

# Exploiting Randomized Prim's Algorithm and Background Contrast for Saliency Detection

Lv Jianyong<sup>\*</sup>, Tang Zhenmin and Xu Wei

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu, 210094, P.R. China

**Abstract:** The geodesic saliency method in the literature was based on the boundary and connectivity priority, which assumed that most of the background regions can touch the image boundaries. It cannot deal with the images with complex backgrounds or variant textures. To address such problem, we propose an improved saliency detection method by involving the important foreground priority. First, the statistical results of randomized Prim's algorithm are used to generate a coarse conspicuity map, which aims to roughly estimate the potential foreground. Then, the image is over-segmented into some individual superpixels and an affinity propagation clustering method is used to group the superpixels having a similar color appearance together. This is followed by the foreground probability map computation through the spatial interaction information between the coarse conspicuity map and superpixel based color clusters. The final saliency map is generated by integrating the above foreground probability map and background color contrast in a unified way. The quantitative and qualitative comparisons on the benchmark dataset MSRA-1000 and SED show that our method outperforms many recent proposed state-of-the-art approaches significantly.

**Keywords:** Background contrast, coarse conspicuity map, foreground probability map, randomized prim's algorithm, saliency detection.

## 1. INTRODUCTION

The selective visual attention mechanism of human beings can effectively and effortlessly guide people to find the most attractive and important regions from a scene. Various image saliency detection methods have been presented to equip computer vision with similar ability of automatically capturing the potential informative and interesting content. Motivated by the importance of saliency results in a variety of applications, such as image compression, object recognition, target tracking, *etc.*, there is rapid increase in the interest in saliency detection recently [1]. Different from predicting human eye fixation [2], salient object detection aims to highlight the whole object, which we focus on in this paper [3].

Most existing bottom-up saliency methods suppose that the distinctive appearance contrast between salient object and redundant background is apparently high, and measure the saliency in a center-surround contrast way [4-8]. As the most famous and influential models, Itti *et al.* [4] combine intensity, color and orientation features across multi-scale by local center-surround differences to measure saliency. Achanta *et al.* [5] construct a two-layer sliding window, and evaluate saliency via the color differences between the inner and outer windows. Based on the similar sliding window assumption, Rahtu *et al.* [6] intuitively utilize a Bayesian framework to compute the color contrast. But due to the lack

of high-level prior knowledge of real object, it directly defines the prior probability in the framework as a constant, which is too idealistic and fragile. To get more accurate results, the convex hull based Bayesian saliency models are presented in [7, 8], which circle the important Harris points to form a coarse salient region (foreground) and detect saliency through color and spatial information. But these methods are still failed when dealing with the pictures with complex backgrounds.

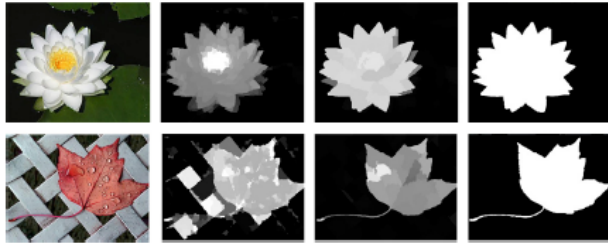
Moreover, some methods which are based on contrast organize the global salient clues in other ways [9-12]. Frequency-tuned method [9] defines the saliency of a pixel by its color difference with respect to the average image color. Cheng *et al.* [10] present a region contrast-based saliency method, and improve it in [11] by incorporating the soft image abstraction technique. Perazzi *et al.* [12] estimate saliency by the efficient N-D Gaussian filtering.

The above methods can achieve acceptable results in their own aspects, but they only consider the main salient features of foreground prior, which is insufficient for the generic salient object detection problem. Thus, from the perspective of background prior, Wei *et al.* [13] suppose that most of the background regions connect to image boundaries and define the geodesic saliency of a region as its shortest path to the pre-assumed background regions. The above idea achieves excellent performance among the state-of-the-art methods, but it does not hold when dealing with the pictures with complex backgrounds and highly textured images. Inspired by the work of [13], Shuai *et al.* [14] directly compute a superpixel's saliency as the summation of its  $k$  minimum color distance to the background superpixels after texture

<sup>\*</sup>Address correspondence to this author at School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu, 210094, P.R. China; Tel: +86 25 84315660-13; E-mail: Lv\_jy@126.com

suppression. Zhu *et al.* [15] make a great improvement via the robust boundary connectivity and saliency optimization.

To overcome the drawback of geodesic saliency, which only pays attention to the background information, we consider both the background contrast prior and potential foreground distribution. The statistical results of randomized Prim's algorithm which is based on object proposals are used to get a coarse location of the salient foreground area. The spatial interaction information between the estimated coarse conspicuity map and superpixel based color clusters is computed to acquire the foreground probability map. Then, the above map is effectively integrated with background color contrast to get the improved saliency results. Some representative saliency maps of [13] and our improved method are shown in Fig. (1), our results can uniformly and accurately highlight the real salient object.



(a) Original image (b) GS-SP [13] (c) Our (d) Ground Truth

Fig. (1). Some saliency maps of GS-SP and our method.

