Graph-Regularized Saliency Detection With Convex-Hull-Based Center Prior

Chuan Yang, Lihe Zhang, and Huchuan Lu

***Abstract*—Object level saliency detection is useful for many content-based computer vision tasks. In this letter, we present a novel bottom-up salient object detection approach by exploiting contrast, center and smoothness priors. First, we compute an initial saliency map using contrast and center priors. Unlike most existing center prior based methods, we apply the convex hull of interest points to estimate the center of the salient object rather than directly use the image center. This strategy makes the saliency result more robust to the location of objects. Second, we refine the initial saliency map through minimizing a continuous pairwise saliency energy function with graph regularization which encourages adjacent pixels or segments to take the similar saliency value (i.e., smoothness prior). The smoothness prior enables the proposed method to uniformly highlight the salient object and simultaneously suppress the background effectively. Extensive experiments on a large dataset demonstrate that the proposed method performs favorably against the state-of-the-art methods in terms of accuracy and efficiency.**

***Index Terms*—Center prior, contrast prior, salient object detection, smoothness prior.**

I. Introduction

SALIENT object detection is the process of identifying the location of the object, which is different with the traditional models which predict human fixations [1]. Salient object detection is related to many applications, including image segmentation, object recognition and content-based image retrieval. Therefore, numerous saliency models have been developed in recent years, which can be categorized as either bottom-up [2]–[6] or top-down [7] approaches. In this letter, we focus on the bottom-up salient object detection.

Bottom-up saliency methods are data-driven and rely on some pre-defined assumptions (i.e., priors) [2]–[5]. The most influential prior is contrast, that is, salient object should present high appearance contrast with background. The contrast prior can be investigated from the local [2], [3] or global point of view [4]–[6]. Accordingly, the contrast measure is computed with respect to local neighbourhoods or the entire image respectively. While contrast prior based methods have achieved promising results, they still have some certain limitations. Typically, the whole object region can not be highlighted uniformly and the background region can not be suppressed effectively. Recently, Perazzi *et al.* [6] successfully suppress the background by using a compactness prior, but some parts of the salient object are also incorrectly suppressed. Center prior has been widely used in saliency detection [2], [3]. They assume that the salient object is often framed near the center of image. However, in many images, the salient objects often appear off the image center, which makes the center prior map incorrectly suppress the salient region far off the image center and highlight some background region near the image center. Furthermore, these methods independently compute the saliency of each image element (i.e., superpixel or pixel) and ignore the interrelationships between the adjacent elements (e.g., similar elements should have similar saliency), which can be critical for uniformly highlighting the whole salient region and adequately suppressing the background region.

Considering all above-mentioned issues, we introduce a convex-hull-based center prior and a smoothness prior into the saliency model. Note that the salient object can appear anywhere in image, we use a recently proposed convex hull technique [8] to estimate the center of the salient object. This makes the center prior more accurate and more robust to the location of salient object. The smoothness prior is inspired by graph based object segmentation technique, which considers similar and adjacent image elements should have the same label. This constraint makes the saliency of the whole object region more uniform and can partially rectify some errors in the background region, thereby suppressing the saliency of the background region.

The rest of this letter is organized as follows. The proposed saliency model is described in Section II. The experiments on a popular large saliency dataset are performed in Section III and the conclusion is made in Section IV.

II. Saliency Model

Our approach considers the superpixel as the element of saliency estimation. First, we over-segment a given image into superpixels using the SLIC method [9]. Then, we compute an initial saliency map based on contrast and center priors. Last, we refine the initial saliency map with smoothness prior. In the following, we describe the details of our method.

A. Initial Saliency Map

1) Contrast Prior Map: Spatially weighted contrast measure has been shown to be effective in saliency detection [5], [6]. For any superpixel i, given its color mean c_i in the CIE LAB color space and average position p_i with pixel coordinates normalized

Manuscript received December 23, 2012; revised April 07, 2013; accepted April 22, 2013. Date of publication April 30, 2013; date of current version May 13, 2013. The work of C. Yang and L. Zhang was supported by the Fundamental Research Funds for the Central Universities (DUT12JS05). The work of H. Lu was supported by the Natural Science Foundation of China under Grants 61071209 and 61272372. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Qian Du.

