International Journal of Advanced Research in Science Management and Technology

ISSN(Online)-2454-4159 Volume 3, Issue 4, April 2017

# "A Review of Saliency Detection via Graph-based Manifold Ranking"

**Swati R.Patil[1], Prof. S.P.Patil [2]**

PG Scholar, Department of E& TC, D. N. Patel College of Engineering, Shahada, Maharashtra, India[1]

Associate Professor, Department of E& TC, D. N. Patel College of Engineering, Shahada, Maharashtra, India[2]

***Abstract***— Most existing bottom-up methods measure the foreground saliency of a pixel or region based on its contrast within a local context or the entire image, whereas a few methods focus on segmenting out background regions and thereby salient objects. Instead of considering the contrast between the salient objects and their surrounding regions, we consider both foreground and background cues in a different way. We rank the similarity of the image elements (pixels or regions) with foreground cues or background cues via graph-based manifold ranking. The saliency of the image elements is defined based on their relevances to the given seeds or queries. We represent the image as a close-loop graph with superpixels as nodes. These nodes are ranked based on the similarity to background and foreground queries, based on affinity matrices. Saliency detection is carried out in a two-stage scheme to extract background regions and foreground salient objects efficiently. Experimental results on database demonstrate the proposed method performs well when against the state-of-the-art methods in terms of accuracy and speed.

***Keywords—*** Foreground saliency, graph-based manifold ranking, a close-loop graph, foreground and background cues.

## I. INTRODUCTION

The task of saliency discovery is to recognize the most important and informative part of a scene. It has been applied to various vision issues including image segmentation, protest acknowledgment, image pressure, content based image retrieval [8], to name a couple. Saliency techniques in general can be categorized as either base up or best down approaches. Base up techniques are data-driven and pre-attentive, while beat down strategies are task driven that entails managed learning with class labels. We take note of that saliency models have been created for eye fixation expectation and salient protest recognition. The previous concentrates on recognizing a couple of human fixation locations on natural images, which is important for understanding human attention. The latter is to accurately identify where the salient protest ought to be, which is helpful for many abnormal state vision tasks.

The main observation is that the distance between a pair of background districts is shorter than that of an area from the salient protest and a locale from the background. The hub labeling task (either salient protest or background) is formulated as a vitality minimization issue based on this criteria. We watch that background regularly shows local or global appearance availability with each of four image boundaries and foreground presents appearance rationality and consistency. In this work, we abuse these cues to figure pixel saliency based on the ranking of superpixels. For each image, we develop a close-loop graph where each hub is a superpixel. We demonstrate saliency location as a manifold ranking issue and propose a two-stage conspire for graph labeling. Figure 1 demonstrates the main strides of the proposed algorithm. In the principal stage, we misuse the boundary earlier [13, 22] by utilizing the hubs on each side of image as labeled background questions. From each labeled outcome, we figure the saliency of hubs based on their relevances (i.e, ranking) to those questions as background labels. The four labeled maps are then integrated to generate a saliency map. In the second stage, we apply binary segmentation on the came about saliency map from the primary stage, and take the labeled foreground hubs as salient questions. The saliency of each hub is processed based on its relevance to foreground questions for the final map. To completely capture inborn graph structure information and incorporate local gathering cues in graph labeling, we utilize manifold ranking procedures to learn a ranking capacity, which is essential to learn an optimal affinity matrix [20]. Unique in relation to [12], the proposed saliency recognition algorithm with manifold ranking requires just seeds from one class, which are initialized with either the boundary priors or foreground cues. The boundary priors are proposed propelled on the current works of human fixations on images [31], which demonstrates that humans tend to gaze at the focal point of images. These priors have also been utilized as a part of image segmentation and related issues [13, 22, 34]. In contrast, the semi-directed technique [12] requires both background and salient seeds, and generates a binary segmentation. Besides, it is hard to decide the number and locations of salient seeds as they are generated by random walks, especially for the scenes with various salient articles. This is a known issue with graph labeling where the outcomes are delicate to the chosen seeds. In this work, all the background and foreground seeds can be easily generated via background priors and ranking background questions (or seeds). As our model incorporates local gathering cues extracted from the whole image, the proposed algorithm generates very much characterized boundaries of salient articles and consistently highlights the entire salient areas. Experimental outcomes utilizing data sets demonstrate that the proposed algorithm performs effectively and favorably against the state-of-the-art saliency identification strategies.

## II.LITERATURE SURVEY

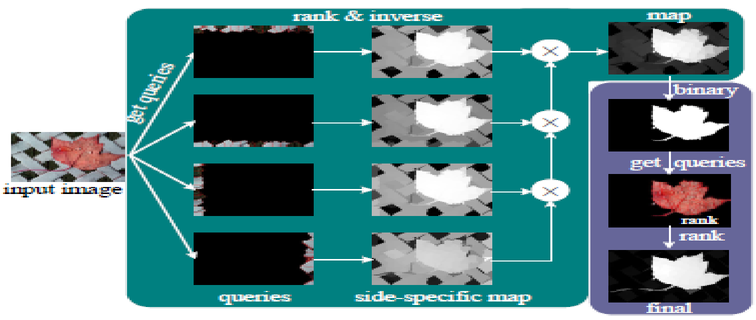
In this paper, we concentrate on the base up salient question recognition tasks. Salient protest identification algorithms usually generate bouncing boxes, binary foreground and background segmentation, or saliency maps which indicate the saliency probability of each pixel. Liu et al. [23] propose a binary saliency estimation demonstrate via training a conditional random field to join an arrangement of novel features. Wang et al. [32] analyze various cues in a brought together vitality minimization framework and utilize a graph-based saliency show [14] to recognize salient articles. In [24] Lu et al. build up a hierarchical graph demonstrate and use concavity setting to register weights between hubs, from which the graph is bi-partitioned for salient question discovery. Then again, Achanta et al. [1] register the saliency probability of each pixel based on its shading contrast to the whole image. Cheng et al. [9] consider the global district contrast as for the whole image and spatial relationships across the locales to extract saliency map. In [11] Goferman et al. propose a setting aware saliency algorithm to identify the image areas that speak to the scene based on four standards of human visual attention. The contrast of the inside and encompass circulation of features is processed based on the Kullback-Leibler disparity for salient question identification [21]. Xie et al. [35] propose a novel model for base up saliency inside the Bayesian framework by misusing low and mid level cues. Sun et al. [30] enhance the Xie's model by presenting boundary and delicate segmentation. As of late, Perazzi et al. [27] demonstrate that the total contrast and saliency estimation can be formulated unifiedly utilizing high-dimensional Gaussian channels. In this work, we generate a full-determination saliency map for each information image. Most above-said techniques measure saliency by measuring local focus encompass contrast and rarity of features over the whole image. In contrast, Gopalakrishnan et al. [12] formulate the protest location issue as a binary segmentation or labeling task on a graph. The most salient seed and several background seeds are distinguished by the behavior of random walks on a total graph and a k-regular graph. At that point, a semi-directed learning procedure is utilized to induce the binary labels of the unlabelled hubs. As of late a technique that adventures background priors is proposed for saliency recognition [34].

## III.PROBLEM DEFINITIONS

The approaches for deciding low-level saliency can be based on biological models or absolutely computational ones. Some approaches consider saliency more than several scales while others operate on a solitary scale. In general, all techniques utilize a few means of deciding local contrast of image areas with their surroundings utilizing at least one of the features of shading, power, and orientation. Usually, separate feature maps are created for each of the features utilized and then joined [8, 11, 6, 4] to obtain the final saliency map. An entire review of all saliency identification and segmentation research is past the extent of this paper, here we talk about those approaches in saliency recognition and saliency-based segmentation that are most relevant to our work. Ma and Zhang [11] propose a local contrast-based strategy for generating saliency maps that operates at a solitary scale and is not based on any biological model. The contribution to this local contrast-based map is a resized and shading quantized CIELuv image, sub-separated into pixel squares. The saliency map is obtained from summing up yielding's of image pixels with their separate encompassing pixels in a small neighborhood. This framework extracts the focuses and areas of attention. A fluffy developing technique then portions salient districts from the saliency map. Hu et al. [6] create saliency maps by thresholding the shading, force, and orientation maps utilizing histogram entropy thresholding analysis instead of a scale space approach. They then utilize a spatial compactness measure, figured as the area of the raised body encompassing the salient district, and saliency thickness, which is a component of the magnitudes of saliency values in the saliency feature maps, to measure the individual saliency maps before consolidating them. Itti et al. [9] have constructed a computational model of saliency-based spatial attention gotten from a biologically plausible architecture. They figure saliency maps for features of luminance, shading, and orientation at deferent scales that aggregate and consolidate information about each location in an image and nourish into a joined saliency map in a base up manner. The saliency maps delivered by Itti's approach have been utilized by different researchers for applications like adapting images on small gadgets [3] and unsupervised question segmentation [5, 10]. For instance, a Markov random field model is utilized to integrate the seed values from the saliency map along with low-level features of shading, surface, and edges to develop the salient protest districts [5]. Ko and Nam [10], then again, utilize a Support Vector Machine trained on the features of image portions to choose the salient locales of enthusiasm from the image, which are then bunched to extract the salient articles. We demonstrate that utilizing our saliency maps, salient protest segmentation is conceivable without requiring such complex segmentation algorithms. As of late, Frintrop et al. [4] utilized integral images [14] in VOCUS (Visual Object Detection with a Computational Attention System) to accelerate computation of focus encompass regard's for Finding salient areas utilizing separate feature maps of shading, force, and orientation. Although they obtain better determination saliency maps as compared to Itti's technique, they resize the feature saliency maps to a lower scale, subsequently losing determination. We utilize integral images in our approach however we resize the Filter at each scale instead of the image and hence maintain the same determination as the original i.