## 2. COARSE CONSPICUITY MAP

Some small-scale variations or textures of background may influence the saliency computation due to their high local contrast. First, we employ the simple Gaussian filter to smooth out these undesired image details. Then, we present the foreground prior generated by randomized Prim's algorithm which is based on object proposals. It can effectively alleviate the dependency on the background assumption of geodesic saliency [13].

The randomized Prim's algorithm is applied to a graph model to get high-quality object proposals quickly with some bounding boxes in [16]. The main principle of this method is that the groups of some connected superpixels included in the same object can be estimated by sampling partial spanning trees of superpixels that have high sum of edge weights. Let  $G = (V, E, \rho)$  be the weighted connectivity graph of an image, where the vertices  $V$  are the superpixels generated by the efficient graph based image segmentation [17], and the edges  $E$  connect two neighbor nodes. The weight  $\rho_{i,j}$  represents the probability that the superpixels  $i$  and  $j$  are included in the same object.

How to set the weight function  $\rho$  is very important. Manen *et al.* [16] utilize the logistic regression to obtain  $\rho_{i,j}$ :

$$\rho_{i,j} = \sigma(W^T \Phi_{i,j} + b) \quad (1)$$

$$\sigma(x) = (1 + \exp(-x))^{-1} \quad (2)$$

where  $\Phi_{i,j}$  is a meaningful feature vector, which measures the appearance similarity between  $i$  and  $j$ . The normalized color consistency, common border ratio and size

are used to constitute the feature vector.  $\sigma$  is the sigmoid function. The weights  $w$  and bias term  $b$  can be learned from the training data.

Then, for the node  $i$ , its neighbor nodes are defined as  $N(i)$ . If there is a set of vertices  $S$ , the frontier  $N(S)$  can be formulized as:

$$N(S) = \cup_{i \in S} N(i) \setminus S \quad (3)$$

Based on the above graph model, the randomized Prim's algorithm [18] is used as an iterative tree growing procedure to construct the partial spanning tree  $T_k$  at the  $k$ -th iteration. At each iteration, a candidate node from frontier  $N(S)$  is added to the tree  $T_k$  according to the edge weight by multinomial sampling strategy. It always selects the edge with maximum weight.

When the randomized stopping criterion is met, *i.e.*, the opportunity of adding a new node to  $T_k$  is low enough, the expected partial spanning tree is generated, which indicates the location of potential object. Then, the bounding box is used to generate the final object proposal (see Fig. (2b)) as the foreground prior.

To get more stable and accurate results, we sample  $N$  bounding boxes based on the randomized Prim's algorithm independently and overlap them. Then, the conspicuous value in the coarse conspicuity map of a pixel  $x$  is:

$$C(x) = \sum_{k=1}^N p_k(x) \quad (4)$$

where the probability  $p_k(x)$  is defined as:

$$p_k(x) = \begin{cases} 1, & x \in B_k \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$B_k$  represents the  $k$ -th bounding box.  $x \in B_k$  means the location of  $x$  is within  $B_k$ . In our experiments,  $N=1000$ .

By normalizing the conspicuous values of all pixels within the range  $[0,1]$ , we get the coarse conspicuity map. A representative example is shown in Fig. (2c), it can roughly provide the location of potential foreground regions.

## 3. FOREGROUND PROBABILITY MAP

The coarse conspicuity map (CCM) can effectively disregard most background while preserving important foreground regions, but it cannot establish well-defined boundaries of the large-scale salient object precisely due to the geometric shape of various bounding boxes (Fig. 2c). Therefore, we introduce the spatial interaction information to produce the foreground probability map.

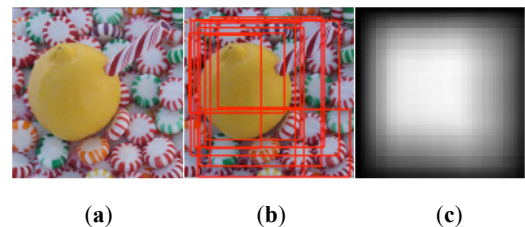


Fig. (2). (a) is the original image, (b) is the object proposal results annotated by bounding boxes, we choose 10 for visualization, and (c) is the coarse conspicuity map.

Human vision tends to take perceptual homogeneous image regions as a unit. To achieve the similar capability for computer vision, we employ Quick Shift [19] method to abstract the image into perceptually uniform superpixels (Fig. 3a); each superpixel is shown with the average color of all pixels in it). Note that this method is different from the graph based image segmentation [17] used in Section II. The reason of our choice is that Quick Shift has the ability of shape well adaptivity, boundary alignment and proper size control of individual superpixels.

To emphasize the whole salient object, we adopt the affinity propagation clustering (APC) to assemble those superpixels together according to their color appearance similarity (Fig. 3b). APC does not need to predefine the cluster number, which is very applicable to process the unknown color distributions of different images.

Now, we have decomposed the image into 2 high-layer hierarchical structure—superpixel layer and cluster layer. Then, we gradually refine the CCM in each layer to get the final foreground probability map.