The authors are with the School of Information and Communication Engineering, Dalian University of Technology, Dalian 116023, China (e-mail: ycscience86@gmail.com; zhanglihe@dlut.edu.cn; lhchuan@dlut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2013.2260737

1070-9908/$31.00 © 2013 IEEE

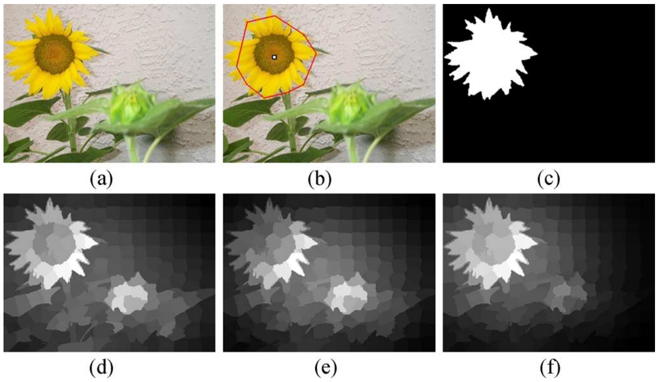

Fig. 1. Example for showing the benefits of convex-hull-based center prior. (a) input image. (b) convex hull and estimated center. (c) ground truth. (d) contrast prior map. (e) initial map S_{in} with image center prior. (f) initial map S_{in} with convex-hull-based center prior.

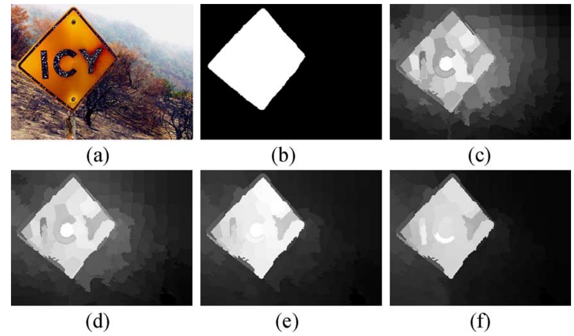

Fig. 2. Illustration of the smoothness prior. (a) input images. (b) ground truth. (c) initial map S_{in}. (d)–(f) refined saliency maps with different weights $\lambda = 2.5$, $\lambda = 25$, $\lambda = 100$. The refined map with $\lambda = 2.5$ is not enough uniform and the one with $\lambda = 100$ suppresses the salient object region. The map with $\lambda = 25$ is best.

to $[0, 1]$, its saliency is defined as its spatially weighted contrast to all other superpixels j:

$$S_{co}(i) = \sum_{j \neq i} \|c_i - c_j\| \cdot \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_p^2}\right), \qquad (1)$$

where σ_p controls the strength of spatial weighting.

2) Convex-Hull-Based Center Prior Map: The contrast prior map often incorrectly detects some background superpixels (See Fig. 1(d)). Thus, we introduce the center prior to alleviate this problem. Previously used center prior [2], [3] assigns higher saliency to the image elements near the image center. However, this principle becomes invalid when the objects are placed far off the image center. In this letter, we first compute a convex hull enclosing interesting points to estimate the location of salient region and then use the centroid of the convex hull as the center to get the convex-hull-based center prior map. Given the center (x_0, y_0), the saliency of a superpixel i is defined as:

$$S_{ce}(i) = \exp\left(-\frac{\|x_i - x_0\|^2}{2\sigma_x^2} - \frac{\|y_i - y_0\|^2}{2\sigma_y^2}\right), \qquad (2)$$

where x_i and y_i denote the mean horizontal and vertical positions of the superpixel i, σ_x and σ_y indicate the horizontal and vertical variances. In our implementation, we use a centered anisotropic Gaussian distribution to model the center prior map, i.e., we set $\sigma_x = \sigma_y$ with pixel coordinates normalized to $[0, 1]$.

The convex hull is first used in [8] as a rough salient region to facilitate inference of Bayesian saliency model, which achieves a good performance. Because the convex hull provides a rough location of the salient object, the convex-hull-based center prior map is more reasonable and robust.

3) Integration: Inspired by the structure adapted from the Feature Integration Theory [10], we fuse the above two prior maps to get the initial saliency map:

$$S_{in}(i) = S_{co}(i) \times S_{ce}(i). \qquad (3)$$

A persuasive example is presented in Fig. 1. By comparing Fig. 1(e) with (f), it can be found that the convex-hull-based center prior works well when the salient object appears far off the image center, while the image center prior incorrectly highlights the background region near the image center. The quantitative comparison is shown in Fig. 5(a). However, the benifits of the convex-hull-based center prior are up to the accuracy of the convex hull. In addition, since we adopt the symmetric distribution with the fixed parameters σ_x and σ_y, the shape asymmetry and scale changes of the salient object make the convex-hull-based center prior produce some incorrect detection in the initial map S_{in}.

To alleviate these issues, further uniformly highlight the salient region and adequately suppress the background region, we introduce a smoothness prior to take the interaction between image elements into account, thereby obtaining a refined saliency map.

B. Saliency Map Refining With Graph Regularization

The use of smoothness constraint has been well documented for graph-based object segmentation [11], [12]. The segmentation models encode the smoothness constraint by adding a pairwise potential to the energy function which encourages neighbouring pixels in the image to take the same label. In this letter, we introduce a similar pairwise potential into saliency model as the smoothness prior. We design a sparsely connected graph $G = (V, E)$, where the nodes V are a set of superpixels and the edges E are the undirected links which exist if two superpixels share a common boundary. The edge between two linked nodes i, j has the following weight $w_{ij} \in W$:

$$w_{ij} = \exp\left(-\frac{\|c_i - c_j\|}{2\sigma_w^2}\right), \qquad (4)$$

where c_i and c_j indicate the mean of the superpixels corresponding to two nodes in the CIE LAB color space, and σ_w controls the strength of the weight. Note that the affinity matrix W is highly sparse. Here, we define a saliency cost function to encode the smoothness prior:

$$E(S) = \sum_i (S(i) - S_{in}(i))^2 + \lambda \sum_{i,j} w_{ij}(S(i) - S(j))^2, \qquad (5)$$

where $S(i)$ and $S(j)$ correspond to the saliency values of node i and j respectively, $S_{in}(i)$ is the initial saliency value for node i computed in Section II-A, and λ is a regularization parameter. The first term of the right-hand side in the cost function is the fitting constraint, which means a good saliency map should not change too much from the initial saliency map. The second term is the smoothness constraint, which means that a good saliency

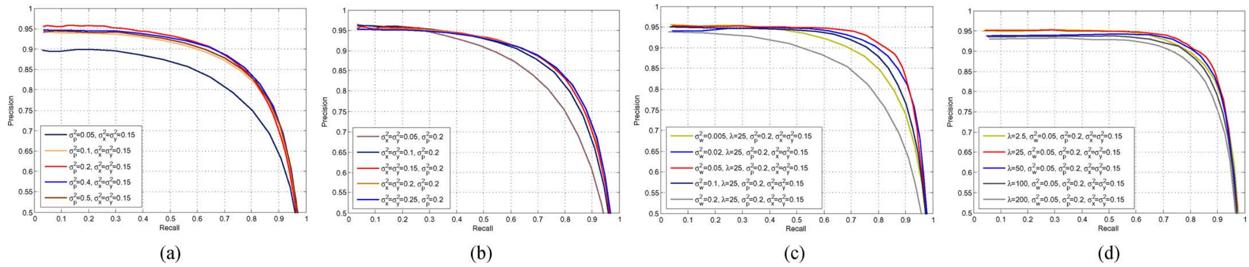

Fig. 3. Precision-recall curves for different parameter settings on the MSRA-1000 dataset. (a) Comparison of initial saliency maps with different σ_p. (b) Comparison of initial saliency maps with different σ_x and σ_y. (c) Comparison of refined saliency maps with different σ_w. (d) Comparison of refined saliency maps with different λ.

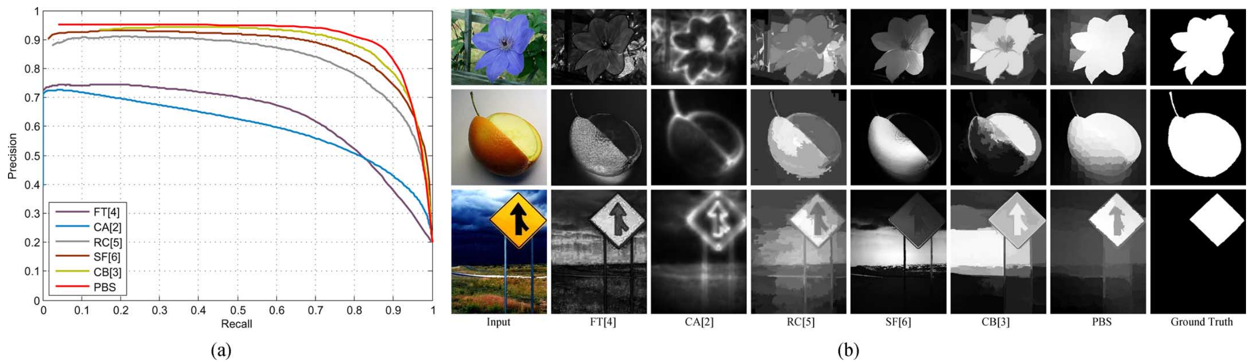

Fig. 4. Comparison with five state-of-the-art bottom-up saliency methods. (a) precision-recall curves of different methods on the MSRA-1000 dataset. (b) Visual comparison of saliency maps.

map should not change too much between neighbouring superpixels. The optimal saliency values of superpixels are computed by minimizing the above cost function, which can be solved by setting the derivative of the above function with respect to S to be zero. The resulting solution is:

$$S^* = \mu(D - W + \mu I)^{-1} S_{in}, \tag{6}$$

where D is the diagonal degree matrix with $d_{ii} = \sum_j (w_{ij})$ and $\mu = 1/(2\lambda)$.

The refined saliency map has several advantages. First, the saliency contrast of foreground and background is further extended, which can generate well-defined boundaries of salient objects. Second, some incorrect detection (e.g., there exist some isolated background elements with high saliency or isolated salient elements with low saliency) can be partially rectified, thereby getting more uniform foreground and background. As a result, it achieves significant performance improvement with a little computational expense increase. Fig. 2 shows the impact of smoothness prior on refining the saliency map. With a smaller λ, the fitting constraint is emphasized more, which means that the information of initial labels plays a more important role. With a larger λ, the smooth constraint is weighted more, which means that each node receives more information from its neighbors.

III. EXPERIMENTS

The experiments are conducted on a publicly available MSRA-1000 dataset of 1000 images [7] with accurate labeled ground truth [4]. Similar to [4], we use precision-recall curves to evaluate the algorithm. Given a saliency map with intensity values normalized to $[0, 255]$, a set of binary foreground segments are obtained by varying the threshold from 0 to 255. The precision-recall curve is then computed based on the ground truth mask.

A. Experimental Setup

We set the number of superpixel nodes $\mathrm{N} = 200$ in all the experiments. There are five parameters in the proposed algorithm: $\sigma_p, \sigma_x, \sigma_y, \sigma_w$ and λ. Since the proposed saliency model is unsupervised, these five parameters are empirically chosen, $\sigma_p^2 = 0.2$, $\sigma_x^2 = \sigma_y^2 = 0.15$, $\sigma_w^2 = 0.05$ and $\lambda = 25$, for all the test images in the experiments. Fig. 3 shows the sensitivity of the proposed model to the parameter settings.

B. Comparison With Other Methods

We compare the proposed prior based saliency (PBS) method with five currently best performing object level saliency detection methods: the FT [4], CA [2], RC [5], CB [3] and SF [6] methods. The FT method defines pixel-wise saliency based on pixel's contrast to average image color. The RC method proposes a region-based contrast measure to compute saliency map. In the SF method, two well-defined contrast measures based on the uniqueness and spatial distribution are formulated in a unified way using high dimensional Gaussian filters. These three methods are based on global contrast prior, while the CB and CA methods exploit local contrast and image center priors. Fig. 4(a) shows the precision-recall curves of all methods. We note that the proposed method outperforms the other five methods. At the high recall values from 0.96 to 1, the curve of the results computed by the SF method is slightly higher than that of our results. That is because the SF method suppresses the background more effectively than our method. However, when the recall values are lower than 0.96, their curve becomes lower, which shows that they can not highlight the whole object region uniformly. A visual comparison of the different methods is shown in Fig. 4(b).

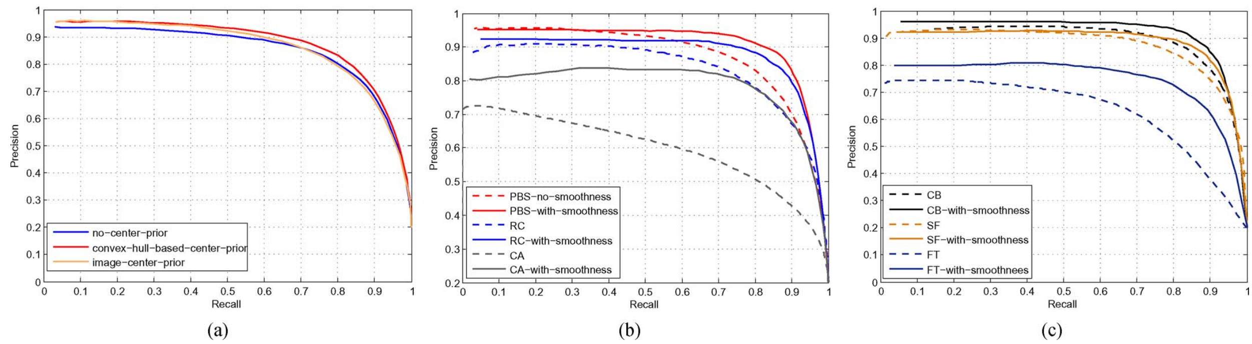

Fig. 5. Evaluation of center and smoothness priors on the MSRA-1000 dataset. (a) precision-recall curves of initial saliency maps with different priors. (b)-(c) precision-recall curves for all the methods with and without smoothness prior.