## IV. PROPOSED SOLUTION

The subject of how to measure the performance of learning algorithms and classifiers has been investigated. This is an intricate question with many aspects to think of one as finish of the analysis is that classifier performance is frequently measured as far as classification accuracy, e.g., with cross validation tests. A few strategies were observed to be general in the way that they can be utilized to evaluate any classifier (regardless of which algorithm was utilized to generate it) or any algorithm (regardless of the structure or representation of the classifiers it generates), while different techniques just are applicable to a certain algorithm or representation of the classifier.




First Stage Second Stage

**Fig.1 Proposed Model**

**1. Manifold Ranking**

In [39], a ranking method that exploits the intrinsic manifold structure of data (such as image) for graph labeling is proposed. Given a dataset X = (x1, . . . , xl, xl+1, . . . , xn) ∈ Rm×n, some data points are labelled queries and the rest need to be ranked according to their relevances to the queries. Let f: X → Rn denote a ranking function which assigns a ranking value fi to each point xi, and f can be viewed as a vector f = [f1, . . . , fn]T .

Let y = [y1, y2, . . . , yn]T denote an indication vector, in which yi = 1 if xi is a query, and yi = 0 otherwise. Next, we define a graph G = (V, E) on the dataset, where the nodes V are the dataset X and the edges E are weighted by an affinity matrix W = [wij]n×n. Given G, the degree matrix is D = diag{d11, . . . , dnn}, where dii = Pj wij. Similar to the PageRank and spectral clustering algorithms [5, 26], the optimal ranking of queries are computed by solving the following optimization problem:

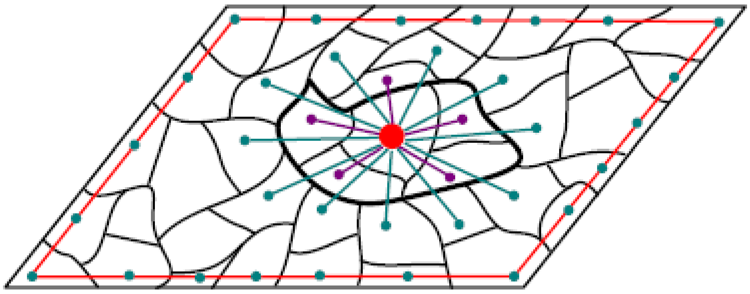


**Fig.2 Our graph model.**

The red line along the four sides indicates that all the boundary nodes are connected with each other.

$$\mathbf{f}^{*}=\arg\min_{f}\frac{1}{2}\Big(\sum_{i,j=1}^{n}w_{ij}\Big\|\frac{f_i}{\sqrt{d_{ii}}}-\frac{f_j}{\sqrt{d_{jj}}}\Big\|^2+\mu\sum_{i=1}^{n}\|f_i-y_i\|^2\Big), \tag{1}$$

where the parameter μ controls the balance of the smoothness constraint (the first term) and the fitting constraint (the second term). That is, a good ranking function should not change too much between nearby points (smoothness constraint) and should not differ too much from the initial query assignment (fitting constraint). The minimum solution is computed by setting the derivative of the above function to be zero. The resulted ranking function can be written as

$$\mathbf{f}^{*}=(\mathbf{I}-\alpha\mathbf{S})^{-1}\mathbf{y}, \tag{2}$$

where I is an identity matrix, α = 1/(1 + μ) and S is the normalized Laplacian matrix, S = D−1/2WD−1/2. The ranking algorithm [39] is derived from the work on semi-supervised learning for classification [38]. Essentially, manifold ranking can be viewed as an one-class classification problem [29], where only positive examples or negative examples are required. We can get another ranking function by using the unormalized Laplacian matrix in Eq. 2:

$$\mathbf{f}^{*}=(\mathbf{D}-\alpha\mathbf{W})^{-1}\mathbf{y} \tag{3}$$

We compare the saliency results using Eq. 2 and Eq. 3 in the experiments, and the latter achieves better performance. Hence, we adopt Eq. 3 in this work.