We assume the image is abstracted into  $M$  superpixels and the  $i$ -th superpixel is  $sp_i$ . Then, the value of  $sp_i$  in the superpixel layer refined CCM is:

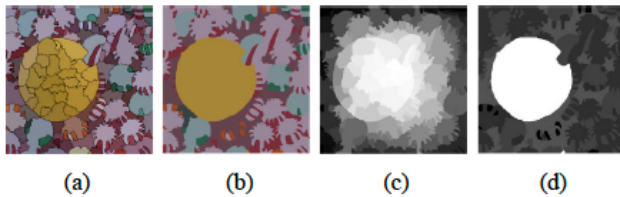
$$CR_i = \frac{1}{M_i} \sum_{j=1}^{M_i} C(x_j), \quad x_j \in sp_i \quad (6)$$

where  $M_i$  is the number of pixels included in  $sp_i$  and  $x_j$  is the  $j$ -th pixel. The result is shown in Fig. (3c).

To greatly enhance the real salient object, the global spatial interaction information of cluster layer should be considered. If all the superpixels aggregate to  $N$  clusters and the  $k$ -th cluster is  $c_k$  with  $N_k$  superpixels in it, the value of  $c_k$  in the foreground probability map is:

$$F_k = \frac{1}{N_k} \sum_{i=1}^{N_k} CR_i, \quad sp_i \in c_k \quad (7)$$

Equation 7 gives an intuitive understanding of highlighting entire meaningful object. Some background superpixels have relative high conspicuous values (Fig. 3c), but they can be significantly reduced by simply considering other cluttered background superpixels belonging to the same cluster in spatial domain (Fig. 3d).

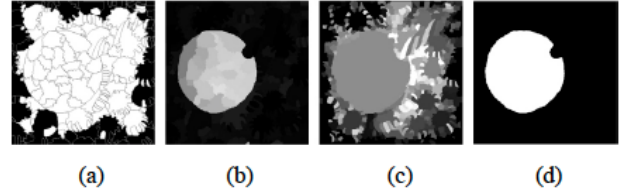


**Fig. (3).** (a) is the Quick Shift superpixel decomposition result, (b) is the APC result, (c) is the refined coarse conspicuity map in superpixel layer, (d) is the foreground probability map

#### 4. FINAL SALIENCY MAP

The foreground probability map can be integrated with the traditional background prior assumption to improve the overall salient object detection performance.

The geodesic saliency method [13] intuitively assumes that the image boundaries are mostly background regions. To solve the problem of some salient object regions touching image boundaries, it presents the “one-dimension rule”, which only computes the saliency of all image boundaries and confirms the real background regions.



**Fig. (4).** (a) is the schematic diagram of background regions, predefined background superpixels are shown in black color, (b) is our final saliency map, (c) is the saliency map of [13], (d) is the Ground Truth

Inspired by it, we adopt the same background assumption (Fig. 4a). But different from using geodesic distance to measure the saliency of each superpixel with respect to the predefined background superpixels, we directly utilize the color difference.

The background contrast of  $sp_i$  is defined as the summation of its  $L$  minimum color distance with respect to the background superpixel  $sp_b$  in CIELab space [14]:

$$B_i = \min \left\{ \sum_{b=1}^L D(sp_i, sp_b) \right\} \quad (8)$$

In our experiments, we set  $L=10$ .

Finally, we combine the foreground probability of  $sp_i$  with its background contrast to get its final salient value:

$$S_i = F_k \cdot B_i, \quad sp_i \in c_k \quad (9)$$

Normalize all the values to  $[0, 1]$ , we get the final saliency map. As shown in Fig. (4b), our saliency map can highlight the entire object uniformly and accurately while geodesic saliency assigns high salient value to cluttered background (Fig. 4c).

#### 5. EXPERIMENTS

We test our approach on two benchmark datasets. The first one is MSRA-1000, which contains accurate human labeled Ground Truth (GT) provided by Achanta *et al.* [9]. The second one is the SED dataset [20], which is divided into: the single object sub-dataset SED1 and two objects sub-dataset SED2. Each sub-dataset contains 100 images with GT.

To show the effectiveness and efficiency of our method, we compare the results of ours with those of 11 state-of-the-art methods, including: Itti [4], AC [5], FT [9], RA [6], RC [10], XL [7], SF [12], GS-SP [13], GC [11], BS [8] and PCAS [21]. Note that the most similar method to ours is GS-SP. GS-SP directly adopts the similar background assumption, while ours uses the superpixel based implementation of geodesic saliency (GS-SP) to get the better results.

##### 5.1. Evaluation on MSRA-1000 Dataset

Firstly, similar to FT [9], we utilize the precision-recall (P-R) curve, which is obtained by varying the thresholds from 0 to 255 and computing the precision and recall under

all the thresholds, as the first quantitative evaluation. As shown in Fig. (5), the P-R curve of our proposed method outperforms others significantly. Note that when the recall is under 0.9, the precision of our method (about 0.96) is apparently higher than that of other methods, especially GS-SP [13], demonstrating the effectiveness of combining important foreground prior and background contrast in our method.

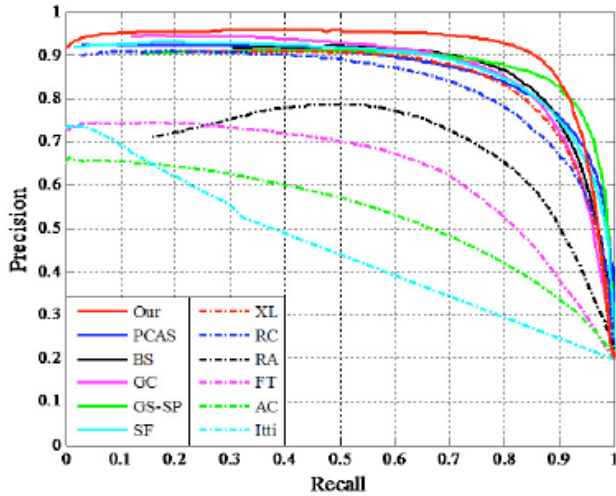


Fig. (5). The P-R curves of our method and 11 state-of-the-art methods.

