Multi-graph Based Salient Object Detection

Idir Filali[1,a], Mohand Said Allili [2,b] and Nadjia Benblidia [1,c]

[1]LRDSI Laboratory, University of Blida 1, Blida, Algeria.
[2]LARIVIA Laboratory, Universite du Quebec en Outaouis, QC, Canada.
[a] inf_tyg@yahoo.fr
[b] mohandsaid.allili@uqo.ca
[c] benblidia@yahoo.com

Abstract. We propose a multi-layer graph based approach for salient object detection in natural images. Starting from a set of multi-scale image decomposition using superpixels, we propose an objective function optimized on a multi-layer graph structure to diffuse saliency from image borders to salient objects. After isolating the object kernel, we enhance the accuracy of our saliency maps through an objectness-like based refinement approach. Beside its simplicity, our algorithm yields very accurate salient objects with clear boundaries. Experiments have shown that our approach outperforms several recent methods dealing with salient object detection.

Keywords: multi-layer graphs, multi-scale segmentation, salient object detection.

1 Introduction

Saliency detection aims to localize most informative objects or regions in images. Saliency detection methods rely either on a local or global contrast estimation. Local contrast based methods [9] assume that regions which stand out from their neighborhoods have high saliency values. These methods are more suitable to highlight salient object boundaries instead of the entire objects. Global contrast based methods [1, 12] express rarity of a regions compared to the overall image in terms of global statistics. They are better at highlighting entire salient regions. However, they are less accurate to detect large-sized objects due to the fact the the object statistics dominate the global statistics of the image.

Recently, several methods have combined local and global contrasts to overcome the aforementioned limitations. Among proposed techniques, graph-based methods have emerged as an excellent tool for salient object detection [11, 8]. In addition to the simplicity they provide for combining several image cues, graphs are efficient in encoding spatial priors such as object contiguity and location. For example, [11, 8] propose a graph-based method to detect salient objects far from the image border. In [4, 5], random walk models are used to extract salient objects on graphs. However, since the graphs are built of the image lattice, these methods incur a huge computation time. To circumvent this limitation, [11] use superpixels instead of pixels and propose to optimize functions on graphs. Furthermore,

to make use of multi-resolution image information, [8] use global contrast and spatial contiguity to generate initial saliency maps. Then, a region merging procedure with dynamic scale control is used to generate the so-called *saliency trees*. This method highlights salient object regions with well-defined boundaries. Similarly, [10] propose a hierarchical model to estimate saliency maps at different image resolutions. These maps are then combined using a weighted color distance. This method yields a good saliency maps, but may assign high saliency values to isolated background regions. Note that the majority of graph-based methods use the image borders to extract the background. One major limitation of these methods is that they can not deal efficiently with variable object scale as the graph construction highly depends on the initial (arbitrary) segmentation scale. Some parts of the salient object can then be confounded with the background in case of over/under segmentation. A multi layered graph based method can provide a solution to this problem.

In this paper we propose a multi-layer graph ranking approach for salient object detection. Starting from a multi-scale image decomposition of the image into superpixels, each layer of the graph will be constructed on superpixels of a given segmentation scale. The image saliency is obtained by optimizing an objective function on the graph, which detects the location of coarse and fine parts of salient objects. We also propose a window based refinement process which narrows the object localization and enhances the overall contrast of the object with the immediate background. Experiments on the standard MSRA dataset containing complex scenes have demonstrated that our approach outperforms several recent state-of-the art methods.

This paper is organized as follows: Section 2 presents our algorithm for multi-graph salient object detection. Section 3 presents some experimental results validating our approach. We then end the paper with a conclusion.

