

# Contrast based Saliency Detection via Manifold Ranking

Shunyao Jin, Xueqing Li

Shandong University

**Keywords:** saliency detection, manifold ranking, superpixel segmentation, foreground region

**Abstract.** Saliency detection has been widely studied in computer vision. In this paper we propose a two-steps method combining contrast assumption and ranking technology to detect saliency region. Firstly, We use the traditional contrast assumption to find foreground cues. Then we rank the nodes with labeled contrast cues by graph-based manifold ranking. We experiment with our method on a large public data set. Our results show the effectiveness of our method, and perform better compared to recent state-of-the-art methods.

## INTRODUCTION

In computer vision, salient region detection is a long-standing problem. It measures the quality of the region standing out to its neighbors. Usually, salient regions are those attracting the most attention when observers glance the scene. It is a complex problem identifying these regions without prior knowledge, and it lays in the interdisciplinary of computer science, biology, psychology and other fields. Hundreds of computational saliency models that following early attention models (e.g., Koch and Ullman [1] and Itti et al. [2]) have been proposed to detect salient region from images or videos. And the results of the saliency detection have been successfully applied in many areas, such as image segmentation [3], object detection [4], image resizing [5] and Image compression [6].

In general saliency methods are categorized as bottom-up [7,8,9] and top-down [10,11] approaches: the bottom-up approaches use low-level features like color, intensity and orientation from pixels or segmented parts and the top-down approaches focus on mid-level or high-level cues such as face detection. In bottom-up salient object detection tasks, most of the methods [21,22] focus on the center-surround contrast assumption, which is that appearance contrasts between salient region and their surrounding regions are high.

In this paper, firstly our method is rely on the fact that an image can be decomposed into basic elements that remove unnecessary detail, and still a clear definition of contrast-based saliency can be found. At the same time, as the nodes reducing, efficiency can improve a lot. So, in practice an input image is decomposed into small patches by superpixel segmentation in our method.

Then we propose a two-stage method for bottom-up saliency detection. In the first stage, we use contrast assumption to find the foreground regions. Image patches are connected in a fully-connect directed graph. A patch ranks high only if it stands out from its neighbors in the image. In the second stage, the saliency results from the first step can be roughly segmented into background and foreground region. We take the labeled foreground nodes as ranking queries. In practice, manifold ranking is used to rank the queries. Based on relevance to queries, the final saliency of each node can be computed.

We test our method on one public available benchmark data sets [12], and the result shows a great improvement comparing to the original contrast-based algorithms. Our method also generates comparable result with the state-of-the-art methods [2,13,14,15,16].

## ALGORITHM

Our method composes of three steps (see Figure 1): first, an input image is decomposed into small patches by segmentation; second a fully connected directed graph is formed to find foreground regions by using contrast assumption. Finally, we take the labelled foreground nodes as salient queries in manifold ranking to get the final saliency map.

## 1.1 Segmentation

Supapixel is the most popular choice in saliency detection algorithms. One reason is that computational elements can be greatly reduced; the other is that it can preserve the object boundary and remove the noisy, which most saliency detection can benefit from.

In our method, the input image is first decomposed into uniform regions in CIELab space by SLIC. The color and spatial feature of each region is computed by calculating mean colour and position of every region.

## 1.2 Contrast Map

Most of saliency detection methods are based on contrast assumption. This assumption, both local and global contrast, measures the rarity of each patch and has been the basis for most previous algorithms. It works well in most cases. Although it has disadvantage in accuracy, it is a simple and fast way to find the salient seed for ranking.

The contrast  $C_{i,j}$  between two patches is defined:

$$C_{i,j} = \text{Exp}\left(-\frac{D_{i,j}}{\sigma}\right) * W_{i,j}$$

Where  $W_{i,j}$  is defined as the L2 distance in CIELab color space between two patches and  $D_{i,j}$  is the L2 distance between the centers of two patches. So we can effectively combine global and local contrast estimation by considering the actual Euclidean distance between two patches in the map. Based on the contrast definition, we form a fully connected directed graph  $G$  by connecting every two patches. We treat the patches as vertices and the weight  $C_{i,j}$  is assigned for each edge connecting vertices  $i$  and  $j$ .

In previous work, there are two main methods dealing with the fully-connected directed graph  $G$ : in [6, 9], the saliency value  $R$  is defined as sum of income from each edge; in [13, 14], the full-connected graph is taken as the Markov chain, and the transition probability is used to simulate human observers eye movement. Here we adapt the first method because it is simple and quick.

So the saliency value  $S$  can be defined as:

$$S_i = \sum_{j=1}^n C_{j,i}$$

As we consider the contrast as the weight between the nodes, the larger the value of  $S$  is, the more possibility it stands for foreground region. So we first normalize  $S$  within  $[0,1]$ , then we use the value  $S$  to construct a binary map. Then the positive nodes that are larger than twice of the mean saliency are labeled foreground nodes in the next step.