Besides the precision and recall, we use the F-measure to evaluate the overall performance:

$$F - measure = \frac{(1 + \beta^2) Precision \times Recall}{\beta^2 Precision + Recall} \quad (10)$$

We also set  $\beta^2=0.3$  to weigh precision more than recall [9]. Then we draw the F-measure vs. T curve by calculating the F-measure at each threshold T ranging from 0 to 255 when drawing the P-R curve. Fig. (6) shows the comparison, where the red line is the result of our saliency map. It is apparent that the the F-measure of our method is higher than any other alternatives when the threshold T varies in the range [25,125]. We also calculate the maximum F-measure and the corresponding threshold T of each compared method. As shown in Table 1, our method generates the highest F-measure 0.8949 when T=74. Note that the best F-measure of GS-SP is only 0.8581, indicating the great improvement of our method.

We adopt the adaptive thresholding [9-12], which is defined as twice of the average image salient values, as the third evaluation. From Fig. (7), we conclude that: (1) the precision of our method (0.91) is the highest value; (2) although the recall of our approach (0.84) is lower than that of GS-SP (0.89), the F-measure of our method has the best performance (0.89), which implies that only using the simple threshold segmentation on our saliency map, we can get more consistent result with the GT.

Finally, we employ the mean absolute error (MAE) provided in [12] to both consider these pixels correctly marked as a non-salient region (denoted as the black pixels in GT) and the pixels labeled as a salient region (denoted as the white pixels in GT). As shown in Fig. (8), the MAE of our

method has the second minimum value (0.1128, slightly higher than that of GC 0.1016), proving that our saliency map is very close to human labeled Ground Truth.

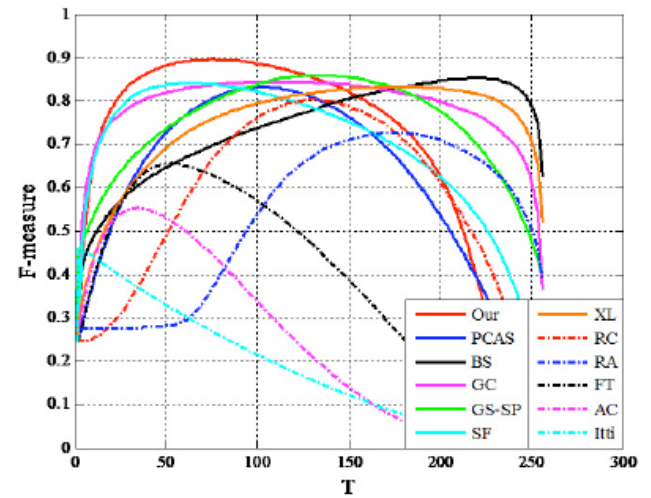


Fig. (6). The F-measure vs. threshold T Curves of our method and 11 state-of-the-art methods.

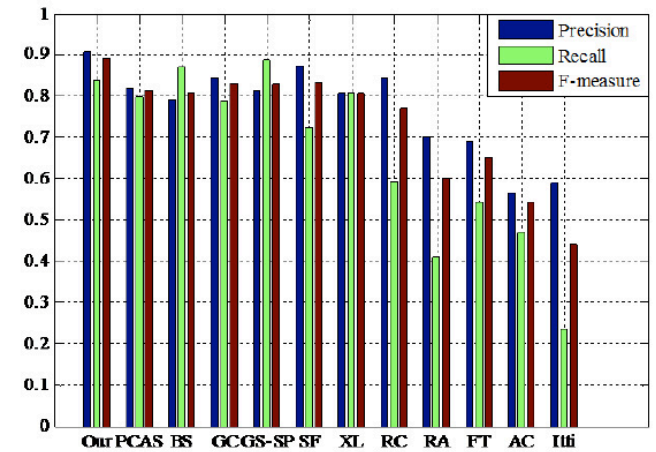


Fig. (7). The Precision, Recall and F-measure after adaptive thresholding.