2 GRAPH RANKING FOR SALIENCY DETECTION

2.1 multi-layers graph for saliency detection

Let L the number of layers composing our multi graph. We suppose then that the segmentations required to this structure are defined by the sets: $\Omega_1, \Omega_2, \ldots, \Omega_L$ where each set $\Omega_\ell, \ell \in \{1, 2, ..., L\}$. Is obtained by applying the SLIC algorithm [2] which produce n_ℓ superpixels. We use one granularity defined by a certain number of superpixels at each layer, so we obtain several resolutions $N = \{n_1, n_2, \ldots, n_L\}$ (see Fig.1). We consider the following minimization to propagate the ranking to all the nodes of the graph:

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} \left[\left(\sum_{\ell=1}^{L} \sum_{i,j=1}^{n_\ell} w_{ij}^{(\ell)} \left(f_i^{(\ell)} / \sqrt{d_i^{(\ell)}} - f_j^{(\ell)} / \sqrt{d_j^{(\ell)}} \right)^2 + \lambda \sum_{i=1}^{n_\ell} (f_i^{(\ell)} - y_i^{(\ell)})^2 \right) + \sum_{\ell=1}^{L-1} \sum_{m=\ell+1}^{L} \left(\sum_{i=1}^{n_\ell} \sum_{j=1}^{n_m} w_{ij}^{(\ell,m)} \left(f_i^{(\ell)} / \sqrt{\tilde{d}_i^{(\ell)}} - f_j^{(\ell)} / \sqrt{\tilde{d}_j^{(m)}} \right)^2 \right) \right], \quad (1)$$

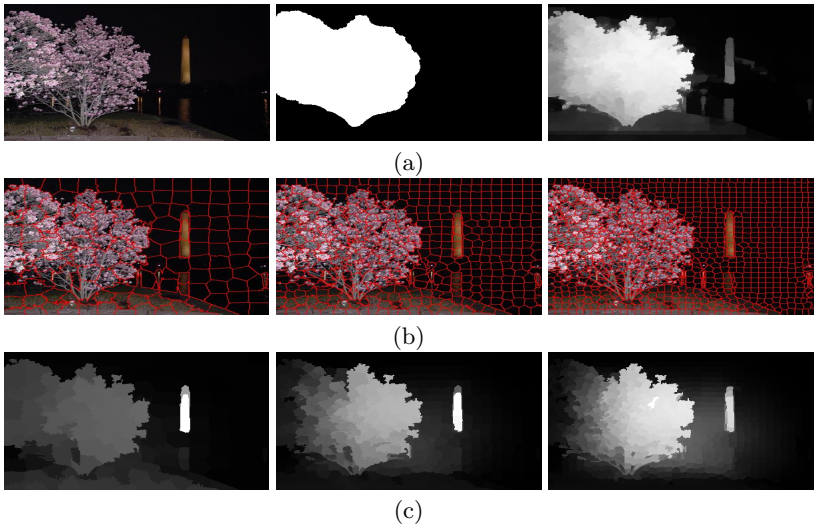


Fig. 1. Comparison between obtained saliency maps using single-layer and multi-layer graphs. (a) represents, from left to right, the original image, the ground truth and the multi-layer graph saliency map. (b) represents, from left to right, the segmentation of the image into 150, 500 and 1000 superpixels, respectively. (c) represents, from left to right, saliency maps generated by a single-layer graph using each segmentation in (b) of the same column.

where $d_i^{(\ell)} = \sum_{k=1}^{n_\ell} w_{ik}^{(\ell)}$, $\tilde{d}_i^{(\ell)} = \sum_{m=1,m\neq\ell}^{L} \sum_{k=1}^{n_m} w_{ik}^{(\ell,m)}$, $w_{ik}^{(\ell)} = e^{-\alpha.\|c_{\ell i} - c_{\ell k}\|}$ and $w_{ik}^{(\ell m)} = e^{-\alpha.\|c_{\ell i} - c_{mk}\|}$ is the weight between superpixels $r_i^{(\ell)}$ and $r_k^{(\ell)}$ located on the same levelΩ_ℓ with α a constant that controls the sensitivity of the weight and $c_{\ell i}$ represents the mean color vector of superpixel $r_i^{(\ell)}$. The weight $w_{ik}^{(\ell,m)} = e^{-\alpha.\|c_{\ell i} - c_{mj}\|}$ links the two regions $r_i^{(\ell)}$ and $r_k^{(m)}$ situated in the levels Ω_ℓ and Ω_m respectively.