## 1.3 Manifold Ranking

The manifold ranking algorithm [19] is initially proposed to rank text or image data. Different from other rank algorithm like page ranking, manifold ranking respect intrinsic manifold structure which is import to saliency detection.

Images might not have such manifold structure. So we construct a weighted graph by using every superpixel point as a vertex and defining weighted edge between them. All superpixel point will be ranked according to the query, a large score meaning high relevance.

The ranking algorithm is defined:

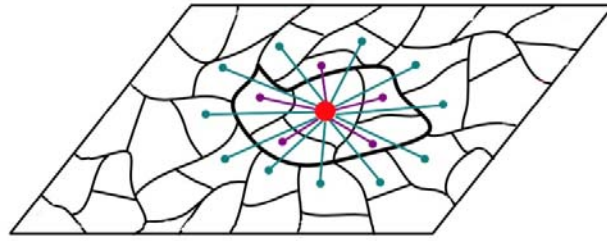


Figure 1

Given a set of point  $V = \{v_1 \dots v_q, v_{q+1} \dots v_n\}$ , the first  $q$  points belong to the queries and the rests need to be ranked. Let  $d : V \times V \rightarrow \mathbb{R}$  be a metric on  $V$  which assigns a distance  $d(v_i, v_j)$  to each pair of points  $v_i$  and  $v_j$ . Let  $f : V \rightarrow \mathbb{R}$  be a ranking function. Every  $v_i$  can be assigned a ranking value  $f_i$  by this function.  $f$  can be viewed as a vector  $f = [f_1 \dots f_n]^T$ . Then we define a vector  $y = [y_1, \dots, y_n]^T$ , if  $x_i$  is a query then  $y_i = 1$ , if  $v_i$  is not a query then  $y_i = 0$ .

We construct a graph  $G = (V, E)$ . In this graph,  $E$  is a set of undirected edges and  $V$  is a set of nodes. Each node is connected to the nodes that have common boundaries. Thus, it also has connection with the nodes surrounding its neighboring node (see in Figure 2).

The distance between two nodes can be defined:

$$d_{i,j} = \exp(-|c_i - c_j|/\sigma^2)$$

Where  $|c_i - c_j|$  is the L2 distance between two patches in CIE Lab color space, and  $\sigma^2$  is constant. The calculation has been shown to be effective in saliency detection [17, 18].

The algorithm [19] is as follows:

1. Sort the pairwise distances among points in ascending order. Repeat connecting the two points with an edge according the order until a connected graph is obtained.
2. Form the affinity matrix  $W$  defined by  $W_{i,j} = \exp[-d_{i,j}^2/2\sigma^2]$  if there is an edge linking  $v_i$  and  $v_j$ . Note that  $W_{ii} = 0$  because there are no loops in the graph.
3. Symmetrically normalize  $W$  by  $S = D^{-1/2}WD^{-1/2}$  in which  $D$  is the diagonal matrix with  $(i, i)$ -element equal to the sum of the  $i$ th row of  $W$ .
4. Iterate  $f(t+1) = \alpha Sf(t) + (1-\alpha)y$  until convergence, where  $\alpha$  is a parameter in  $[0, 1)$ .
5. Let  $f_i^*$  denote the limit of the sequence  $\{f_i(t)\}$ . Rank each point  $x_i$  according its ranking scores  $f_i^*$  (largest ranked first).

We use the results obtained in first step as the queries and the rest as the unlabeled data. In the end, we can get the final saliency map by the manifold ranking algorithm with queries. The ranking scores represent the salient possibility.

## EXPERIMENTAL RESULTS

We evaluated detection method in a benchmark data-set provided by Achanta et al. which contains 1000 color images with segmentation ground truth [11].

We compare our experiments with several state-of-art methods. The methods for comparison include three classical methods IT [2], SR [13], GB [21]. Three recent state-of-art methods including HC [15], RC [15] and SF [16] are also considered.

For each image, first we compare binary masks for every threshold in the range  $[0..255]$ . Then we employ the adaptive threshold as twice of the mean saliency, F-measure where  $F = ((1 + \beta^2)P * R) / (\beta^2 P + R)$  ( $P$ =Precision,  $R$ =Recall) is computed for evaluation (See Figure. 3)