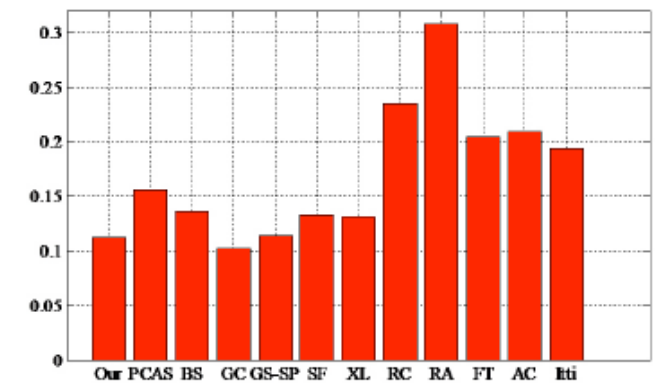
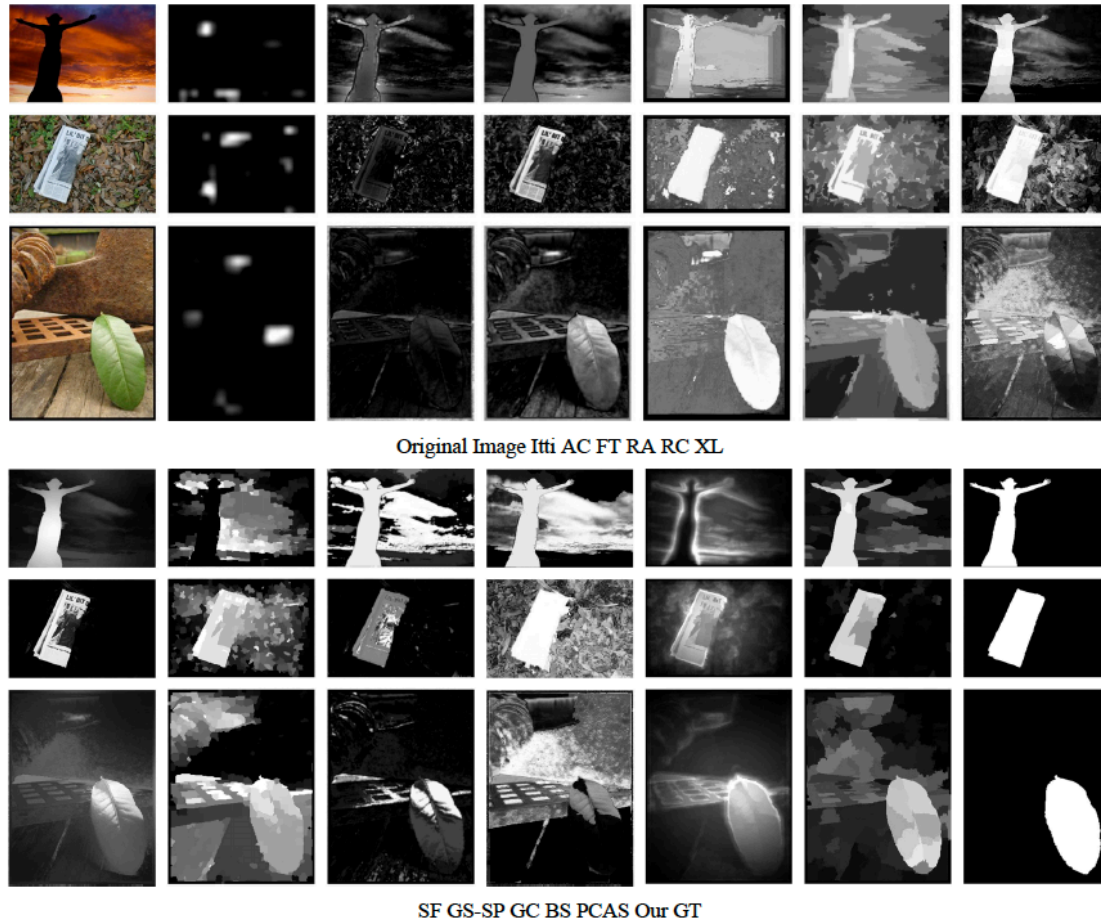


Fig. (8). The MAE histograms.

**Table 1.** The comparison of the maximum F-measure of different methods on MSRA-1000.

Method	Our	PCAS	BS	GC	GS-SP	SF	XL	RC	RA	FT	AC	Itti
F-measure	<b>0.8949</b>	0.8309	0.8515	0.8443	0.8581	0.8395	0.8317	0.8031	0.7272	0.6550	0.5528	0.4633
T	74	102	222	106	131	59	176	130	173	53	34	3

**Fig. (9).** The qualitative visual comparison of 11 state-of-the-art methods and our method.

Some representative saliency maps are shown in Fig. (9). Observed from the visual comparison results, we find that our method can correctly and uniformly assign high salient values to the real salient object while effectively suppress diversified background regions. Note that when compared with GS-SP, only based only on the background prior, our method has obviously better visual effect.

## 5.2. Evaluation of SED Dataset

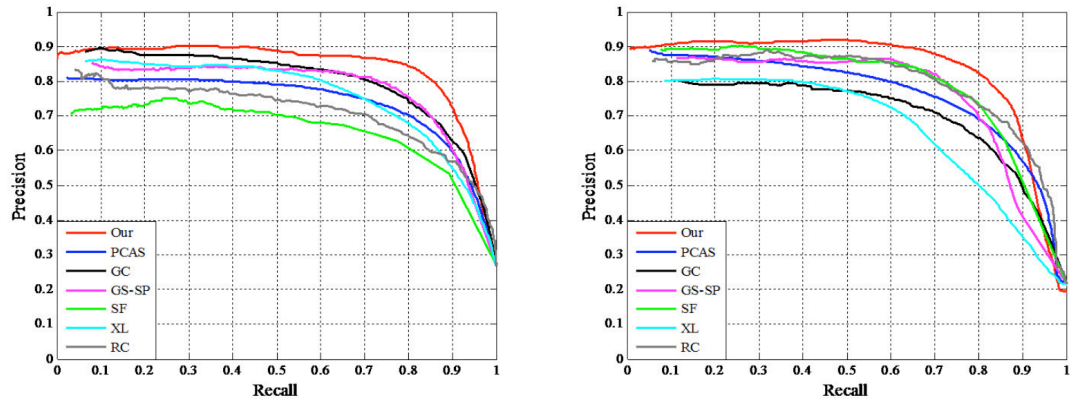
SED dataset is much more complex than MSRA-1000. Similar to [22], on this single object dataset SED1 and two objects dataset SED2, we carefully compare our method with 6 recent proposed state-of-the-art methods for their better performance evaluation, including RC [10], XL [7], SF [12], GS-SP [13], GC [11] and PCAS [21].

