International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 02
February 2017

Salient Region Detection via Integrating Diffusion-Based Compactness and Local Contrast

[1]Mr.D.VENKATARAMI REDDY, [2]Mr.SK SUBHAN, [3]Mr. P.SANDEEP

[1]Associate Professor & HOD, ECE, Madhira Institute of Technology & science, Chilkur, Suryapet, India
[2]Assistant Professor in ECE, Madhira Institute of Technology & science, Chilkur, Suryapet, India
[3]PG Scholar, ECE, Madhira Institute of Technology & science, Chilkur, Suryapet, India
[1]dvenkataramireddy@gmail.com; [2]subhanece15@gmail.com; [3]sandeeppillutla33@rediffmail.com;

ABSTRACT

Salient region detection is a challenging problem and an important topic in computer vision. It has a wide range of applications, such as object recognition and segmentation. Many approaches have been proposed to detect salient regions using different visual cues, such as compactness, uniqueness, and objectness. However, each visual cue-based method has its own limitations. After analyzing the advantages and limitations of different visual cues, we found that compactness and local contrast are complementary to each other. In addition, local contrast can very effectively recover incorrectly suppressed salient regions using compactness cues. Motivated by this, we propose a bottom-up salient region detection method that integrates compactness and local contrast cues. Furthermore, to produce a pixel-accurate saliency map that more uniformly covers the salient objects, we propagate the saliency information using a diffusion process. Our experimental results on four benchmark data sets demonstrate the effectiveness of the proposed method. Our method produces more accurate saliency maps with better precision-recall curve and higher F-Measure than other 19 state-of-the-arts approaches on ASD, CSSD, and ECSSD data sets.

I. INTRODUCTION

Visual attention is an important mechanism of the human visual system. It filters out redundant visual information and effectively selects highly relevant subjects, which are called the salient objects. Visual attention is considered to involve two mechanisms: stimulus driven [1] and task driven . The stimulus-driven mechanism is often called bottom-up, and is fast, involuntary, and purely based low-level visual stimuli. The task-driven mechanism is called top-down, and is based on high-level information such as prior knowledge of the task, emotions, and expectations. Accordingly,

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 02
February 2017

computational visual attention methods can be categorized into bottom-up and top-down methods. In this paper, we focus on bottom-up salient region detection tasks. Salient region detection methods aim to completely highlight entire objects of interest and sufficiently suppress background regions. Their output can be used for numerous computer vision problems such as image classification object detection and recognition image compression and image segmentation As a fundamental computer vision task, salient region detection has been extensively studied over the past few years, and a number of algorithms have been proposed Most bottom-up salient region detection methods rely on visual cues to consistently separate the salient object and background. These cues include uniqueness compactness and background.

Most uniqueness-based methods use low-level features of the image (such as intensity, color, and orientation) to determine the contrast between image pixels or regions and their surroundings. According to the contrastive reference regions, these methods can be roughly divided into local- and global contrast-based methods. Local contrast-based methods consider the uniqueness of pixels (or superpixels, image regions) with respect to their surrounding regions or local neighborhoods, whereas global contrast-based methods consider contrast relationships over the entire image. Unlike uniqueness-based methods, which consider the uniqueness of low-level features in the feature space, compactness based methods consider the spatial variance of features. Ideally, salient pixels (or superpixels, image regions) tend to have a small spatial variance in the image space, whereas the background is distributed over the entire image and has a high spatial variance. Background based methods use boundary and connectivity priors derived from common backgrounds in natural images . These methods are primarily motivated by the psychophysical observations that salient objects seldom touch the image boundary, and most background regions can be easily connected to each other. Although the above-mentioned methods have achieved good results in some aspects, each method has its own limita-

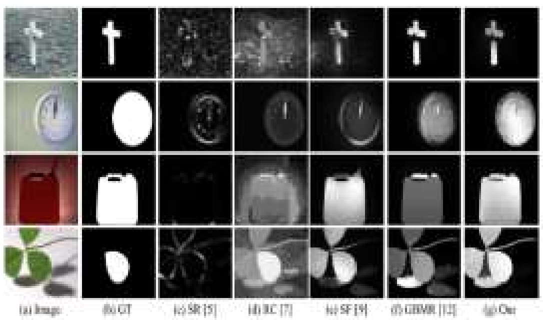

Fig. 1. Visual limitations of different methods. (a) Input image. (b) Ground truth salient regions. (c) Saliency maps using local contrast based method . (d) Saliency maps using global contrast based method [7]. (e) Saliency maps using compactness based method . (f) Saliency maps using background based method . (g) Our method. tions.

For example, Fig. 1 illustrates the saliency detection results using four state-of-the-art methods Figure 1(c) shows that the local contrast-based method tends to highlight the salient object's edges instead of uniformly propagating the saliency to the interior. The global contrast based method sometimes produces high saliency values for non-salient regions, especially for regions with complex patterns or rare background distractors. This is shown in the first example of Fig. 1(d), where some grass regions in the background are highlighted. A typical limitation of the compactness based method is that some salient regions may be wrongly suppressed when the foreground objects and background are similar. In the second example of Fig. 1(e), the inner smooth parts of the clock are wrongly suppressed. Finally, background based methods can perform very well. However, they fail when the salient objects touch the image boundary, as illustrated in the last two examples of Fig. 1(f). From the above discussion, we can conclude that single visual cue based salient region detection methods all have their own limitations. To determine these limitations, different visual cues should be integrated into a unified framework. Motivated by this approach, some methods integrate multiple visual cues. Perazzi et al. [9] proposed a saliency filters method, which unifies uniqueness and compactness (of the spatial distribution) into a single, high-dimensional, Gaussian filtering framework. However, global contrast and compactness based methods have difficulty distinguishing between similar colors in the foreground and background. Consequently, the saliency filters method fails when foreground objects and the background are similar (e.g., the second example in Fig. 1(e)). In this work, we integrated local contrast and compactness visual cues to generate saliency maps. Compared with the global contrast method, the local contrast method is a more appropriate complement to compactness. When the foreground is similar to some background regions, global contrast and compactness methods may wrongly suppress the foreground region. However, local contrast methods may properly highlight the foreground region based on

.II. RELATED WORK

Our work focuses on bottom-up salient region detection. A comprehensive survey of visual attention and saliency detection can be found in, and a quantitative analysis of different methods was provided in . According to the type of visual cue, bottom-up salient region detection methods can be broadly classified into uniqueness, compactness, and background based. Furthermore, uniqueness-based methods can be roughly divided into local and global contrast-based techniques. One of the first local contrast-based methods was themodel of Itti et al. They used a difference of Gaussians approach to extract multi-scale color, intensity, and orientation information from images. This information was then used to define saliency by calculating center-surround differences. Ma and Zhang proposed an alternate local contrast analysis for generating saliency maps. They directly computed center-surround color differences in a fixed neighborhood for each pixel, and then extended the saliency map using a fuzzy growth model. Harel et al. proposed a graph based visual saliency method for non-linearly combining local uniqueness maps from different feature channels to concentrate conspicuity. Hou and Zhang introduced a model in the frequency domain, which defines the saliency of a location based on the difference between the log-spectrum feature and its surrounding local average. Achanta et al. calculated the saliency by computing center-surround contrasts of the average feature vectors, between the inner and outer sub-regions of a sliding square window. Liu et al. computed center surround histograms over windows of various sizes and aspect ratios in a sliding window. They trained a conditional random field to combine different features for salient object detection.Jiang et al. used the difference between the color histogram of a region and its immediately neighboring regions to evaluate the saliency score.

Global contrast-based methods compute the saliency of individual pixels or image regions using contrast relationships over the complete image. Zhai and Shah computed

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 02
February 2017

pixel-level saliency using the contrast with all other pixels. Bruce and Tsotsos exploited Shannon's selfinformation measure in a local context to compute saliency. Achanta et al. achieved globally consistent results based on a frequency-tuned method, which directly defines pixel saliency using the difference from the average image color. Goferman et al. highlighted salient objects with their contexts by simultaneously modeling local low-level clues, global considerations, visual organization rules, and high-level features. Cheng et al. proposed a regional contrast-based saliency extraction algorithm, which simultaneously considers the global region contrast over the entire image in the Lab color space and the spatial coherence, and used them to compute a saliency map.

III. DIFFUSION PROCESSES

There has recently been a growing interest in using diffusion processes to propagate saliency information throughout a Harel et al. used graph algorithms and a measure of dissimilarity to compute saliency in their graph-based visual saliency model. Yang et al. cast saliency detection into a graph-based ranking problem. Jiang et al. formulated saliency detection via an absorbing Markov chain on an image graph model. Gopalakrishnan et al. used Markov random walks on two different graphs to detect the salient seed nodes. In Lu et al. proposed a method for learning optimal seeds for object saliency using a diffusion process. In all diffusion processes, the image is mapped into a graph G = (V, E) with N nodes {v1, v2, · · · , vN }, and edges E weighted by an affinity matrix W = _wi j _N×N . Node vi corresponds to the i th image superpixel or patch and edge ei j link nodes vi and v j to each other. After the graph is constructed, using a given vector of N saliency observations or saliency seeds y = [y1, y2, · · · , yN]T , diffusion process spread the seeds through the graph based on the defined affinity matrix. Finally we derive an object saliency map f = [f1, f2, · · · , fN]T . In the following, we review the two most commonly used diffusion processes in salient region detection: random walk and manifold ranking. A detailed review of diffusion processes focusing on image retrieval was presented in.

A. Random Walk

Random walk is a popular diffusion process used in salient detection. The model interprets the diffusion processes as a random walk on the graph G, where a so-called transition matrix defines the probabilities for walking from one node to its linked nodes. Finally, the saliency is formulated as the equilibrium distribution of the random walk. The random walk transition matrix is defined as P = D−1W (1) where D = diag {d11, d22, · · · , dNN}, and dii is the degree of nodes vi (i.e. dii = _j wi j). Then, a single step of the diffusion process is characterized by the simple iteration f t+1 = P T f t . This standard random walk can be modifiedby introducing a random jump, such that at each step t, a random walk occures with probability α, while a random jump to an arbitrary node occures with probability 1−α. Thus, the diffusion process is f t+1 = αP T f t + (1 − α)y (2) where 1 − α is the random jump probability, and y are the saliency seeds that define the probabilities of randomly jumping to corresponding nodes. Following the iterated diffusion process can converge to an equilibrium distribution

f ∗ = (1 − α) _I − αPT _ −1

y (3)

where I is the identity matrix.

B. Manifold Ranking

Inspired by the standard PageRank approach, Zhou et al. proposed a manifold ranking method that exploits the intrinsic manifold structure of data. The main procedure of this method is as follows. We first construct a graph on the data with an affinity matrix W, and assign a ranking score to each query node (which is the saliency seed in this paper). Then, a diffusion process propagates the ranking score of each node to their nearby neighbors via the graph. The diffusion process repeats until a global stable state is achieved, and all nodes are ranked according to their final ranking scores. As illustrated in the optimal ranking is computed by minimizing an energy

IV. PROPOSED METHOD

In this section, we present an efficient and effective saliency region detection method that integrates diffusion-based compactness and local contrast. We first abstract the image into superpixels and construct a graph. Next, we compute two complementary saliency maps using the compactness visual cue and local contrast. The resulting saliency maps are propagated using a diffusion process and the constructed graph. Finally, we integrate the two computed saliency maps to generate a pixel-wise saliency map.

A. Graph Construction

Following the observation of Perazzi et al. that abstracting an input image into homogeneous superpixels can improve the performance of salient object detection, we used the SLIC model to abstract the input image into uniform and compact regions. After abstracting the image, we construct a graph G = (V, E), Each node corresponds to a superpixel generated by the SLIC model. Most existing connect each node to neighboring nodes and nodes that share common boundaries with neighboring nodes (k-regular graph). However, in this graph, each node is only connected to its neighboring nodes. Additionally, each pair of boundary nodes are connected to each other to reduce the geodesic distance of similar superpixels. In this work, we define the Lab color space distance li j between nodes vi and v j as li j =__ ci − c j__ (9) where ci and c j are the mean of superpixels corresponding to nodes vi and v j in the Lab color space. Note that the distance matrix L = _li j _N×N is normalized to the interval [0, 1]. The affinity matrix W is defined as

wi j =e −li j σ2 if j ∈ Ni 0 otherwise (10)

where σ is a constant, and Ni denotes the set of neighbors of node vi . Note that all nodes around the image borders are considered neighbors of each other. Given the affinity matrix W, the saliency propagation is implemented using Equation (8).

B. Diffusion-Based Compactness

Salient objects generally correspond to real objects, therefore they are grouped together into connected regions. Therefore, salient objects typically have compact spatial distributions, whereas background regions have a wider distribution over the entire image. Motivated by this, we calculate the spatial variance of the superpixels. We first define the similarity ai j between a pair of superpixels vi and v j, using ai j = e −li j σ2 (11)

To describe the similarity between superpixels more precisely, we propagate the similarity using the manifold ranking through the constructed graph. That is, HT = (D − αW) −1A (12) where A = [ai j] N×N, and H = [hi j] N×N is the similarity matrix after the diffusion process. Salient objects are generally surrounded by background regions. Thus, in the spatial domain, the colors of background regions typically have a larger spread over the whole image, when compared with salient colors. Colors that exhibit large spatial variance across the image are less likely to be salient. We calculate the spatial variance of superpixel vi using sv (i) = _Nj =1 hi j · n j ·__

b j − µi__ _Nj =1 hi j · n j (13) where n j is the number of pixels that belong to superpixel v j ,

b j = _bxj

C. Diffusion-Based Local Contrast

Although compactness based methods achieve good results in some aspects, they have limitations. When the foreground objects and background have similar appearances, some salient

regions may be wrongly suppressed. To mitigate this, some approaches integrate the compactness visual cue with other cues. Perazzi et al. [9] unified the compactness and uniqueness into a single high-dimensional Gaussian filtering framework. However, we found that the local contrast is more complementary to compactness than the global contrast. When a foreground region is similar to some background regions, global contrast and compactness methods may wrongly suppress the foreground, whereas local contrast methods can highlight the foreground based on the contrast with its neighbor. In this section, we determine the saliency using the local contrast of an image superpixel with respect to its neighboring superpixels. Although the local contrast can highlight the foreground regions that are wrongly suppressed by the compactness

method, it may also highlight some background regions. Considering this, we use the saliency map calculated using compactness to suppress the incorrectly highlighted background regions. We define the Lab color space distance A smaller ld(i) value corresponds to a higher probability

that superpixel vi belongs to the background.We set any value of ld that is less than the mean ld to zero. The ld saliency map. To enhance the reliability of the foreground detection (especially for complicated images) and improve the overall quality of salient region segmentation, we define the distribution measure for a superpixel vi with respect to the centroid of saliency map

D. Saliency Map Integration

The compactness and local contrast saliency cues methods efficiently produce two different saliency maps, Scom and Sloc. These maps are complementary to each other. We directly integrate these two different saliency maps to define the final saliency map,

S = Norm(Scom + Sloc)

V. CONCLUSION

In this paper, we proposed a bottom-up method for detecting salient regions in images by integrating two complementary visual cues (compactness and local contrast) with diffusion processes. After considering the advantages and limitations of different visual cues, we found that compactness and local contrast are complementary to each other. Additionally, local contrast can effectively recover the incorrectly suppressed salient regions using compactness cues. To produce a pixelaccurate saliency map that more uniformly covers the salient objects, we propagate the saliency information using a manifold ranking diffusion process on a graph. Our experimental results using four benchmark datasets demonstrated the effectiveness of the proposed method; it produced more accurate saliency maps with better precision-recall curves and higher F-measures than 19 state-of-the-art approaches, when applied to the ASD, CSSD, and ECSSD datasets.

REFERENCES

[1] H. E. Egeth and S. Yantis, "VISUAL ATTENTION: Control, representation, and time course," Ann. Rev. Psychol., vol. 48, pp. 269–297, Feb. 1997.

[2] J. M. Henderson, "Human gaze control during real-world scene perception," Trends Cognit. Sci., vol. 7, no. 11, pp. 498–504, 2003.

[3] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[4] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in Proc. Adv. Neural Inf. Process. Syst., 2006, pp. 545–552.

[5] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2007, pp. 1–8.

[6] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequency-tuned salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 1597–1604.

[7] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 409–416.

[8] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, and S. Li, "Automatic salient object segmentation based on context and shape

prior," in Proc. Brit. Mach. Vis. Conf., 2011, pp. 1–12.

[9] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 733–740.

[10] Y. Wei, F. Wen, W. Zhu, and J. Sun, "Geodesic saliency using background priors," in Proc. 12th Eur. Conf. Comput. Vis., 2012, pp. 29–42.

[11] Y. Xie, H. Lu, and M.-H. Yang, "Bayesian saliency via low and mid level cues," IEEE Trans. Image Process., vol. 22, no. 5, pp. 1689–1698, May 2013.

[12] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2013, pp. 3166–3173.

[13] C. Kanan, M. H. Tong, L. Zhang, and G. W. Cottrell, "SUN: Top-down saliency using natural statistics," Vis. Cognit., vol. 17, nos. 6–7, pp. 979–1003, 2009.

[14] J. Yang and M.-H. Yang, "Top-down visual saliency via joint CRF and dictionary learning," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 2296–2303.

[15] C. Siagian and L. Itti, "Rapid biologically-inspired scene classification using features shared with visual attention," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 2, pp. 300–312, Feb. 2007.

[16] G. Sharma, F. Jurie, and C. Schmid, "Discriminative spatial saliency for image classification," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 3506–3513.

[17] D. Walther, U. Rutishauser, C. Koch, and P. Perona, "Selective visual attention enables learning and recognition of multiple objects in cluttered scenes," Comput. Vis. Image Understand., vol. 100, nos. 1–2, pp. 41–63, 2005.

[18] B. Alexe, T. Deselaers, and V. Ferrari, "Measuring the objectness of image windows," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2189–2202, Nov. 2012.

[19] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," IEEE Trans. Image Process., vol. 19, no. 1, pp. 185–198, Jan. 2010.

[20] L. Wang, J. Xue, N. Zheng, and G. Hua, "Automatic salient object extraction with contextual cue," in Proc. IEEE Int.

Conf. Comput. Vis., Nov. 2011, pp. 105–112.
pp. 1107–1114.

[15] T. Michaeli and M. Irani, "Blind deblurring using internal patch recurrence," in Proc. IEEE Eur. Conf. Comput. Vis. (ECCV), Sep. 2014, pp. 783–798.

[16] S. Cho and S. Lee, "Fast motion deblurring," ACM Trans. Graph., vol. 28, no. 5, 2009, Art. ID 145.

[17] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2009, pp. 1964–1971.

[18] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2011, pp. 2657–2664.

[19] A. Rav-Acha and S. Peleg, "Two motion-blurred images are better than one," Pattern Recognit. Lett., vol. 26, no. 3, pp. 311–317, 2005.

[20] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," ACM Trans. Graph., vol. 26, no. 3, 2007, Art. ID 1. Editor-in-Chief of the SIAM Journal on Imaging Sciences.

AUTHOR'S PROFILE:

Mr.Devireddy Venkatarami Reddy, received the Master of Technology degree in EMBEDDED SYSTEMS from the DR.PAULRAJ ENGINEERING COLLEGE -JNTUH, he received the Bachelor of Engineering degree from S.A. ENGINEERING COLLEGE-ANNA UNIVERSITY. He is currently working as Associate Professor and a Head of the Department of ECE with Madhira Institute of Technology And Sciences, kodad. His interest subjects are Embedded Systems, Microprocessors, Communication Systems, Digital Electronics and etc.

Email id: dvenkataramireddy@gmail.com

Mr.SK SUBHAN, received the Master Of Technology degree in embedded systems from the VIDYA VIKAS INSTITUTE OF TECHNOLOGY-JNTUH, he received the Bachelor Of Technology degree from VNR-VJIE, JNTUH. He is currently working as

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 02
February 2017

assistant professor in the Department of ECE with Madhira Institute Of Technology And Sciences, kodad. His interest subjects are Embedded Systems, Microprocessors, measurement and instrumentation Systems.
Email Id:subhanece15@gmail.com

Mr. P.SANDEEP was completed his B.Tech in Electronics and Communication Engineering at Madhira institute of technology and sciences, Kodad, Suryapet, Telangana. Currently he is pursuing his M.Tech in Electronics and Communication Engineering at Madhira institute of technology and sciences, Kodad, Suryapet, Telangana. His interests are image and video processing.
Email Id:sandeeppillutla33@rediffmail.com