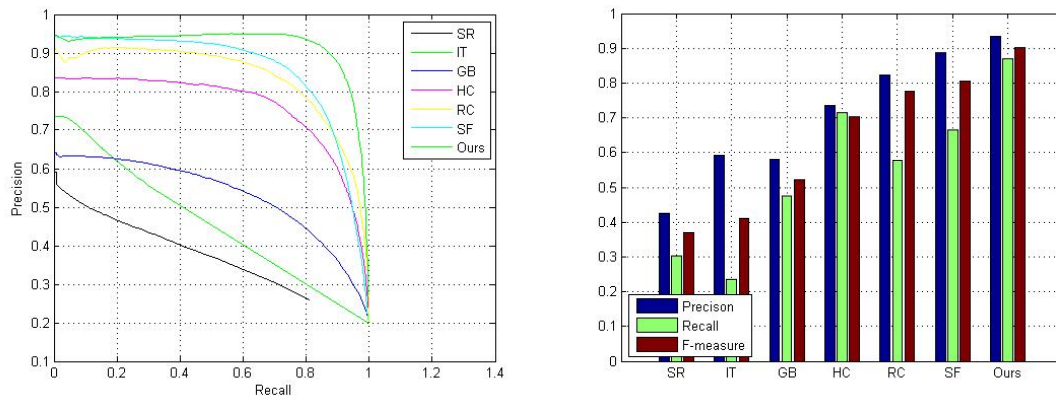


Figure 2

## CONCLUSION

In this paper, we propose we propose a two-steps method combining contrast assumption and ranking technology to detect saliency region. Experimental results demonstrate better saliency detection performance comparing to several state-of-art methods. We will try to improve the accuracy and efficiency of our current implementations in the future.

## REFERENCES

- [1] C. Koch and S. Ullman. Shifts in selective visual attention:towards the underlying neural circuitry. *Human Neurbiology*,4:219–227, 1985.
- [2] L. Itti, C. Koch, and E. Niebur. A model of saliency-basedvisual attention for rapid scene analysis. *IEEE TPAMI*,20(11):1254–1259, 1998.
- [3] Ueli Rutishauser, Dirk Walther, Christof Koch, andPietro Perona, “Is bottom-up attention useful for objectrecognition?,” in *Computer Vision and Pattern Recognition,2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. IEEE, 2004, vol.2
- [4] Kossinets, Gueorgi, and Duncan J. Watts. “Origins of Homophily in an Evolving Social Network.” *American Journal of Sociology* 115 (2009): 405-50. Accessed February 28, 2010. doi:10.1086/599247
- [5] Shai Avidan and Ariel Shamir, “Seam carving forcontent-aware image resizing,” in *ACM Transaactionson graphics (TOG)*. ACM, 2007, vol. 26,p.10
- [6] Charilaos Christopoulos, Athanassios Skodras, andTouradj Ebrahimi, “The jpeg2000 still image codingsystem: an overview,” *Consumer Electronics, IEEETransactions on*, vol. 46, no. 4, pp. 1103–1127, 2000.
- [7] L. Wang, J. Xue, N. Zheng, and G. Hua. Automatic salient objectextraction with contextual cue. In *ICCV*, 2011. 1
- [8] W. Wang, Y. Wang, Q. Huang, and W. Gao. Measuring visualsaliency by site entropy rate. In *CVPR*, 2010. 1
- [9] Y. L. Xie, H. C. Lu, and M. H. Yang. Bayesian saliency via low andmid level cues. *IEEE TIP*, 2013. 1, 6
- [10] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum. Learning to detect a salient object. *IEEE PAMI*, 2011. 1, 5
- [11] J. Yang and M. Yang. Top-down visual saliency via joint crf anddictionary learning. In *CVPR*, 2012. 1

- [12] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In CVPR, 2009. 1, 4, 5, 6, 7
- [13] Xiaodi Hou, Liqing Zhang, “Saliency detection: Spectral residual approach,” in Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on. IEEE, 2007, pp. 1–8.
- [14] Jonathan Harel, Christof Koch, and Pietro Perona, “Graph-based visual saliency,” in Advances in neural information processing systems, 2006, pp. 545–552.
- [15] Ming-Ming Cheng, Guo-Xin Zhang, Niloy J Mitra, Xiao-Lei Huang, and Shi-Min Hu, “Global contrast based salient region detection,” in Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011, pp. 409–416.
- [16] Federico Perazzi, Philipp Krahenbuhl, Yael Pritch, and Alexander Hornung, “Saliency filters: Contrast based filtering for salient region detection,” in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE
- [17] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In CVPR, 2009. 1, 4, 5, 6, 7 □
- [18] A. Borji, D. Sihite, and L. Itti. Salient object detection: A benchmark. In ECCV, 2012. 4, 6 □
- [19] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Scholkopf. Ranking on data manifolds. In NIPS, 2004. 2, 3 □
- [20] Jonathan Harel, Christof Koch, and Pietro Perona, “Graph-based visual saliency,” in Advances in neural information processing systems, 2006, pp. 545–552. □
- [21] Y. Tian, J. Li, S. Yu, and T. Huang, “Learning complementary saliency priors for foreground object segmentation in complex scenes,” IJCV, 2014.
- [22] P. Jiang, H. Ling, J. Yu, and J. Peng, “Salient region detection by ufo: Uniqueness, focusness and objectness,” in ICCV, 2013.