The P-R curves on the sub-datasets SED1 and SED2 are shown in Fig. (10). We can see that our method outperforms others. Note that the precision of our method is much higher

than GS-SP when the recall is at any location on SED1 and SED2, demonstrating the improvement of the proposed method when dealing with more complex images.

As observed from the F-measure vs. Threshold T curves in Fig. (11), the F-measure of the proposed method on SED1 and SED2 datasets are highest when T in the range [60,160] and [50, 200], respectively. And more specifically, as shown in Table 2, our method produces the best F-measure 0.8350 on SED1 and 0.8298 on SED2.

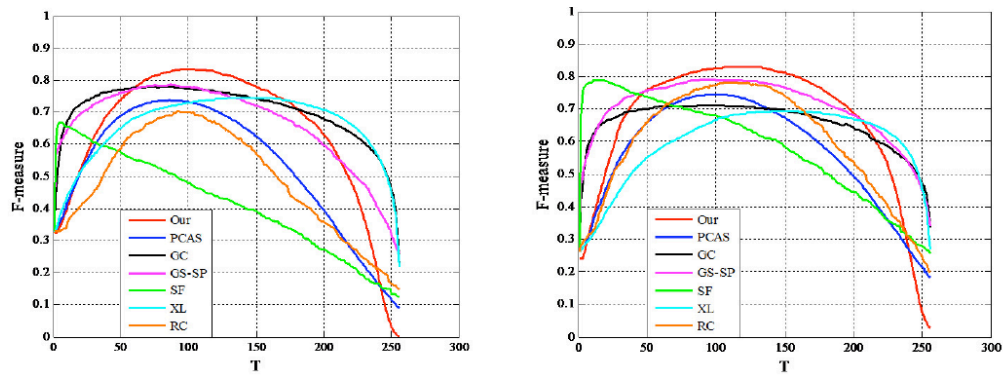
After adaptive thresholding, the average F-measure of our method is apparently higher than any other state-of-the-art methods on both SED1 and SED2 datasets as shown in Fig. (12). Note that on SED2, SF has higher precision value, because it tends to pop out the most salient regions at the expense of low recall, while our method can avoid it.



(a) The P-R curves on SED1 sub-dataset

(b) The P-R curves on SED2 sub-dataset

Fig. (10). The P-R curves of our method and 6 state-of-the-art methods.



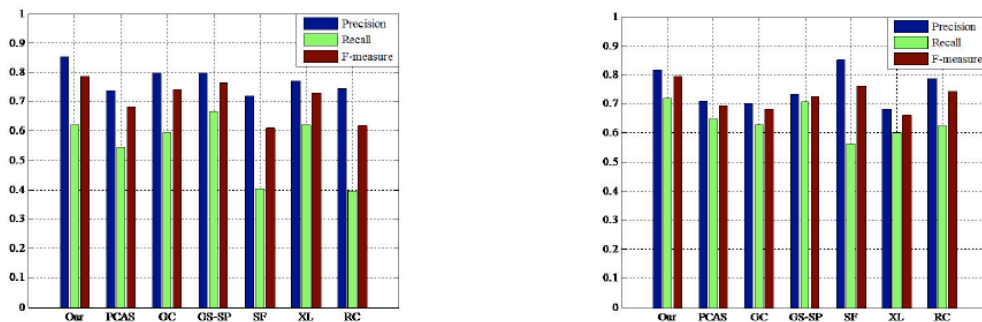
(a) F-measure vs T curves on SED1 sub-dataset

(b) F-measure vs T curves on SED2 sub-dataset

Fig. (11). The F-measure vs. T curves of our method and 6 state-of-the-art methods.

Table 2. The comparison of the maximum F-measure of different methods on SED1 and SED2 datasets.

	Method	Our	PCAS	GC	GS-SP	SF	XL	RC
SED1 dataset	F-measure	<b>0.8350</b>	0.7368	0.7797	0.7866	0.6685	0.7451	0.7041
	T	101	86	86	87	6	151	92
SED2 dataset	F-measure	<b>0.8298</b>	0.7447	0.7126	0.7907	0.7890	0.6938	0.7803
	T	121	101	95	94	13	150	109



(a) The results after adaptive thresholding on SED1 sub-dataset

(b) The results after adaptive thresholding on SED2 sub-dataset

Fig. (12). The Precision, Recall and F-measure of our method and 6 state-of-the-art methods.

