Neurocomputing 152 (2015) 359–368

Contents lists available at ScienceDirect

Neurocomputing

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

Saliency detection via background and foreground seed selection

Jianpeng Wang, Huchuan Lu*, Xiaohui Li, Na Tong, Wei Liu

School of Information and Communication Engineering, Dalian University of Technology, China

ARTICLE INFO

Article history:
Received 21 April 2014
Received in revised form
21 August 2014
Accepted 26 October 2014
Communicated by X. Li
Available online 27 November 2014

Keywords:
Saliency detection
Background
Foreground
Seed
Integration
Unification

ABSTRACT

In this paper, we propose a bottom-up visual saliency detection algorithm. Different from most previous methods that mainly concentrate on image object, we take both background and foreground into consideration. First, we collect background seeds from image border superpixels by boundary information and calculate a background-based saliency map. Second, we select foreground seeds by segmenting the first-stage saliency map via adaptive threshold and compute a foreground-based saliency map. Third, the two saliency maps are integrated by the proposed unified function. Finally, we refine the integrated result to obtain a more smooth and accurate saliency map. Moreover, the unified formula also proves to be effective in combining the proposed approach with other models. Experiments on publicly available data sets demonstrate that the proposed algorithm performs favorably against the state-of-the-art methods.

1. Introduction

Image saliency detection is an important preprocessing step in computer vision to reduce computational load by focusing on salient regions, which has been widely used in numerous applications including object detection [1] and recognition [2], image segmentation [3] and compression [4], video summarization [5], and so on. Saliency detection aims to estimate the probability of an image region appearing as part of foreground. Previous saliency detection approaches can be categorized as two main kinds: bottom-up and top-down. Top-down measures [6–8] learn task-driven models through training process, which requires specific prior knowledge. Bottom-up methods [9–12] also known as stimuli-driven models, mainly detect saliency by global or local contrast without any prior knowledge.

Recent years have witnessed significant progress in visual saliency detection. Itti et al. [9] present a saliency map by combining three feature maps including color, intensity and orientation at different scales. Liu et al. [13] integrate three saliency detection results including multi-scale contrast, regional and global distribution by the weights derived from Conditional Random Field (CRF) learning. Achanta et al. [14] measure saliency by the dissimilarity of each pixel's color with the average color over the entire scene, which is simple and efficient but suffers a great difficulty in handling cluttered background. Their saliency maps have low resolution and tend to highlight object edge rather than interior. Taking context information into consideration, Goferman et al. [15] implement local contrast by computing the dissimilarity of a region only with its relevant context, defined as a set of appearance-based similar regions. The global region contrast based method [10] utilizes spatially weighted color contrast to measure saliency, which uniformly highlights the interior object but is sensitive to object size. Saliency estimation is defined as a Bayesian inference model in [11,12]. Different from [11], Xie et al. [12] propose a coarse-to-fine strategy to roughly locate the object region by a convex hull and compute a prior map instead of the constant prior in [11]. Previous saliency detection models mainly focus on image object which is difficult to be dealt with due to its wide variation. However, prior knowledge based on background information could be more accurate and effective as reported by [16,17], which measure saliency based on background priors.

In this paper, we propose a novel saliency detection algorithm based on both background and foreground priors. First, we obtain a border set by collecting the image border superpixels, which have been proved to be good visual cues for background priors in saliency detection [16,17]. In addition, we remove the superpixels with strong image edges out of the border set to reduce the foreground noise (e.g., when objects appear at the image border) and thus obtain the background superpixel seeds. We calculate the color and spatial distances between each superpixel and the background seeds to obtain the background-based saliency map. Second, the foreground-based saliency detection is implemented based on the first-stage saliency map. We segment the map using an adaptive threshold and select the superpixels which are more likely to be the object as the foreground seeds. Then the foreground-based saliency map is computed by the color and spatial similarity with the foreground seeds, similar to the background-based saliency detection. Third, we

* Corresponding author.
E-mail address: lhchuan@dlut.edu.cn (H. Lu).

http://dx.doi.org/10.1016/j.neucom.2014.10.056
0925-2312/

propose a unified formula to integrate the two saliency maps: the background-based one which could highlight the objects and the foreground-based one which could suppress the background noises, therefore the unified map could benefit from both the two maps. Refined by saliency diffusion and Gaussian falloff, the final result could be more accurate. The main steps of the proposed method are shown in Fig. 1.

The main contributions of this work are three-fold:

1. By the use of image edge information, we remove the foreground noise in the image borders to obtain more stable and reliable background prior knowledge.
2. We devise two saliency detection models by computing the similarity of each image region with the selected background and foreground seeds, respectively.
3. The proposed unification mechanism achieves more favorable results and is demonstrated to be effective for saliency detection by further experiments on other algorithms.

The rest of this paper is organized as follows. In Section 2, we describe the details of our algorithm, including the background and foreground based saliency maps and saliency unification and refinement. Experimental results of the proposed algorithm and comparisons with several previous methods are presented in Section 3. Section 4 concludes this paper.

2. Our approach

In this section, we present the details of the proposed saliency measure. To better capture the structural information of an input image, we use superpixels as the minimum processing units, which are generated by the simple linear iterative clustering (SLIC) algorithm [18].