TABLE I
AVERAGING RUNNING TIME (SECONDS PER IMAGE) OF DIFFERENT METHODS MEASURED ON THE MSRA-1000 DATASET

Method	PBS	CB [3]	CA [2]	RC [5]	FT [4]	SF [6]
Time(s)	0.886	3.417	64.313	0.312	0.078	0.458
Code	Matlab	Matlab	Matlab	C++	C++	C++

C. Validation of Center and Smoothness Priors

In order to demonstrate the effect of the convex-hull-based center prior, we compute three precision-recall curves for the initial saliency maps with different priors (i.e., only contrast prior, contrast and previously used image center priors, contrast and convex-hull-based center priors). The resulting curves are shown in Fig. 5(a). Note that, at the higher recall values (i.e., small threshold), the curve of image center prior based results is slightly lower than that of results generated with only contrast prior. That is because the image center prior assigns high weights to the region near the image center, which are often background as the salient objects are often placed off the center (one-third rule in professional photography). Thus, the binary segment obtained with small threshold contains more background, which leads to lower precision.

We perform two experiments to test the validation of the smoothness prior. First, we compare the precision-recall curve of the initial saliency maps with that of the refined maps. Second, we use the saliency maps of the other five competitors as the initial maps respectively, and then investigate their refined results. The value of $S_{in}(i)$ in (5) is computed by averaging all the saliency values of the pixels within superpixel i. We compare the original results with the refined results. All the precision-recall curves are shown in Fig. 5(b) and (c). We note that by post-processing with the smoothness prior, all the methods achieve performance improvements and the refined results initialled by the CB method perform best.

D. Run Time

The average run times of all the methods on the MSRA-1000 dataset are presented in Table I based on a machine with Pentium Dual Core E5400 2.7 GHz CPU and 2GB RAM. Our run time is much faster than that of the other matlab implementations. Specifically, our method spends 0.227 s, 0.642 s and 0.017 s (about 2%) on superpixel generation, initial map computation and saliency map refining respectively. Note that the time of the SF method is a bit slower than that reported in their paper, since both the superpixel segmentation and the filter are not fully optimized in the code which we use to compute the run time.

IV. CONCLUSION AND DISCUSSION

In this letter, we present a bottom-up object level saliency detection model by exploiting contrast, center and smoothness priors. We estimate a more accurate center to fit the change of object location and introduce a smoothness prior to improve the result. The saliency maps on a large public database demonstrate that the proposed method can highlight the whole object region uniformly and suppress the background region effectively. In addition, the proposed method performs favourably against the other state-of-the-art methods in both accuracy and speed, which shows that the proposed convex-hull-based center prior and smoothness prior are helpful for saliency detection. In the future work, we will investigate a hierarchical image model to encode smoothness prior at different scales.

REFERENCES

[1] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, 1998.
[2] S. Goferman, L. Zelnik-Manor, and A. Tal, “Context-aware saliency detection,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 32, no. 10, pp. 1915–1925, 2012.
[3] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, and S. Li, “Automatic salient object segmentation based on contex and shape prior,” in *BMVC*, 2011.
[4] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, “Frequency-tuned salient region detection,” in *CVPR*, 2009.
[5] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang, and S. M. Hu, “Global contrast based salient region detection,” in *CVPR*, 2011.
[6] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, “Saliency filters: Contrast based filtering for salient region detection,” in *CVPR*, 2012.
[7] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum, “Learning to detect a salient object,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 33, no. 2, pp. 353–367, 2011.
[8] Y. L. Xie and H. C. Lu, “Visual saliency detection based on Bayesian model,” in *ICIP*, 2011.
[9] R. Achanta, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, “SLIC Superpixels,” Tech. Rep. EPFL, 2010, Tech.Rep. 149300.
[10] A. M. Treisman and G. Gelade, “A feature integration theory of attention,” *Cogn. Psych.*, vol. 12, pp. 97–136, 1980.
[11] C. Rother, V. Kolmogorov, and A. Blake, ““grabcut”: Interactive foreground extraction using iterated graph cuts,” *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004.
[12] L. Ladicky, C. Russell, P. Kohli, and P. H. Torr, “Associative hierarchical crfs for object class image segmentation,” in *ICCV*, 2009.