The variable $\ell \in \{1, 2, \ldots, L\}$ correspond to the index of the segmentation located in ℓ-th layer and n_ℓ represent the number of segments generated using the SLIC algorithm. We note that $y^\ell = [y_1^\ell, y_2^\ell, \ldots, y_n^\ell]$ is the indicator vector in which $y_i^\ell = 1$ if a region r_i^ℓ in the layer ℓ is a query and $y_i^\ell = 0$ otherwise. We note that $f = \bigcup_{\ell=1}^{L} f^\ell$ where $f^\ell = \{f_1^\ell, f_2^\ell, \ldots, f_{n_\ell}^\ell\}$ represent the ranking of regions at the ℓ-th level and λ is a regularization constant. Let $D_i^{(\ell)} = \tilde{d}_i^{(\ell)} + d_i^{(\ell)}$, $W = [W^{\ell,m}]$, with ℓ and m, $(\ell, m) \in (\{1, 2, \ldots, L\})^2$ are the line and column numbers respectively and $W^{\ell,m} = [W_{ij}^{\ell,m}]$, $i \in \{1, 2, \ldots, n_\ell\}$ and $j \in \{1, 2, \ldots, n_m\}$. The minimization of the equation 1 is done as follows: $\mathbf{f}^* = (\mathbf{I} - \alpha\mathbf{L})^{-1}\mathbf{y}$, where $\alpha = 1/(1+\lambda)$ and $\mathbf{L}$ is the Laplacian matrix given by: $\mathbf{L} = \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{1/2}$, where D is a diagonal matrix $\mathbf{D} = \text{diag}[D_1^{(1)}, ..., D_{n_1}^{(1)}, ..., ..., D_1^{(L)}, ..., D_{n_L}^{(L)}]$. The saliency value at the superpixel r_i^ℓ is computed as follows: $s_i^{(\ell)} = 1 - f_i^{(\ell)}$.

We perform a propagation process from each of the four borders of the image (top, down, left and right) by initializing its appropriate positions to 0 in y_i^ℓ. We

then obtain four saliency maps (S_T,S_D, S_L,S_R) that are combined by multiplication as follows: $\mathbf{S} = \mathbf{S}_T \circ \mathbf{S}_D \circ \mathbf{S}_R \circ \mathbf{S}_L$, where $\circ$ designates the Hadamard product between matrices. We then perform a second propagation process from the most salient elements extracted from $\mathbf{S}$ (the kernel of the object) to the rest of the image. This is achieved by a proper initialization of function (1) based on the object kernel.

2.2 Window based saliency refinement

For more accurate saliency estimation, we narrow the saliency area to a restricted space that is more likely to contain the salient object (i.e., objectness-like measure [3]). For this purpose, we define a window initialized to entire image and progressively reduced to fit tightly to the area containing the salient object with high plausibility. Let (x_0, y_0) be respectively the height and the width of the image. The window extent is defined as $\mathcal{X}_{\delta_1} = [\delta_1, x_0 - \delta_1]$ and $\mathcal{Y}_{\delta_2} = [\delta_2, y_0 - \delta_2]$, with δ_1 and δ_2:

$$(\delta_1, \delta_2) = \arg\max_{\delta_1,\delta_2} \left\{ \sum_{x \notin \mathcal{X}_{\delta_1}} \sum_{y \notin \mathcal{Y}_{\delta_2}} S(x,y) \leq \eta * \sum_{x=1}^{x_0} \sum_{y=1}^{y_0} S(x,y) \right\}, \tag{2}$$

where η is a threshold set experimentally (usually $\eta = 0.99$). Eq. (2) aims to define a window that contains most salient parts of the image and discard the background with weak saliency percentage compared to the over all image. We

Fig. 2. Example on a window-based refinement: (a) The input image, (b) the ground truth, (c) saliency map obtained using Eq. 1, (d) Window position (Eq. 2), (e) saliency map obtained using window refinement process.

perform a second multi-layer graph based saliency detection process in the inner part of the window by minimizing function (1) (see Fig. (2) for illustration).