2.1. Background based saliency detection

2.1.1. Background seeds selection

Based on the observation that the object is likely to appear at or near the center of an image, we firstly extract the superpixels along the image borders as prior background regions. Nevertheless, there may be some foreground noises in the border regions, leading to negative effects on saliency detection. Therefore, we propose a mechanism based on image boundary information to remove the foreground noises and select background seeds from the border superpixels.

Suppose that an input image is over-segmented into N superpixels as shown in Fig. 2(b). The centroid location vector and mean CIELab color vector of the ith superpixel are denoted by $\mathbf{l}_i$ and $\mathbf{c}_i$, respectively. We choose the superpixels whose centroids locate within a certain number of pixels to the image borders (see Fig. 2(d)) to be the border set. Since the most distinct boundary of an image is likely to be the contour between the object and background, we can roughly remove the image superpixels with strong boundaries, which are regarded as the foreground noises, out of the border set.

We first adopt the probability of boundary (PB) [19] to detect image boundary (see Fig. 2(c)). The boundary feature of the ith superpixel is calculated by the average PB value of pixels along the edge contour of superpixel i, as follows:

$$PB_i = \frac{1}{|B_i|} \sum_{I \in B_i} I^{pb}, \tag{1}$$

where B_i is the edge pixel set of superpixel i and $|B_i|$ denotes its cardinality. The PB value of pixel I is denoted by I^{pb}. Since the superpixel with large boundary feature is more likely to be the object, we remove the superpixels whose boundary features are larger than the adaptive gray threshold derived by [20]. Then the

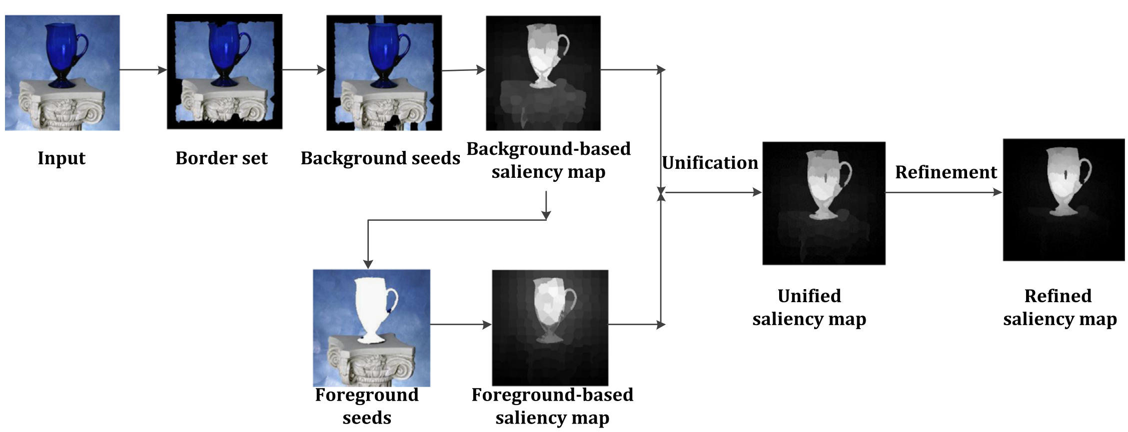

Fig. 1. Main steps of our approach, including the background and foreground seed selection, the two saliency maps, saliency unification and refinement.

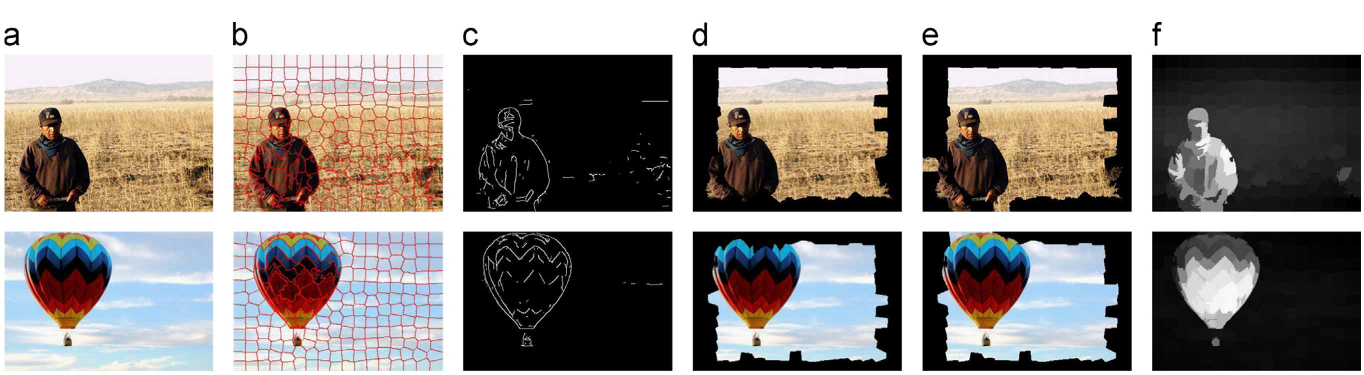

Fig. 2. From left to right: (a) Input image, (b) Superpixel, (c) PB map [19], (d) border set (the black regions along image borders), (e) background seeds (the black regions along image borders) and (f) background-based saliency map.