J. Vis. Commun. Image R. 33 (2015) 378–388

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Surroundedness based multiscale saliency detection ☆

Beiji Zou, Qing Liu, Zailiang Chen *, Hongpu Fu, Chengzhang Zhu

School of Information Science and Engineering, Central South University, Ministry of Education-China Mobile Joint Laboratory For Mobile Health, China

ARTICLE INFO

Article history:
Received 21 April 2015
Accepted 28 September 2015
Available online 23 October 2015

Keywords:
Saliency detection
Figure-ground segregation
The Gestalt principles
Image decomposition
Homogeneous region extraction
Surroundedness
Multiscale
Contour confidence

ABSTRACT

In this paper, a surroundedness-based multiscale saliency method is proposed based on the Gestalt principles for figure-ground segregation, which states that (1) surrounded regions are more likely to be perceived as figures, (2) the humans understand the external stimuli as whole rather than the sum of their parts. First, an image is characterized by a set of binary images, which is generated by a simple and effective homogeneous region extraction method with well contour preservation. And the contour confidence map is obtained by a fast contour detection method. Then for each connect homogeneous region in a binary map, surroundedness is defined by the average outer contour confidence. Finally, integrating the background priors, multiscale saliency maps are generated and combined to the final saliency map. The proposed method is evaluated on two widely used public datasets with pixel accurate salient region annotations using both precision and recall analysis and ROC analysis. And the experimental results show that the proposed method outperforms 14 alternative methods.

1. Introduction

Confronted with an overwhelming amount of visual information received from the natural world, humankind has evolved a remarkable ability to select relevant information out of irrelevant noise, in which attention as a selective process is the key, and turns looking into seeing [1]. Aiming at the same target in computer vision field, saliency detection is to predict what attracts the humankind most in an image via mimicking the human visual system such that the complexity of scene analysis for further processing can be reduced efficiently and effectively. It has been researched for a long time for its wide applications in image retargeting [2,3], image and video compression [4], object recognition [5], image classification [6] and picture collage [7].

Inspired by the second biologically-plausible architecture [8] and feature integration theory [9], Itti et al. [10] propose the first computational saliency detection model, in which local contrast is computed via center-surround difference operator to measure the saliency of each pixel. Thereafter, lots of saliency detection methods based on contrast including local contrast and global contrast are proposed, such as [11–16]. Though those methods have achieved amazing performance, local contrast based methods tend to highlight the borders of the salient objects while global contrast based methods always fail to generate clear saliency map when the image is cluttered.

Recent researches show that attention depends on figure-ground organization. That is to say that figures draw attention while shapes of the ground tend to be ignored. Therefore, inspired by the Gestalt principles, saliency cues, such as distinctness, uniqueness, background priors and surroundedness are proposed. For example, Scharfenberger [17] use the sparse texture distinctness as the saliency cue while Margolin et al. [18] propose to adopt the pattern distinctness in PCA coordinate via L1 norm and the color distinctness to other regions in CIE-LAB color-space via L2 distance. Jiang et al. [19] integrate the uniqueness, focusness and objectness to a saliency map. Wei et al. [20] present a geodesic saliency model, in which background priors including boundary and connectivity priors are first proposed to detect the salient objects. Thereafter, Zou et al. [21] extend the geodesic saliency model using multi-level connectivity. Zhang et al. [22] propose a novel Boolean map based saliency model, in which the surroundedness cue in a Boolean map is defined as a property of a connected region with a closed outer contour. Those methods highlight the salient object uniformly. However, the distinctness and uniqueness based methods fail to detect the salient objects when the background is cluttered while background priors and surroundedness based methods face challenges when the salient objects slightly touch the image borders since the image borders are assumed to be the background. Fig. 1 shows some failed examples of those methods.

To solve the notorious problems, we focus on how human visual system segregates the figure from the image background and find

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author. Fax: +86 0731 888877701.
E-mail addresses: qing.liu.411@gmail.com (Q. Liu), zailiangchencs@gmail.com (Z. Chen).

