Robust Saliency Detection via Regularized Random Walks Ranking

Changyang Li[1], Yuchen Yuan[1], Weidong Cai[1], Yong Xia[2], and David Dagan Feng[1]
[1]The University of Sydney, [2]Northwestern Polytechnical University

Saliency detection refers to the exploration and localization of the region in a given image that attracts most human attention when presented [4]. In computer vision, saliency detection algorithms are usually categorized into bottom-up [1] and top-down [3] approaches, in which graph-based bottom-up approaches prevails the current research . However, many graph-based bottom-up saliency detection algorithms heavily depend on the boundary prior and the pre-processed superpixel segmentation, which in general is of low robustness, and leads to significant sacrifice of detail information from the original image.

In this paper, we propose a novel saliency detection method, which takes both region-based estimations and pixel-wise image details into account. Our innovations come in two aspects: on the one hand, we suggest the erroneous boundary removal, which filters out one of the four boundaries that most unlikely belonging to the background; on the other hand, we propose the regularized random walks ranking, which is independent of the superpixel segmentation, and can generate pixel-wised saliency maps that reflects full details of the image. The workflow of our algorithm is introduced as follows.

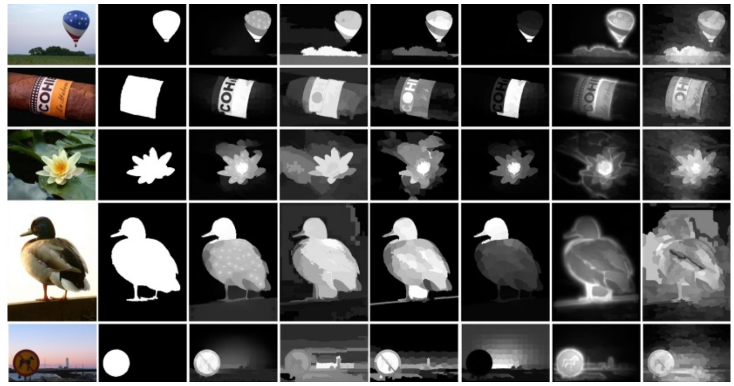
Firstly, we conduct the erroneous boundary removal. The input image is segmented into superpixels by the SLIC method [2], and all of the superpixels on each boundary are treated as a connected region. The normalized RGB histogram of each boundary is calculated respectively afterwards. We then compute the Euclidean distance of any two of the four boundary histograms. If part of the salient object is boundary-adjacent, such boundary is expected to be significantly different to the others; thus the boundary which has the largest histogram difference to the other three is determined as the erroneous boundary, and will be removed from the following saliency estimation.

Secondly, we conduct the saliency estimation with the graph-based manifold ranking [9], which consists of both background saliency estimation and foreground saliency estimation. However, only the three remaining boundaries after the erroneous boundary removal will be used in the background saliency estimation. The saliency estimation result will be generated as $S(i), i=1,...,n,$ where n is the total superpixel number.

Finally, we conduct the proposed regularized random walks ranking, which is extended from the random walks model. The image is treated as a dataset $\chi=\{x_1,...,x_n\}\in\mathbb{R}^m$, where n is the total pixel number. We first mark s elements from χ as the seed nodes. Without loss of generality, we assume that the first s elements of χ are the seeds, so that $\chi=[x_M^T,x_U^T]$, in which x_M are the seed nodes and x_U are the unseeded nodes. We define the weight matrix W, degree matrix D, and the Laplacian matrix L similarly to [9]; then let $p^k=\left[p_1^k,\dots,p_n^k\right]^T$ denote the probability vector of χ for label k, where $k=1,2$ stand for background/foreground, respectively. p^k can thus be partitioned as $p^k=\left[\left(p_M^k\right)^T,\left(p_U^k\right)^T\right]$, where p_M^k is for the seed nodes, having fixed value as 1 for the corresponding label. The optimized p^k is achieved by minimizing our modified Dirichlet integral,

$$Dir\left[p^k\right]=\frac{1}{2}\left(p^k\right)^T L\left(p^k\right)+\frac{\mu}{2}(p^k-Y)^T(p^k-Y)= \\ \frac{1}{2}\left[\left(p_M^k\right)^T\left(p_U^k\right)^T\right]\begin{bmatrix}L_M & B\\ B^T & L_U\end{bmatrix}\begin{bmatrix}p_M^k\\ p_U^k\end{bmatrix}+\frac{\mu}{2}\left(\begin{bmatrix}p_M^k\\ p_U^k\end{bmatrix}-\begin{bmatrix}Y_M^k\\ Y_U^k\end{bmatrix}\right)^T\left(\begin{bmatrix}p_M^k\\ p_U^k\end{bmatrix}-\begin{bmatrix}Y_M^k\\ Y_U^k\end{bmatrix}\right). \tag{1}$$

where μ is a controlling parameter, and Y is a pixel-wise indication vector inheriting the values of S. We define two thresholds t_{high} and t_{low},



Images GT Ours CB[5] GS[8] MR[9] PCA[6] LR[7]
Figure 1: Example saliency maps of state-of-the-art methods and our proposed method.

$$t_{high}=\frac{\text{mean}(S)+\max(S)}{2},\ t_{low}=\text{mean}(S), \tag{2}$$

to select seed pixels for p_M^k, with $Y>t_{high}$ as foreground seeds and $Y<t_{low}$ as background seeds. The optimized solution is obtained as,

$$p_U^k=\left(L_U+\mu I\right)^{-1}\left(-B^T p_M^k+\mu Y_U^k\right). \tag{3}$$

p_U^k and p_M^k are then combined to form p^k. We set $k=2$ to select the foreground possibility as the final foreground saliency output.

Figure 1 shows the effect of the proposed method in comparison with several state-of-the-art saliency detection methods. The details of our method are described in the full paper. The conclusion is that our method outperforms 12 state-of-the-art methods on two public datasets, in terms of both accuracy and robustness.

[1] R. Achanta, F. Estrada, *et al.*, "Salient region detection and segmentation," in *Computer Vision Systems*, ed: Springer, 2008, pp. 66.

[2] R. Achanta, A. Shaji, *et al.*, "SLIC superpixels compared to state-of-the-art superpixel methods," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 34, pp. 2274, 2012.

[3] A. Borji, D. N. Sihite, *et al.*, "Probabilistic learning of task-specific visual attention," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 470.

[4] L. Itti, C. Koch, *et al.*, "A model of saliency-based visual attention for rapid scene analysis," *IEEE transactions on pattern analysis and machine intelligence,* vol. 20, pp. 1254, 1998.

[5] H. Jiang, J. Wang, *et al.*, "Automatic salient object segmentation based on context and shape prior," in *BMVC*, 2011, p. 7.

[6] R. Margolin, A. Tal, *et al.*, "What Makes a Patch Distinct?," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 1139.

[7] X. Shen and Y. Wu, "A unified approach to salient object detection via low rank matrix recovery," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 853.

[8] Y. Wei, F. Wen, *et al.*, "Geodesic saliency using background priors," in *Computer Vision–ECCV 2012*, ed: Springer, 2012, pp. 29.

[9] C. Yang, L. Zhang, *et al.*, "Saliency detection via graph-based manifold ranking," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 3166.

This is an extended abstract. The full paper is available at the CVF webpage (http://www.cv-foundation.org/openaccess/CVPR2015.py).