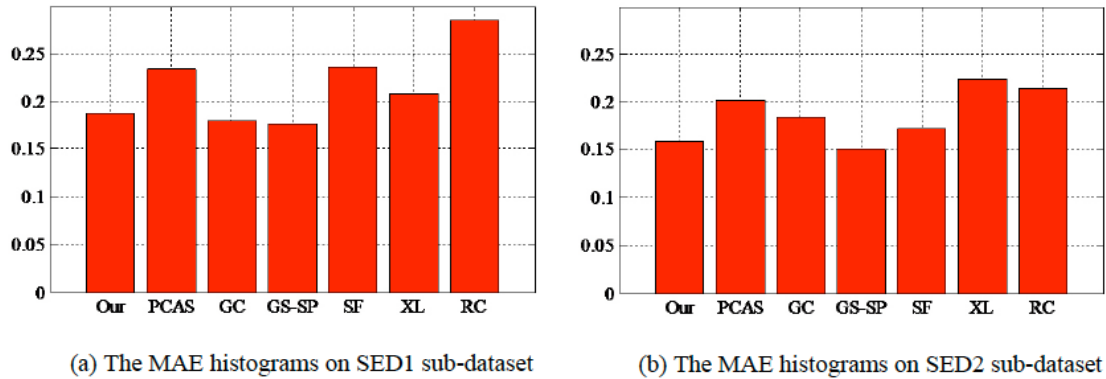


Fig. (13). The MAE histograms of our method and 6 state-of-the-art methods.

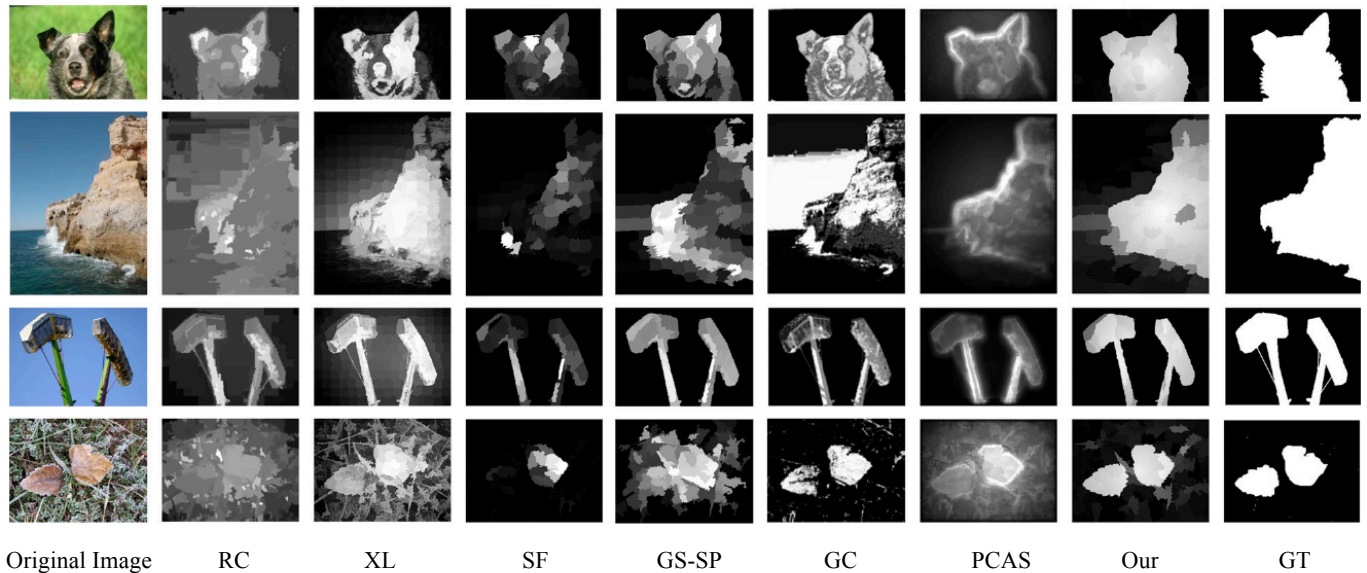


Fig. (14). The representative visual comparison of different methods on SED dataset. The first two rows are from SED1 sub-dataset, the last two rows are from the SED2 sub-dataset.

Table 3. Comparison of average running times on MSRA-1000 dataset.

Method	Our	PCAS	BS	GC	GS-SP	SF	XL	RC	RA	FT	AC	Itti
Time(s)	8.54	6.15	155.86	0.09	7.41	0.15	2.20	0.13	9.45	0.19	0.07	0.34

As shown in Fig. (13), the MAE value of our method is slightly higher than those of GC and GS-SP on SED1. On SED2, our method has the second minimum MAE value. This is because our method tends to assign some background regions relatively higher salient value compared to GS-SP, while highlights the most important salient objects. But considering the above three quantitative evaluation comparison results, our method improves the background prior based saliency method GS-SP significantly.

Some representative visual examples are shown in Fig. (14). No matter one or two objects existing in images, our method can highlight the whole object accurately and uniformly while effectively suppresses the background regions compared to the 6 state-of-the-art methods, i.e., our saliency map is more consistent with human labeled GT. We take the second image for an example, RC and GC assign some

background regions with undesired high salient values while lacking most real salient regions in the object. SF wrongly takes the small background region as the real object. XL, GS-SP and PCAS produce low salient value in some important regions within the object.

