Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Salient Object Detection based on Multiple Priors Fusion *

Zhongli Wang, Guohui Tian*
Department of Control Science and Engineering
Shandong University
Jinan, Shandong Province, China
g.h.tian@sdu.edu.cn

***Abstract*—Salient object detection can be utilized to detect the most significant regions in various environments, which has been regarded as foundation of computer vision. Different saliency models use different prior or knowledge. We propose a multi priors fusion method for saliency measure, which integrates background prior with foreground prior and center prior. Firstly, through each boundary of the image, we can get four saliency maps, and fuse them to get the background prior saliency map. Secondly, we utilize boundary extension method to highlight regions, and these regions can be regarded as the queries of manifold ranking for the foreground prior saliency map. Thirdly, the corners on the image are obtained, filtered by the foreground region, and then clustered into a point as the center of Gaussian model, which is used to calculate the center prior saliency map. Finally, the above three kinds of prior-based saliency maps are fused via the proposed fusion framework to gain a better final saliency map. Compared with fifteen methods, the experimental results on ECSSD and MSRA10K show that our proposed method achieves better saliency detection results.**

***Index Terms*—salient object detection, multiple priors fusion, background prior, foreground prior, center prior**

I. INTRODUCTION

Salient object detection is of great significance in image compression [1], object recognition [2], and other applications, so ti has been widely concerned and studied. Based on the mechanism of salient object detection, this algorithm includes two types of methods, one is bottom-up, generally driven by data. This type are designed via use of a variety of low-level cues, such as texture, intensity, color,etc.. The other is top-down, generally driven by task or supervised learning. Salient object detection based on the top-down methods always have something to do with a special task or purpose. The top-down algorithms can work only when the target characteristics are known in advance. Therefore, the top-down algorithms require a supervised learning approach. However, the bottom-up methods adopt prior information about low-level visual information without the need to learn specific target information in advance [3–6], such as background prior, foreground prior and central prior. In addition to prior information, traditional methods also use a prior-based contrast [7, 8], whether local or global. A prior information or a prior method is effective in saliency target detection [10]. However, there is still an obvious problem. The problem is that computing a saliency model requires consideration of various prior information. It is not clear how to integrate it with other cues for saliency detection.

For the above universal problem, salient object detection model based on multi priors fusion by integrating background prior with foreground prior and center prior is proposed in this paper. The overall diagram of our proposed salient object detection model are present in Fig. 1. Firstly, an image is segmented into a series of superpixels which can be used for saliency measure by simple linear iterative clustering algorithm(SLIC). We regard all superpixels located as each boundary of the image as labeled item of manifold ranking based on graph to gain four saliency maps in turn and multiply their complementary terms to gain the background prior saliency map. Secondly, boundary extension is used to extend boundary and highlight foreground regions, which are regard these foreground regions as queries to rank with other image superpixels for foreground prior saliency measure. Thirdly, in order to gain center prior saliency map, we use corner detector to detect corner points on the image, reduce the number of corner points on the image with foreground saliency area, and then cluster for saliency calculation. Finally, we integrate three kinds of prior-based saliency maps for better saliency results.

II. DETAILS OF OUR PROPOSED METHOD

A. Brief introduction of manifold ranking

A graph can be represented in symbolic form as $G = (V, E)$, where V and E represent the set of nodes and the set of undirected edges between each node, respectively. The manifold ranking [9] based on graph can calculate the correlation between each node. Some nodes are regarded as queries, and the rest are regraded as unlabeled nodes, which need to be ranked based on their similarity between them. The weight between each adjacent node can be expressed in a matrix form Let the affinity matrix $W_{i,j} = [w_{i,j}]_{N\times N}$, and element $w_{i,j}$ of the affinity matrix is defined as:

$$w_{i,j} = \exp(-\frac{d_c^2(v_i, v_j)}{\sigma_2}) \quad (1)$$

*This work is supported by National Natural Science Foundation of China(U1813215, 61773239), and the Taishan Scholars Program of Shandong Province