3 Experimental results

The performance of the proposed model is evaluated quantitatively on a widely used MSRA dataset [7] which contains 5,000 images with their ground truth. We compare our multi graph based manifold ranking (MMR) and our window based refinement (WMMR) with five state of the art saliency detection methods

: ST [8], MSA [13], ORBD [14], HDCT [6] and MR [11]. Fig. (3) shows some examples that compare our method to those of the state of the art. We evaluate the performance of our method using the *precision*, *recall* and *F-Measure* metrics. The precision/recall curves are obtained by binarizing the saliency map using thresholds in the range from 1 and 254. The *F-Measure* is the weighted harmonic sum of the *precision* and *recall* computed as follows:

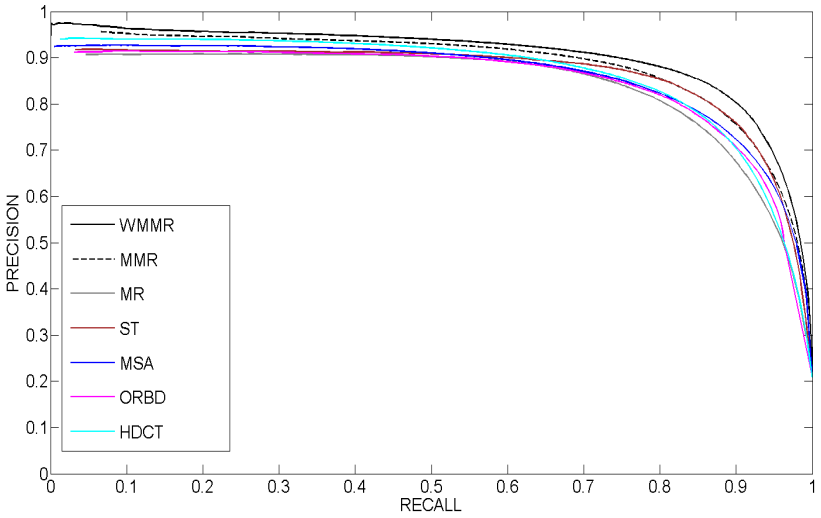
$$F_\alpha = \frac{(1+\alpha) \times precision \times recall}{\alpha \times precision + recall}, \tag{3}$$

where we set $\alpha = 0.3$ for all the compared methods. Fig. 4.a shows the *precision* / *recall* comparison of our method with five state of the art methods and Fig. 4.b shows the best average *F-measure*, *precision* and *recall* measures obtained on all saliency maps generated on the MSRA dataset by each compared method. The highest *F-measure* value is returned by our WMMR method which is 0.87 followed by ST which returns 0.84 for the same measure. This demonstrates clearly the performance of our method with regard to the compared ones. We can note also that our method has obtained more quality saliency maps where objects are more contrasted with the background and have clearer boundaries.

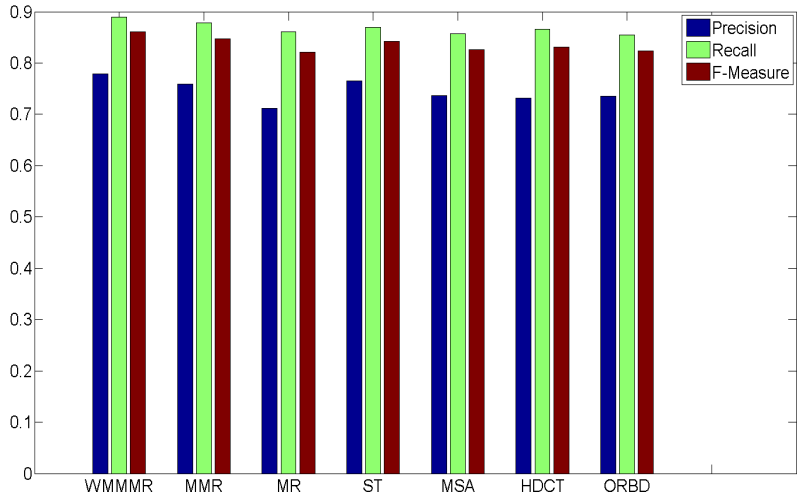
Fig. 3. Visual comparison between our method to five state of the arts approaches.