http://dx.doi.org/10.1016/j.jvcir.2015.09.017
1047-3203/

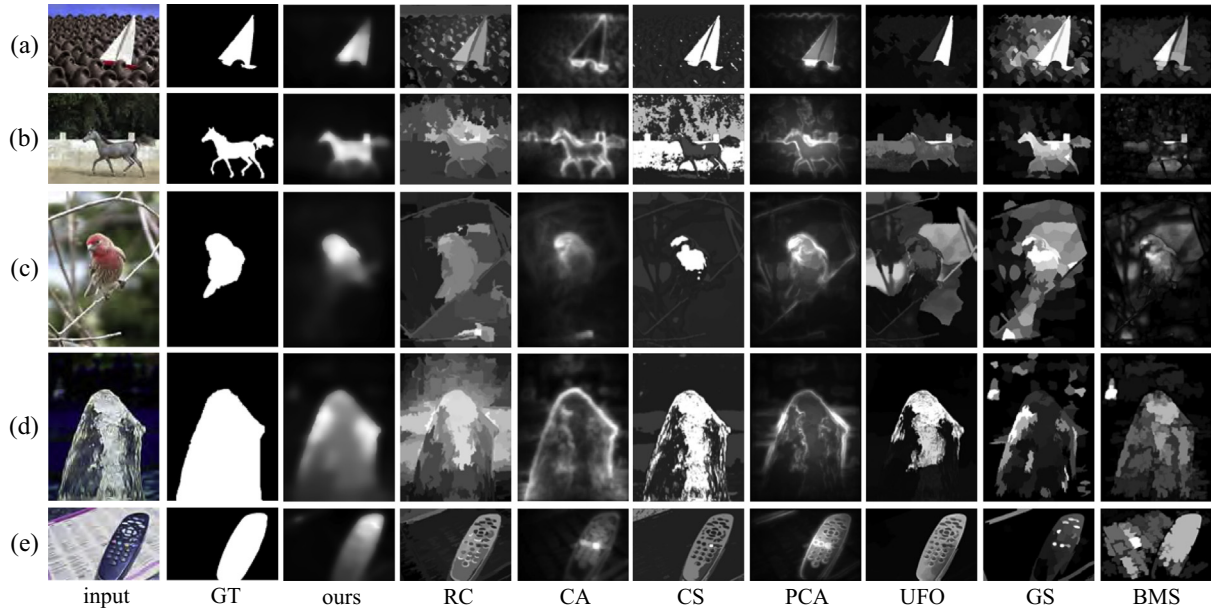

Fig. 1. Saliency maps obtained by different methods. From left to right: input images, ground truth images (GT), saliency maps by our method, saliency maps generated by contrast based methods RC [13], CA [15], CS [16], and saliency maps generated by methods using the Gestalt principles PCA [18], UFO [19], GS [20], BMS [22]. It can be seen that those methods are challenging for cluttered images or salient objects located on the image border.

foundations from the Gestalt principle. As is stated in [23–25], a surrounded region (e.g. the region labeled as 'salient region' in Fig. 2(c)) is more likely to be perceived as a figure while a background region always is the surrounding region (e.g. 'bkg region 1' in Fig. 2(c)). Furthermore, external stimuli is understood as whole rather than the sum of their parts. Meanwhile, we observe that: (1) surrounded region has strong outer contour since there always exists abrupt changes in features at locations that the surrounded region and its surrounding regions abut; (2) a surrounding region is inclined to have weak outer contour since the contour confidence at the image boundary is zero or the feature contrast to other regions is low; (3) the stronger the outer contour is, the more salient the region is.

Based on the above observations, a region-based saliency detection method is proposed in this paper, in which the surroundedness of a region is used as a saliency cue. By definition, the surroundedness of a region refers to its average outer contour confidence. Accordingly, the saliency of salient region located on the

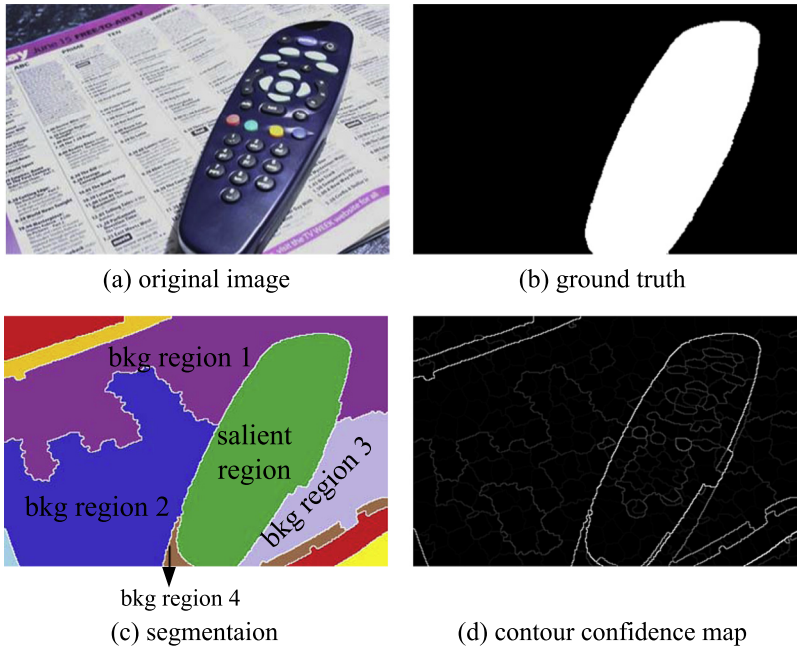

Fig. 2. Example of surrounded region and surrounding region. (a) Input image. (b) Ground truth image. (c) An ideal region proposal. (d) Contour confidence map. The region labeled as 'salient region' in (c) is surrounded region since it is surrounded by other regions, i.e. 'bkg region 1', 'bkg region 2', 'bkg region 3' and 'bkg region 4'. Furthermore, the 'salient region' has strong outer contour confidence as shown in (d) because of its high feature contrast to its surrounding regions while the surrounding regions has weak outer contour confidence because of their low feature contrast to their surrounding regions or zeros contour confidence on the image boundary.