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

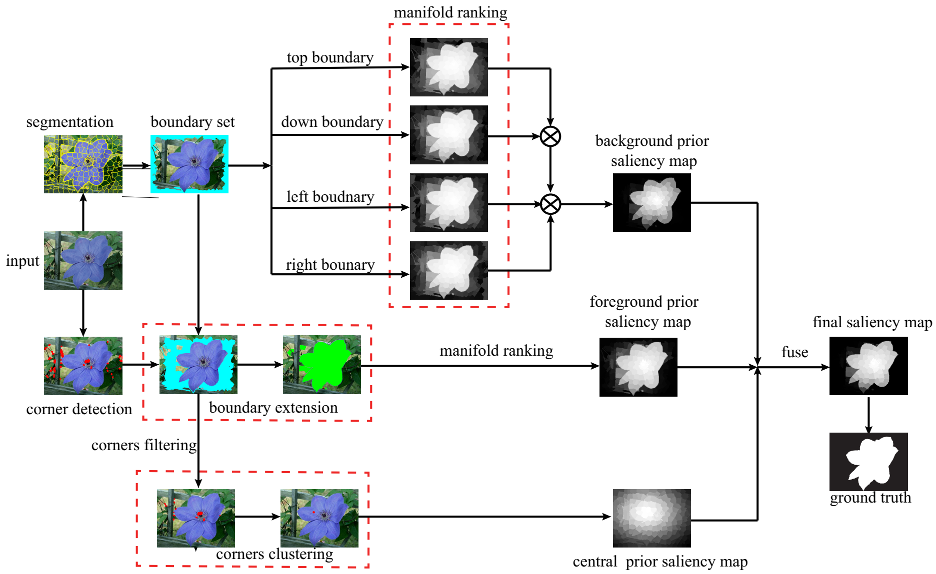


Fig. 1. Overall diagram of our proposed saliency detection algorithm

Where d_c represents the Euclidean distance of the nodes v_i and v_j.

Indication vector are presented as $y = [y_1, y_2, \cdots, y_n]$. Where $y_i = 1$ indicates that node is labelled as the query item, otherwise, $y_i = 0$. Let D denote the degree of Affinity matrix W, that is, $D = diag\{d_{11}, d_{22}, \cdots, d_{mm}\}$, where $d_{ii} = \sum_{j=1}^{n} w_{ij}$, α denote the relative contributions of neighborhood to ranking score and initial ranking score. Therefore, manifold ranking algorithm, which indicates the degree of association between unlabel nodes and labeled nodes, is used to estimate its relevance to the query item. We get ranking function f:

$$f = (D - \alpha W)^{-1} y \tag{2}$$

B. Background prior saliency measure

SLIC divides an image into a series of image patches, namely, a series of superpixels that can be regarded as nodes in a graph. So, nodes of graph can be composed of image superpixels. For the calculation of the background prior saliency map, an image boundary is selected as an example. Superpixels on the boundary are regarded as queries, and the rest superpixels as unlabel nodes. In this way, we get the given indication vector y, affinity matrix is computed based on the (1), where v is the mean color of superpixel. According to the (2), we rank all nodes and get the vector f, whose dimensional is equal to the number of superpixels on the image. Each element in vector f represents the relevances between the superpixel on the image boundary and other image superpixels that are not on this image boundary. The complement of vector f is obtained, and it is taken as the value of each superpixel. We normalize this vector, and obtain saliency map s_b:

$$s_b^{\ell}(i) = 1 - \bar{f}_b^{\ell}(i), i = 1, 2, \cdots, N \tag{3}$$

Where i represents a superpixel, ℓ represents each boundary of an image and $\ell \in \{top, bottom, left, right\}$.

Four saliency maps are generated according to four boundaries of the image, which can be further fused to generate a saliency map with better performance, that is, background prior saliency map S_b:

$$S_b(i) = \prod_{\ell=1}^{4} s_b^j(i) \tag{4}$$

An example of background prior saliency measure is presented in Fig. 2. Background prior saliency map(Fig. 2(e)) has better performance than saliency maps that are presented in Fig. 2(d) ranked by each image boundary respectively. Although,Fig. 2(e) shows that background prior saliency map can highlight salient object, there still exists some background regions.

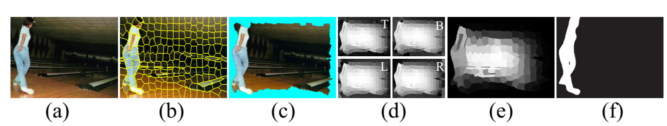


Fig. 2. An example of background prior saliency measure. (a)input image,(b)image superpixels,(c)image boundary set(regions marked as blue color), (d)saliency measure based on each boundary, (e)background prior saliency map, (f)ground truth.

C. Saliency measure based on foreground prior and center prior

1) Foreground prior saliency measure based on boundary extension: There may exist more noise in background prior saliency map, which can fail to highlight foreground regions. Therefore, boundary extension are utilized to extended to background regions for foreground regions with better performance.

Suppose that there are two superpixels i and j,d_p is used to represent the Euclidean distance between center position coordinates of superpixel. According to the width and height of the image, the two-dimensional coordinate values of the image superpixel are normalized respectively. The mean color in CIELab is expressed as d_c. Define the difference measures between image superpixels as follows:

$$mer(i, j) = \exp(\frac{d_p^2(i, j)}{\sigma_3^2})d_c(i, j) \tag{5}$$

where σ_3 is a constant that adjusts the intensity of differences between image superpixels. Its value is set to range of 0.5 to 1 based on the stability of boundary extension experiment.

Let B denote the set of boundary superpixels of an image and v denote image superpixel except the boundary superpixels. Based on the (5), define the difference measure between image superpixels located at the image boundary and the remaining superpixels as follows:

$$\psi(v, B) = \min_{v \notin B, j \in B} mer(v, j) \tag{6}$$

If the superpixel v satisfies the conditions $\psi_(v, B)$, it can be extended to the background region. Defined $\psi_(v, B)$ as:

$$\psi(v, B) < \frac{1}{N}\sum_{i=1}^{N} mer(i, B) \tag{7}$$

Where N means the number of superpixels not including the image boundary.

We use boundary extensions to highlight foreground regions, which are selected as queries. Image background regions composed of extended regions and boundary regions are treated as unlabeled nodes. The other Manifold ranking algorithm is used to rank the correlation between the foreground regions and image background. After manifold ranking, we can obtain a saliency map s_f:

$$s_f(i) = \bar{f}_f(i) \tag{8}$$

Where $\bar{f}_f$ represents a normalized vector with a range of values from 0 to 1.

Some examples of boundary extension and foreground prior saliency maps are presented in Fig. 3. Most of background regions are covered via the boundary extensions,as is shown in Fig. 3(e). The regions marked as blue color in Fig. 3(g) are the foreground regions after boundary extension and boundary regions are removed from the image. Compared with the ground truth(Fig. 3(f)), boundary extension can highlight the foreground regions to a certain extent. After manifold ranking, we achieve a foreground prior saliency map that is presented in 3(f).

Fig. 3. Examples of foreground prior saliency.(a)input image,(b)image superpixels,(c)image boundary set(regions marked as blue color),(d)boundary extension regions(regions marked as blue color),(e)foreground regions(regions marked as green color),(f)foreground prior saliency map,(g)ground truth

2) Center prior saliency measure based on corners clustering: By using Moravec algorithm for reference, Harris corner detection algorithm is developed, and its core viewpoint is shown in [11]. In our paper, the improved Harris corner detection [12] is used as a corner detection algorithm to detect image corner points. Harris corner detector can detect multiple corners that may be scattered anywhere in the image. If all corners are used, the calculation is large and the efficiency is low. In order to effectively make a saliency measure, the foreground area obtained by the boundary extension are used to reduce the number of corner points, and highlight the corners on the foreground regions.

Use the boundary extension to highlight foreground regions, which can be used to filter image corners. Although there are still some corners on the background regions or some corners that belong to the foreground area are deleted after corner filtering, these situations can be ignored. In this way, most of the corners on the foreground regions are obtained and denoted as (a_i, b_i),$i = 1, 2, \cdots, k$. K-means clustering is utilized to cluster them into one corner 0with a center of (a_0, b_0). Finally, a two-dimensional Gaussian model with a center point as a peak value is established to compute center prior saliency value $\bar{S}_c$:

$$\bar{S}_c(i) = \exp(-\frac{(x(i) - a_0)^2}{2\sigma_x^2} - \frac{(y(i) - b_0)^2}{2\sigma_y^2}) \tag{9}$$

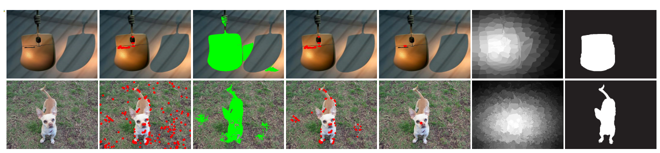
Where x_i and y_i is the center coordianates of superpixel i. σ_x and σ_y are horizontal and vertical variances of images, respectively.

In addition, by judging the number of corners, the saliency value is corrected as follows:

$$S_c = \begin{cases} \bar{s}_c & k \geq m \\ 1 & k < m \end{cases} \tag{10}$$

Where k is the number of remaining corners. m is a threshold indicating the number of corner points that determine whether to use the Gaussian model.

Fig. 4 shows center prior saliency results. Corners on the image are detected by Harris corner detector, which are marked as red color. Corners on the image are presented in Fig. 4(b), Fig. 4(d) and Fig. 4(e). Compared with Fig. 4(b) and Fig. 4(d), the number of corners is reduced to some corners on the foreground regions. Due to the background noise in the foreground regions, a small part of the corner points fall on the background regions. In all, number of corners can be reduced by screening whether corners are located at the foreground regions (Fig. 4(c)). After corners clustering, one corner pointer fall in the foreground This corner(Fig. 4e), which is regarded as a center point for Gaussian calculation. From the Fig. 4(f), we can see that it can locate the salient regions.



(a) (b) (c) (d) (e) (f) (g)

Fig. 4. Some samples of center prior saliency results.(a)input image.(b)corners(points marked in red color).(c)foreground regions(regions marked in green color),(d)corners screening(points marked in red color),(e)corners clustering(one point marked in red color).(f)center prior saliency map.(g)ground truth

D. Fusion

In the final state of our proposed method, we propose a fusion framework to fuse three kinds of saliency maps that are measured by the background prior, foreground prior and center prior respectively. Although they can highlight salient regions at a certain. they are fused by our proposed fusion framework to gain a final saliency map with better performance $S(i)$:

$$S(i) = S_f(i) + S_f(i) \cdot (1 - \exp(-S_b(i) \cdot S_c(i))) \tag{11}$$

III. EXPERIMENT

In the section, we verify the performance of our proposed method via comparison with the classic or representative methods on two benchmark datasets: ECSSD [13], MSRA10K [14]. There are 1000 images with complex background on ECSSD dataset, and 10000 images on MSRA10K dataset. The fifteen methods we compare are IT98 [4], SR07 [5], FT09 [6], SEG10 [15], MSS10 [16], CA10 [17], FES11 [8], HC11 [18], SF12 [7], PCA13 [19],GMR13 [20], GC13 [21], MBD15 [22], LPS15 [23], RCRR18 [24]. For a fair comparison, we use a salient map that was obtained from the author or generated by the author's code using default parameters.

A. Evaluation metrics

The performance of our proposed method are verified using a standard precision-recall curve. Saliency map is binarized using a threshold between 0 and 255 and set a threshold for every 5 values to obtain a series of binary plots, then calculate a binary plot against the ground truth. An effecti

ve method requires not only higher precision values, but also higher recall values, so the combination of our calculation accuracy and recall rate(called F-measure) is defined as:

$$F_\beta = \frac{(1+\beta^2) \cdot precision \cdot recall}{\beta^2 \cdot precision + recall} \tag{12}$$

Where the value of β^2 is a constant, which is set to 0.3 in [25] to put emphasis on the precision.

B. Performance evaluation

We compare our proposed method with fifteen methods based on the performance evaluation metrics of precision-recall curves , F-measure on ECSSD and MSRA10K. Fig. 5 and Fig.6 shows the experimental results on ECSSD datset and MSRA10K dataset, respectively.

As can be seen from Fig. 5, precision-recall curves that obtained by our proposed method achieve better performance. The precision of our method is only at LPS, inferior to RCRR and MBD. But, the recall achieves the best results. For the F-measure, our method achieves the best results. Based on these results, our proposed method can achieve better performance on ECSSD.

As can be seen from Fig. 6, the precision-recall curves that obtained by our method cover most of the candidate methods, while F-measure is in the second place. Combining all these facts, our proposed method can achieve better performance on MSRA10K.

To further verify the performance of our method, we have listed some visual comparison results, which are presented in Fig.7. The saliency map that obtained by our proposed method is visually similar to the ground truth. It has strong robustness in complex backgrounds, and can highlight the outline of the salient object. Experimental results show that the three kinds of prior-based saliency maps can achieve better performance after fusion by our fusion framework.