### 5.3. Running Time

We compare the average running time of the proposed method with those of the 11 state-of-the-art methods on MSRA-1000 dataset. The experimental environment is a computer with Inter(R) Core(TM)i5-2410M 2.8GHz CPU and 8GB RAM and we implement the proposed method by Matlab. As shown in Table 3, our method is only faster than BS and RA. But when considering both the quantitative evaluation results and the average running, our method still has some advantage.

## CONCLUSION

We proposed a novel saliency detection method by exploiting the statistical results of randomized Prim's algorithm which is based on the object proposal and background contrast prior assumption. We found that the coarse conspicuity map generated by the randomized Prim's algorithm can greatly emphasize the potential salient object at a rough scale. The foreground probability map produced by simple superpixel and cluster based refinement can provide more accurate estimation of real object. The above information can be used to optimize the naive background contrast saliency measure successfully. Extensive experiments demonstrate that our method is apparently superior to 11 state-of-the-art alternatives, especially GS-SP which is built on the previous background prior. In the future, we will consider high level saliency clues to improve the detection accuracy.

## CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

## ACKNOWLEDGEMENTS

This research was supported by the Natural Science Foundation of China (NSFC), No. 61473154.

## REFERENCES

- [1] A. Toet, "Computational versus psychophysical bottom-up image saliency: a comparative evaluation study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2131-2146, 2011.
- [2] A. Borji, D.N. Sihite, and L. Itti, "Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 55-69, 2013.
- [3] A. Borji, D.N. Sihite, and L. Itti, "Salient object detection: a benchmark", In: *Proceedings European Conference on Computer Vision*, Springer Press, 2012, pp. 414-429.
- [4] 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, 1998.
- [5] R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk, "Salient region detection and segmentation", In: *Proceedings IEEE International Conference on Computer Vision Systems*, Springer Press, 2008, pp. 66-75.
- [6] E. Rahtu, J. Kannala, M. Salo, and J. Heikkilä, "Segmenting salient objects from images and videos", In: *Proceedings European Conference on Computer Vision*, Springer Press, 2010, pp. 366-379.
- [7] Y. Xie, and H. C. Lu, "Visual saliency detection based on Bayesian model", In: *Proceedings IEEE International Conference on Image Processing*, IEEE Press, 2011, pp. 645-648.
- [8] Y. Xie, H. Lu, and M. Yang, "Bayesian saliency via low and mid level cues", *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1689-1698, 2013.
- [9] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequency-tuned salient region detection", In: *Proceedings IEEE International Conference on Computer Vision and Pattern Recognition*, IEEE Press, 2009, pp. 1597-1604.
- [10] M. Cheng, G. Zhang, N. J. Mitra, X. Huang, and S. Hu, "Global contrast based salient region detection", In: *Proceedings IEEE International Conference on Computer Vision and Pattern Recognition*, IEEE Press, 2011, pp. 409-416.
- [11] M. Cheng, W. Jonathan, W. Lin, Z. Shuai, V. Vibhav, and C. Nigel, "Efficient salient region detection with soft image abstraction", In: *Proceedings IEEE International Conference on Computer Vision*, IEEE Press, 2013, pp. 1529-1536.
- [12] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection", In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Press, 2012, pp. 733-740.
- [13] Y. Wei, F. Wen, W. Zhu, and S. Jian, "Geodesic saliency using background priors", In: *Proceedings European Conference on Computer Vision*, Springer Press, 2012, pp. 29-42.
- [14] J. Shuai, L. Qing, J. Miao, Z. Ma, and X. Chen, "Salient region detection via texture-suppressed background contrast," In: *Proceedings IEEE International Conference on Image Processing*, IEEE Press, 2013, pp. 2470-2474.
- [15] W. J. Zhu, S. Liang, Y. C. Wei, and J. Sun, "Saliency optimization from robust background detection", In: *Proceedings IEEE International Conference on Computer Vision and Pattern Recognition*, IEEE Press, 2014, pp. 2814-2821.
- [16] S. Manen, M. Guillaumin, and L. V. Gool, "Prime object proposals with randomized Prim's algorithm", In: *Proceedings IEEE International Conference on Computer Vision*, IEEE Press, 2013, pp. 4321-4328.
- [17] P. F. Felzenszwalb, and D. P. Huttenlocher, "Efficient graph-based image segmentation", *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167-181, 2004.
- [18] R.C. Prim, "Shortest connection networks and some generalizations", *Bell System Technology Journal*, vol. 36, no. 6, pp. 1389-1401, 1957.
- [19] A. Vedaldi, and S. Soatto, "Quick shift and kernel methods for mode seeking", In: *Proceedings European Conference on Computer Vision*, Springer Press, 2008, pp. 705-718.
- [20] R. B. S. Alpert, M. Galun, and A. Brandt, "Image segmentation by probabilistic bottom-up aggregation and cue integration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 315-327, 2012.
- [21] R. Margolin, A. Tal, and L. Zelink-Manor, "What makes a patch distinct? ", In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Press, 2013, pp. 1139-1146.
- [22] W. Jiang, L. H. Zhang, H. C. Lu, C. Yang, and M.-H. Yang, "Saliency detection via absorbing markov chain", In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Press, 2013, pp. 1665-1672.

Received: May 26, 2015

Revised: July 14, 2015

Accepted: August 10, 2015

© Jianyong et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.