J. Vis. Commun. Image R. 48 (2017) 110–121

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Combining multi-layer integration algorithm with background prior and label propagation for saliency detection ☆

Chenxing Xia [a], Hanling Zhang [a,b,*], Xiuju Gao [a]

[a] College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China
[b] Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science & Technology, Nanjing 210044, China

ARTICLE INFO

Article history:
Received 11 January 2017
Revised 13 May 2017
Accepted 16 June 2017
Available online 17 June 2017

Keywords:
Corner
Objectness
Energy function
Integration algorithm
Multi-layer

ABSTRACT

In this paper, we propose a novel approach to automatically detect salient regions in an image. Firstly, some corner superpixels serve as the background labels and the saliency of other superpixels are determined by ranking their similarities to the background labels based on ranking algorithm. Subsequently, we further employ an objectness measure to pick out and propagate foreground labels. Furthermore, an integration algorithm is devised to fuse both background-based saliency map and foreground-based saliency map, meanwhile an original energy function is acted as refinement before integration. Finally, results from multiscale saliency maps are integrated to further improve the detection performance. Our experimental results on five benchmark datasets demonstrate the effectiveness of the proposed method. Our method produces more accurate saliency maps with better precision-recall curve, higher F-measure and lower mean absolute error than other 13 state-of-the-arts approaches on ASD, SED, ECSSD, iCoSeg and PASCAL-S datasets.

1. Introduction

Saliency detection, aimed at identifying the most important and conspicuous object regions in an image, has attracted much attention in recent years. There are various applications for salient object detection, including image segmentation [1], object recognition [2], image compression [3], image retrieval [4], dominant color detection [5] and image forensics [6] etc.

Existing work of saliency detection can be roughly divided into two categories: top-down and bottom-up approaches. Top-down methods [7–9] are task-driven which generally require supervised learning with manually labeled ground truth. To better distinguish salient object from background, high-level information and supervised methods are incorporated to improve the accuracy of saliency map. The accuracy of the method is high while the operator is complex and has a slow speed. On the other hand, bottom-up methods [10–12] usually exploit low-level cues such as features, colors and spatial distances to construct saliency maps. Now more and more methods formulate their algorithms based on boundary prior, assuming that regions along the image boundary are more likely to be the background modeling the so-called center prior from [13,14]. However, it is not appropriate to put all nodes on the boundary into a class, which will inevitably lead to noise. Other representations based on the low-level features try to exploit the intrinsic textural difference between the foreground and background, including focusness [15], textual distinctiveness [16], and structure descriptor [17]. They perform well in many cases, but can still struggle in complex images.

Due to the shortcomings of low-level features, many algorithms have turned to take higher-level features into consideration. One type of higher-level representations that the notion of objectness is employed [18]. Jiang et al. [15] computed a saliency measure by combining the objectness values of many overlapping windows. Li et al. [19] devised a co-transduction algorithm to fuse both boundary and objectness labels based on an inter propagation scheme. Jia et al. [20] obtained high-level saliency prior with the objectness algorithm to find potential object candidates, which was used to combine with low-level appearance for saliency detection. However, using the objectness measure directly to compute saliency, it may not work well in complex scenes when the objectness score fails to predict true salient object regions. It is a better way to consider the scores as hints of the foreground when employing high-level objectness. In addition, some low-level features and high-level features may be not robust. In other words, some background cues are falsely selected as foreground informa-

☆ This paper has been recommended for acceptance by Zicheng Liu.

* Corresponding author at: College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China.

E-mail addresses: starry@hnu.edu.cn (C. Xia), starry1614@163.com (H. Zhang), s131010061@hnu.edu.cn (X. Gao).

http://dx.doi.org/10.1016/j.jvcir.2017.06.009
1047-3203/

tion, thus leading to inaccurate detection results. Hence, it is necessary to designed refinement mechanism to obtain better results with robust information. Furthermore, the accuracy of the saliency map is sensitive to the number of superpixels while salient objects are likely to appear at different scales. An integration algorithm should be developed to effectively take advantage of multiple saliency maps.

In this paper, we put forward a unified approach to incorporate low-level features and the objectness measure for saliency detection via label propagation. We observe that the corner cues can be used to estimate the appearance of the background while the objectness cues focus on the characteristics of the salient object. Firstly, based on a neoteric background prior selecting four corners of an image as background, corner-based background saliency map (CBSM) can be obtained by a unique affinity covered color, spatial and texture. Inspired by reverse-measurement methods to improve the accuracy of measurement in engineering, we take the objectness labels as foreground prior based on part of information of CBSM to construct objectness-based foreground saliency map (OFSM). Further, an original energy function is applied to optimize both of them respectively and a single-layer saliency map (SLSM) is formed by merging the above two maps. In addition, we propose a new threshold segment method by size prediction when using energy function. Finally, to deal with the scale problem, we obtain our multi-layer saliency map (MLSM) by presenting an integration algorithm to take advantage of multiple saliency maps.