4 Conclusion

We have presented a multi-layer graph based algorithm for salient object detection. Our method is able to localize accurately fine and coarse parts of salient objects. It also uses a refinement process using an objectness-like measure narrowing the salient object search. Experimental results on the well-known MSRA dataset have demonstrated that our method outperforms recent sate-of-the-art methods. Qualitatively, our salient objects are generally more emphasized and the backgrounds are more efficiently discarded, which make our results more suitable to applications such as object segmentation and selective object recognition.



(a)



(b)

Fig. 4. Comparative results between our method and five state-of-the-art saliency detection methods using the MSRA dataset.

References

1. R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-Tuned Salient Region Detection. *IEEE Conf. on Comp. Vision and Pattern Recognition*, 1597-1604, 2009.
2. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Sösstrunk. SLIC Superpixels Compared to State-of-the-art Superpixel Methods. *IEEE T. on Pattern Aanalysis and Machine Intelligence*, 34(11):2274-2282, 2012.
3. B. Alexe, T. Deselaers and V. Ferrari. Measuring the Objectness of Image Windows. *IEEE T. on Pattern Aanalysis and Machine Intelligence*, 34(11):2189-2202, 2012.
4. V. Gopalakrishnan, Y. Hu and D. Rajan. Random Walks on Graphs for Salient Object Detection in Images. *IEEE T. on Image Processing*, 19(12):3232-3242, 2010.
5. B. Jiang, L. Zhang, H. Lu, C. Yang, and M.-H. Yang. Saliency Detection via Absorbing Markov Chain. *IEEE Int'l Conf. on Computer Vision*, 1665-1672, 2013.
6. J. Kim, D. Han, Y.W Tai and J. Kim. Salient Region Detection via High-Dimensional Color Transform and Local Spatial Support. *IEEE T. on Image Processing*, 25(1): 9-23, 2016.
7. T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum. Learning to Detect a Salient Object. *IEEE T. on Pattern Analysis and Machine Intelligence*, 33(2):pp. 353-367, (2011).
8. Z. Liu, W. Zou and O. Le Meur. Saliency Tree: A Novel Saliency Detection Framework. *IEEE T. on Image Pocessing*, 23(5):1937-1952, 2014.
9. H. J. Seo and P. Milanfar. Static and space-time visual saliency detection by self-resemblance. *IEEE Conf. on Comp. Vision and Pattern Recognition Workshops*,9(12): 45-52, 2009.
10. Q. Yan, L. Xu, J. Shi and J. Jia. Hierarchical Saliency Detection. *IEEE Conf. on Comp. Vision and Pattern Recognition*, 1155-1162, 2013.
11. C. Yang, L. Zhang, H. Lu and X. Ruan. Saliency Detection via Graph-Based Manifold Ranking. *IEEE Conf. on Comp. Vision and Pattern Recognition*, 3166-3173, 2013.
12. L. Zhang, Z. Gu and H. Li. SDSP: A Novel Saliency Detection Method by Combining Simple Priors. *IEEE Int'l Conf.on Image Processing*, 171-175, 2013.
13. L. Zhu, D.A. Klein, S. Frintrop, Z. Cao and A.B. Cremers. A Multi-size Superpixel Approach for Salient Object Detection based on Multivariate Normal Distribution Estimation. *IEEE T. on image processing*, 23(12):5094-5107, 2014.
14. W. Zhu, S. Liang, Y. Wei and J. Sun. Saliency Optimization from Robust Background Detection. *IEEE Conf. Comp. Vision and Pattern Recognition*, 2814-2821, 2014.