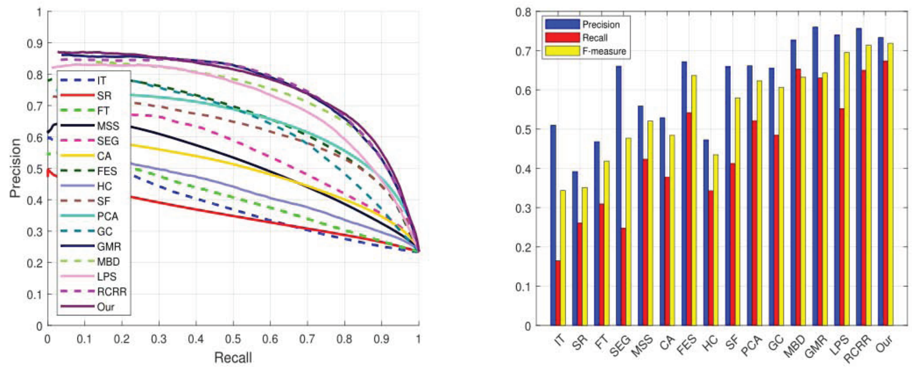


Fig. 5. Precision-recall curves and F-measure comapring with different state-of-the-art methods on ECSSD dataset

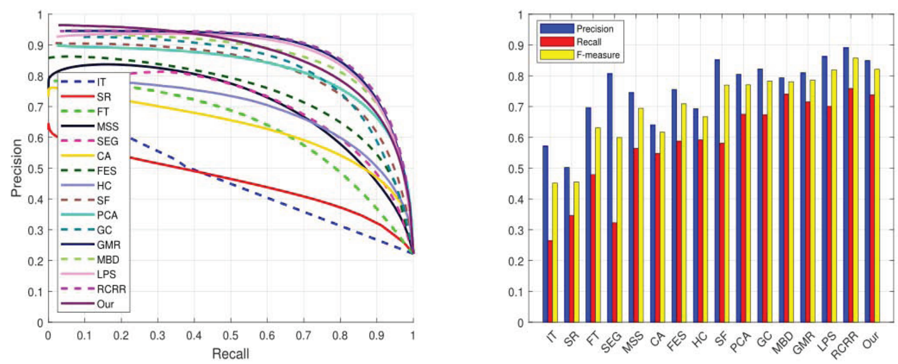


Fig. 6. Precision-recall curves and F-measure comapring with different state-of-the-art methods on MSRA10K dataset

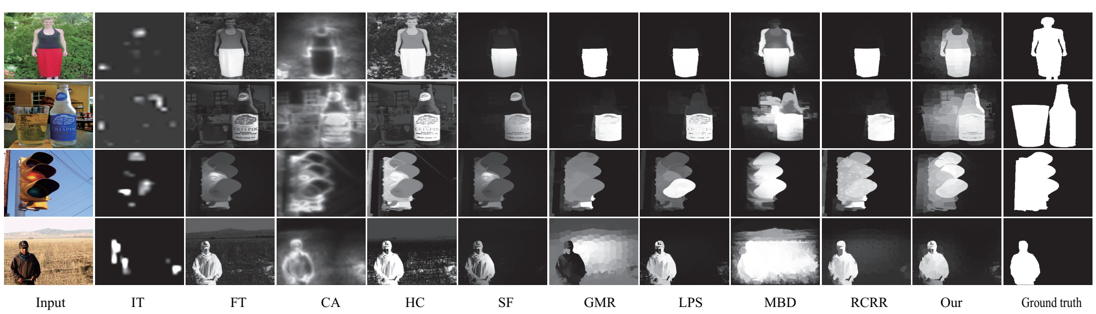


Fig. 7. Sample saliency maps of the compared methods

IV. Conclusion

A saliency detection method based on a variety of prior information fusions are proposed in this paper. The prior information includes background prior, foreground prior, and center prior. Firstly, by using each boundary of the image as the background prior information, we obtain the background prior saliency map by the four main saliency maps through the manifold ranking algorithm. Secondly, we use the boundary expansion to highlight the foreground regions that are regarded as the foreground prior information, and use the manifold ranking to obtain the foreground prior saliency map. Thirdly, the corner detector is used to obtain the corners on the image, and obtaining the center prior saliency map through screening, clustering, and Gaussian model calculation.Finally, the above three kinds of prior-based saliency maps are fused via the proposed fusion framework to gain a better final saliency map. Compared with the other fifteen methods on ECSSD and MSRA10K, Experimental results show that our proposed method achieve better saliency detection results. In the future, we will utilize top-down method to achieve a higher accuracy location and object integrity of salient object detection form supervised learning method.

REFERENCES

[1] Y.Fang ,Z. Chen ,W. Lin , et al, Saliency detection in the compressed domain for adaptive image retargeting, IEEE Transactions on Image Processing.21(9)(2012)3888-3901.