In summary, the main contributions of our work include: (1) We proposed a novel background prior to construct a saliency map via a unique affinity. (2) A creative energy function is put forward to optimize the background prior map and foreground prior map before incorporation. (3) A new threshold segment method is proposed by size prediction when using energy function. (4) Mutil-layer Integration algorithm is proposed to integrate multiple saliency maps into a more favorable result.

A preliminary version of our model was described in [21]. This paper differs from [21] in several ways. First, edge information is incorporated when we define the graph affinity entry whereas [21] only considered color and spacial features. Besides, background-based saliency is obtained by manifold ranking based on the graph affinity entry while [21] defined the saliency of a superpixel as its color contrast to the corner superpixels. In addition, A new threshold segment method is proposed by size prediction when using energy function. More technical details are provides, showing the proposed method outperforms the previous method in [21].

2. Related work

Recently, numerous bottom-up saliency detection methods have been proposed, which prefer to generate the saliency map by utilizing the boundary information. In [22], the contrast against image boundary was used as a new regional feature vector to characterize the background. Zhu et al. [23] proposed a more robust boundary-based measure, which took the spatial layout of image patches into consideration.Yang et al. [24] selected the four boundaries of an image as background cues to get foreground queries via manifold ranking (MR). Perazzi et al. [25] weighted the initial prior map with boundary contrast to obtain the coarse saliency map. Inspired by previous academics, in this paper, we choose the information of four corners of an input image as background prior.

Generic object detection methods aim at generating the locations of all category independent objects in an image. We observe that object detection is closely related to saliency object segmentation. In [18], saliency was utilized as objectness measurement to generate object candidates. Chang et al. [10] used a graphical model to exploit the relationship of objectness and saliency cues for saliency object detection. Li et al. [26] predicted the saliency score of an object candidate via training a random forest model. In this work, we select the objectness labels as foreground prior.

In addition, there has recently been a growing interest in using diffusion processes to propagate saliency information throughout a graph. Harel et al. [27] used graph algorithms and a measure of dissimilarity to compute saliency in their graph-based visual saliency model. Yang et al. [24] cast saliency detection into a graph-based ranking problem. Jiang et al. [28] formulated saliency detection via an absorbing Markov chain on an image graph model. In [29], Lu et al. proposed a method for learning optimal seeds for object saliency using a diffusion process. Li et al. [30] proposed regularized random walks ranking, which consists of a pixel-wise graph term and a newly formulated fitting constraint to take local image data and prior estimation into account. In this paper, we choose manifold ranking as our propagation algorithm.

In many cases, different features (or multiview data) can be obtained, and how to efficiently utilize them is a challenge. In order to solve this problem, many multiple feature integration methods were proposed [31–34]. Yu et al. [31] developed a deep multi-task learning algorithm to jointly learn more representative deep convolutional neural networks and more discriminative tree classifier, so that they achieved fast and accurate detection of large numbers of privacy-sensitive object classes. A novel deep multimodal distance metric learning (Deep-MDML) method is developed to assign optimal weights for different modalities [32]. In [33], a high-order distance-based multiview stochastic learning (HD-MSL) is proposed, which effectively combines varied features into a unified representation and integrates the labeling information based on a probabilistic framework. Apart from detecting salient object in a single layer, salient object detection also has been extended to identifying common salient objects shared in multiple layers. Tong et al. [35] concluded six principles for effective saliency computation and fused them into a single framework via combining with Bayesian framework. Qin et al. [36] proposed multi-layer Cellular Automata to integrate multiple saliency maps into a better result under the Bayes framework. In [37], the final strong saliency map was obtained by the way of calculating the mean value. All of them achieve very good results demonstrating the effectivity of multi-layer in the accuracy of saliency detection.

3. The proposed approach

Fig. 1 shows the main steps of the proposed salient object detection algorithm. In this section, we give the details about our model. To better capture intrinsic structure information and improve computational efficiency, an input image is over-segmented at M scales. At any scale m, an image is segmented into N small superpixels by the simple linear iterative clustering (SLIC) algorithm [38]. Some important notations are presented in Table 1.

3.1. Contrast-based background saliency map (CBSM)

3.1.1. Affinity matrix construction

At first, the similarity of two nodes is measured by considering their color and distance. Based on the intuition that neighboring regions are likely to share similar appearances and that remote ones do not bother to have similar saliency values even if the appearance of them are highly identical. Our affinity between node i and node j is considered from the color and spatial characteristics. From the perspective of the color features, we define the affinity entry c_{ij} of node i to a certain node j as: