Link)
[7] L. Shi, J. Wang, L. Xu, H. Lu and C. Xu, "Context saliency based image summarization," in *Proc. of Proceedings of the IEEE International Conference on Multimedia and Expo*, New York, USA, pp. 270-273, July, 2009. Article (CrossRef Link)
[8] M. Donoser, M. Urschler, M. Hirzer and H. Bischof, "Saliency driven total variation segmentation," in *Proc. of Proceedings of the IEEE International Conference on Computer Vision*, Kyoto, Japan, pp. 817-824, September, 2009. Article (CrossRef Link)
[9] R. Achanta, F. Estrada, P. Wils and S. Süsstrunk, "Salient region detection and segmentation," in *Proc. of Proceedings of the International Conference on Computer Vision Systems*, Santorini, Greece, pp. 66-75, May, 2008. Article (CrossRef Link)
[10] C. Guo, L. Zhang, "A Novel Multiresolution Spatiotemporal Saliency Detection Model and Applications in Image and Video Compression," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 185-198, January, 2010. Article (CrossRef Link)
[11] Y. Nagai, "From bottom-up visual attention to robot action learning," in *Proc. of Proceedings of the IEEE International Conference on Development and Learning*, Shanghai, China, pp. 1-6, June, 2009. Article (CrossRef Link)
[12] J. Zhang, , M. Wang, S. Zhang, X. Li and X. Wu, "Spatiochromatic Context Modeling for Color Saliency Analysis," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 6, pp. 1177-1189, June, 2016. Article (CrossRef Link)
[13] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang and X. Y. Shum, "Learning to detect a salient object," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 353-367, February, 2011. Article (CrossRef Link)
[14] M. Cheng, N.J. Mitra, X. Huang, P. H. S. Torr and S. H. Hu, "Global contrast based salient region detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569 – 582, March, 2015. Article (CrossRef Link)
[15] R. Achanta, S. Hemami, F. Estrada and S. Susstrunk, "Frequency tuned salient region detection," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Miami, USA, pp. 1597- 1604, June, 2009. Article (CrossRef Link)

[16] J. Harel, C. Koch and P. Perona, "Graph-based visual saliency," *Advances in Neural Information Processing Systems*, pp. 545- 552, December, 2007. Article (CrossRef Link)
[17] M. Jian, K. M. Lam, J. Dong and L. Shen, "Visual-Patch-Attention-Aware Saliency Detection." *IEEE Transactions on Cybernetics*, vol. 45, no. 8, pp. 1575 – 1586, August, 2015. Article (CrossRef Link)
[18] M. Jian, Q. Qi, J. Dong, S. Sun and K.M. Lam, "Saliency detection using quaternionic distance based weber descriptor and object cues," in *Proc. of Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Asia-Pacific, jeju, south Korea, December 13-16, 2016. Article (CrossRef Link)
[19] V. Setlur, T. Lechner, M. Nienhaus and B. Gooch, "Retargeting Images and Video for Preserving Information Saliency," *IEEE Computer Graphics and Applications*, vol. 27, no. 5, pp. 80 – 88, September, 2007. Article (CrossRef Link)
[20] M. Guttmann, L. Wolf and C. O. Danny, "Content aware video manipulation," *Computer Vision and Image Understanding*, vol. 115, no. 12, pp. 1662-1678, December, 2011. Article (CrossRef Link)
[21] B. C. Ko and J. Y. Nam, "Object-of-interest image segmentation based on human attention and semantic region clustering," *Journal of the Optical Society of America A*, vol. 23, no. 10, pp. 2462-2470, 2006. Article (CrossRef Link)
[22] J. Han, K.N. Ngan, M. Li and H. J. Zhang, "Unsupervised extraction of visual attention objects in color images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 1, pp. 141-145, January, 2006. Article (CrossRef Link)
[23] U. Rutishauser, D. Walther, C. Koch and P. Perona, "Is bottom-up attention useful for object recognition?" in *Proc. of Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)*, Washington, DC, USA, July, 2004. Article(CrossRef Link)
[24] J. S. Kim, J. H. Kim and C. S. Kim, "Adaptive image and video retargeting technique based on fourier analysis," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09)*, Miami, FL, USA, pp. 1730-1737, June, 2009. Article (CrossRef Link)
[25] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America. A*, vol. 4, no.12, pp. 2379-2394, December, 1987. Article (CrossRef Link)
[26] J. V. D. Weijer, T. Gevers and A. D. Bagdanov, "Boosting color saliency in image feature detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 150-156. January, 2006. Article (CrossRef Link)
[27] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. of Fourth Alvey Vision Conference*, pp. 147-151, 1988. Article (CrossRef Link)
[28] P. Kovesi, "Image features from phase congruency," *Videre: Journal of Computer Vision Research*, vol. 1, pp. 1-30, June, 1995. Article (CrossRef Link)
[29] E. Rahtu, J. Kannala, M. Salo and J. Heikkila, "Segmenting salient objects from images and videos," in *Proc. of Proceedings of European Conference on Computer Vision (ECCV'10)*, Crete, Greece, pp. 366-379, September, 2010. Article(CrossRef Link)
[30] Y. Xie, H. Lu and M. H. Yang, "Bayesian saliency via low and mid level cues," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1689 -1698, May, 2013. Article (CrossRef Link)
[31] L. Zhang, M. H. Tong, T. K. Marks, H. Shan and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *Journal of Vision*, vol. 8, no. 7, pp. 1-20, December, 2008. Article (CrossRef Link)
[32] T. N. Vikram, M. Tscherepanow and B. Wrede, "A saliency map based on sampling an image into random rectangular regions of interest," *Pattern Recognition*, vol. 45, no. 9, pp. 3114-3124, September, 2012. Article (CrossRef Link)
[33] R. Achanta and S. Susstrunk, "Saliency detection using maximum symmetric surround," in *Proc. of Proceedings of IEEE International Conference on Image Processing*, Hong Kong, pp. 2653-2656, September, 2010. Article (CrossRef Link)