[2] Z. Ren, S. Gao ,L. T. Chia , et al, Region-based saliency detection and its application in object recognition, IEEE Transactions on Circuits and Systems for Video Technology. 24(5)(2014) 769-779.

[3] J. Linhua ,Z. Hui ,L. Xiao , Saliency detection via boundary prior and center prior, International Robotics& Automation Journal. 2(4)(2017)2-8.

[4] L.Itti , C. Koch,E. Niebur , A model of saliency-based visual attention for rapid scene analysis, IEEE Transactions on Pattern Analysis & Machine Intelligence.(11)(1998) 1254-1259.

[5] X. Hou,L. Zhang , Saliency detection: A spectral residual approach, in: 2007 IEEE Conference on computer vision and pattern recognition. IEEE, 2007.pp. 1-8.

[6] R. Achanta ,S. Hemami ,F Estrada , et al, Frequency-tuned salient region detection,in: IEEE international conference on computer vision and pattern recognition , 2009 .pp.1597-1604.

[7] A.Hornung ,Y. Pritch ,P. Krahenbuhl ,Saliency filters: Contrast based filtering for salient region detection,in: IEEE Conference on Computer Vision and Pattern Recognition, USA, 2012.pp. 733-740.

[8] M. M.Cheng ,G. X. Zhang ,N. J. Mitra , et al, Global Contrast Based Salient Region Detection, Computer Vision and Pattern Recognition. 2011.

[9] D. Zhou. Learning with Local and Global Consistency, NIPS, 2004, 16.

[10] J. Tang,G.Yang,et al. Salient object detection of dairy goats in farm image based on background and foreground priors[J]. Neurocomputing. 332(2019)270-282.

[11] C. G.Harris,M. Stephens , A combined corner and edge detector,Alvey vision conference. 15(50)(1988)10-5244.

[12] X.Zhang ,B. Li ,D. Yang , A novel harris multi-scale corner detection algorithm, Journal of Electronics & Information Technology. 29(7)(2007)1735-1738.

[13] Q. Yan ,L. Xu ,J. Shi , et al, Hierarchical saliency detection,in:Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013.pp.1155-1162.

[14] A. Borji ,M. M. Cheng ,Q. Hou , et al. Salient Object Detection: A Survey[J]. Eprint Arxiv, 16(7)(2014),3118.

[15] E.Rahtu ,J. Kannala , M. Salo, et al, Segmenting salient objects from images and videos,in:European conference on computer vision. Springer, Berlin, Heidelberg, 2010.pp. 366-379.

[16] R. Achanta,S, Ssstrunk, Saliency detection using maximum symmetric surround in:IEEE International Conference on Image Processing, IEEE, 2010.pp. 2653-2656.

[17] S.Goferman ,L. Zelnikmanor , A. Tal,Context-Aware Saliency Detection, IEEE Trans Pattern Anal Mach Intell. 34(10)(2012)1915-1926.

[18] M. M.Cheng,N. J. Mitra ,X. Huang , et al, Global contrast based salient region detection, IEEE Transactions on Pattern Analysis and Machine Intelligence.37(3)(2015)569-582.

[19] R.Margolin ,A. Tal ,L. Zelnik-Manor , What makes a patch distinct? In: Computer Vision and Pattern Recognition (CVPR),2013. pp. 1139-1146.

[20] C.Yang,L. Zhang,H. Lu, et al, Saliency detection via a graph-based manifold ranking,in:Proceedings of the IEEE conference on computer vision and pattern recognition. 2013.pp. 3166-3173.

[21] Y.Wei ,F. Wen ,W. Zhu ,et al, Geodesic Saliency Using Background Priors,in: European Conference on Computer Vision, Italy, 2012.pp. 29-42.

[22] J.Zhang ,S. Sclaroff ,Z. Lin , et al, Minimum barrier salient object detection at 80 fps,in: Proceedings of the IEEE international conference on computer vision, 2015.pp. 1404-1412.

[23] H. Li ,H. Lu ,Z. Lin , et al, Inner and Inter Label Propagation: Salient Object Detection in the Wild, IEEE Transactions on Image Processing.24(10)(2015)3176-3186.

[24] Y. Yuan ,C. Li ,J. Kim , et al, Reversion correction and regularized random walk ranking for saliency detection, IEEE Transactions on Image Processing. 27(3)(2018) 1311-1322.

[25] F. Perazzi, et al. Saliency filters: Contrast based filtering for salient region detection, in: 2012 IEEE conference on computer vision and pattern recognition. IEEE, 2012.pp. 733-740.