**2. Saliency Measure**

Given an input image represented as a graph and some salient query nodes, the saliency of each node is defined as its ranking score computed by Eq. 3 which is rewritten as f ∗ = Ay to facilitate analysis. The matrix A can be regarded as a learnt optimal affinity matrix which is equal to (D − αW)−1. The ranking score f ∗(i) of the i-th node is the inner product of the i-th row of A and y. Because y is a binary indicator vector, f ∗(i) can also be viewed as the sum of the relevances of the i-th node to all the queries. In the conventional ranking problems, the queries are manually labelled with the ground-truth. However, as

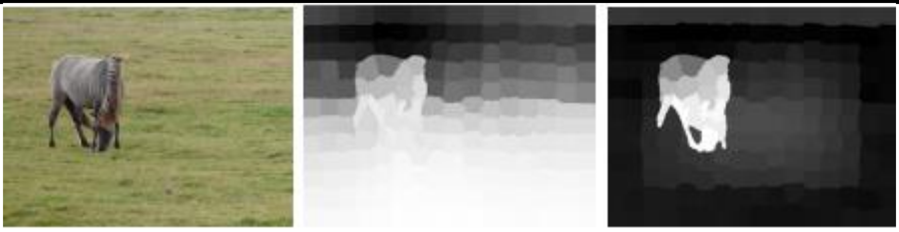


**Fig.3 Graph labeling results using the top boundary prior. Left: input images. Center: Results without enforcing the geodesic distance constraints. Right: Results with geodesic distance constraints.**

Queries for saliency detection are selected by the proposed algorithm; some of them may be incorrect. Thus, we need to compute a degree of confidence (i.e., the saliency value) for each query, which is defined as its ranking score ranked by the other queries (except itself). To this end, we set the diagonal elements of A to 0 when computing the ranking score by Eq. 3. We note that this seemingly insignificant process has great effects on the final results. If we compute the saliency of each query without setting the diagonal elements of A to 0, its ranking value in $f*$ will contain the relevance of this query to itself, which is meaningless and often abnormally large so as to severely weaken the contributions of the other queries to the ranking score. Lastly, we measure the saliency of nodes using the normalized ranking score $f*$ when salient queries are given, and using $1 - f*$ when background queries are given.

## V. ALGORITHM

**Bottom-up Saliency based on Manifold Ranking**

**Input:** An image and required parameters

1: Segment the input image into superpixels, construct a graph G with superpixels as nodes, and compute its degree matrix D and weight matrix W by Eq. 4.

2: Compute $(D - \alpha W)-1$ and set its diagonal elements to 0.

3: Form indicator vectors y with nodes on each side of image as queries, and compute their corresponding side-specific maps by Eq. 3 and Eq. 5. Then, compute the saliency map Sbq by Eq. 6.

4: Bi-segment Sbq to form salient foreground queries and an indicator vector y. Compute the saliency map Sfq by Eq. 3 and Eq. 7.

**Output:** a saliency map Sfq representing the saliency value of each superpixel.

## VI. EXPECTED RESULTS

1. To propose a bottom-up method to detect salient regions in images through manifold ranking on a graph, this incorporates local grouping cues and boundary priors.
2. To adopt a two-stage approach with the background and foreground queries for ranking to generate the saliency maps.
3. To prove the proposed algorithm is computationally efficient.

## VII. CONCLUSION

We propose a base up strategy to distinguish salient areas in images through manifold ranking on a graph, which incorporates local gathering cues and boundary priors. We adopt a two-stage approach with the background and foreground inquiries for ranking to generate the saliency maps. We evaluate the proposed algorithm on datasets and demonstrate promising outcomes with comparisons to state-of-the-art strategies. Moreover, the proposed algorithm is computationally proficient. Our future work will concentrate on integration of numerous features with applications to other vision issues.

ISSN(Online)-2454-4159 Volume 3, Issue 4, April 2017

## REFERENCES


[1] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk. Salient region detection and segmentation. In ICVS, 2008.

[2] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequencytuned salient region detection. In CVPR, 2009.

[3] R. Achanta, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. Slic superpixels. Technical report, EPFL, Tech.Rep. 149300, 2010.

[4] A. Borji, D. Sihite, and L. Itti. Salient object detection: A benchmark. In ECCV, 2012.

[5] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. Computer networks and ISDN systems, 30(1):107–117, 1998.

[6] N. Bruce and J. Tsotsos. Saliency based on information maximization. In NIPS, 2005.

[7] K.-Y. Chang, T.-L. Liu, H.-T. Chen, and S.-H. Lai. Fusing generic objectness and visual saliency for salient object detection. In ICCV, 2011.

[8] T. Chen, M. Cheng, P. Tan, A. Shamir, and S. Hu. Sketch2photo: Internet image montage. ACM Trans. on Graphics, 2009.

[9] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang, and S. M. Hu. Global contrast based salient region detection. In CVPR, 2011.

[10] J. Feng, Y.Wei, L. Tao, C. Zhang, and J. Sun. Salient object detection by composition. In ICCV, 2011.

[11] S. Goferman, L. Zelnik-Manor, and A. Tal. Context-aware saliency detection. In CVPR, 2010.

[12] V. Gopalakrishnan, Y. Hu, and D. Rajan. Random walks on graphs for salient object detection in images. IEEE TIP, 2010.

[13] L. Grady, M. Jolly, and A. Seitz. Segmenation from a box. In ICCV, 2011.

[14] J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. In NIPS, 2006.

[15] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. In CVPR, 2007.

[16] L. Itti. Automatic foveation for video compression using a neurobiological model of visual attention. IEEE TIP, 2004

[17] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. IEEE PAMI, 1998.

[18] 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.

[19] T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In ICCV, 2009.

[20] T. H. Kim, K. M. Lee, and S. U. Lee. Learning full pairwise affinities for spectral segmentation. In CVPR, 2010.

[21] D. Klein and S. Frintrop. Center-surround divergence of feature statistics for salient object detection. In ICCV, 2011.

[22] V. Lempitsky, P. Kohli, C. Rother, and T. Sharp. Image segmentation with a bounding box prior. In ICCV, 2009.

[23] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum. Learning to detect a salient object. IEEE PAMI, 2011.

[24] Y. Lu, W. Zhang, H. Lu, and X. Y. Xue. Salient object detection using concavity context. In ICCV, 2011.

[25] Y. Ma and H. Zhang. Contrast-based image attention analysis by using fuzzy growing. ACM Multimedia, 2003.

[26] A. Ng, M. Jordan, Y. Weiss, et al. On spectral clustering: Analysis and an algorithm. In NIPS, pages 849–856, 2002.

[27] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung. Saliency filters: Contrast based filtering for salient region detection. In CVPR, 2012.

[28] U. Rutishauser, D. Walther, C. Koch, and P. Perona. Is bottom-up attention useful for object recognition? In CVPR, 2004.

[29] B. Scholkopf, J. Platt, J. Shawe-Taylor, A. Smola, and R. Williamson. Estimating the support of a high-dimensional distribution. Neural Computation, 2001.

[30] J. Sun, H. C. Lu, and S. F. Li. Saliency detection based on integration of boundary and soft-segmentation. In ICIP, 2012.

[31] B. Tatler. The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. Journal of Vision, 2007.

[32] L. Wang, J. Xue, N. Zheng, and G. Hua. Automatic salient object extraction with contextual cue. In ICCV, 2011.

[33] W. Wang, Y. Wang, Q. Huang, and W. Gao. Measuring visual saliency by site entropy rate. In CVPR, 2010.

[34] Y. C. Wei, F. Wen, W. J. Zhu, and J. Sun. Geodesic saliency using background priors. In ECCV, 2012.

[35] Y. L. Xie, H. C. Lu, and M. H. Yang. Bayesian saliency via low and mid level cues. IEEE TIP, 2013.

[36] J. Yang and M. Yang. Top-down visual saliency via joint crf and dictionary learning. In CVPR, 2012.

[37] Y. Zhai and M. Shspatiotemporal cues. ACM Multimedia, 2006.

[38] D. Zhou, O. Bousquet, T. Lal, J. Weston, and B. Scholkopf. Learning with local and global consistency. In NIPS, 2003.

[39] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Scholkopf. Ranking on data manifolds. In NIPS, 2004.