[34] H. J. Seo, P. Milanfar, "Static and space-time visual saliency detection by self-resemblance," *Journal of Vision*, vol. 9, no.12, pp.1-27, November, 2009. Article (CrossRef Link)
[35] T. N. Vikram, M. Tscherepanow and B. Wrede, "A random center surround bottom up visual attention model useful for salient region detection," in *Proc. of Proceedings of the IEEE Workshop on Applications of Computer Vision*, Kona, HL, USA, pp. 166-173, February, 2011. Article (CrossRef Link)
[36] W. H. Tsai, "Moment-preserving thresholding: a new approach, Computer Vision," *Graphics and Image Processing*, vol. 29, no. 3, pp. 377-393, March, 1985. Article (CrossRef Link)
[37] R. B. S. Alpert, M. Galun, and A. Brandt, "Image segmentation by probabilistic bottom-up aggregation and cue integration," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, MN, USA, July, 2007. Article (CrossRef Link)

Naeem Ayoub received his master degree in information technology from university of education Lahore, Multan campus in 2013. Now he is pursuing his Ph.D. at computer science and technology at Dalian university of technology, Dalian, China. His research interests include computer vision, image processing, IOT and WSNs.

Zhenguo Gao is a Professor in Dalian University of Technology, Dalian, China. He has been a visiting scholar in University of Illinois at Urbana-Champaign and University of Michigan in 2010 and 2011. He received his BS and MS degree in Mechanical and Electrical Engineering from Harbin Institute of Technology, Harbin, China, in 1999 and 2001, respectively. Then he received his Ph.D. degree in Computer Architecture from Harbin Institute of Technology, Harbin, China, in 2006. He is a senior member of China Computer Federation. He received National Science Foundation Career Award of China in 2007 and Outstanding Junior Faculty Award of Harbin Engineering University in 2008. He is severing as a reviewer for project proposals to National science foundation of China, Ministry of Education of China. His research interests include wireless ad hoc network, cognitive radio network, network coding, image processing.

Danjie Chen is a postgraduate student, pursing her Master degree in College of Software, Beijing University of Technology, Beijing, China. Her main research interests include image processing, cooperative communication, and cognitive radio networks.

Rachida Tobji received her master degree university of technology, sidi bel Abbes, Algeria, in 2014. Now she is pursuing her Ph.D degree in computer science and technology at Dalian university of technology, Dalian, China. Her research Interests include Image processing, Iris detection.

Nianmin Yao is a Professor in Dalian University of Technology, Dalian, China. He has been a visiting scholar in University of Connecticut in 2010. He is a senior member of China Computer Federation. His primary research interests include image processing, wireless sensor networks and network storage.