Signal Processing: Image Communication 44 (2016) 69–83

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Co-saliency detection via inter and intra saliency propagation

Chenjie Ge [a], Keren Fu [a], Fanghui Liu [a], Li Bai [b], Jie Yang [a,*]

[a] *Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, China*
[b] *University of Nottingham, UK*

ARTICLE INFO

Article history:
Received 6 August 2015
Received in revised form
19 March 2016
Accepted 20 March 2016
Available online 24 March 2016

Keywords:
Co-saliency detection
Inter-saliency propagation
Intra-saliency propagation
Fusion

ABSTRACT

The goal of salient object detection from an image is to extract the regions which capture the attention of the human visual system more than other regions of the image. In this paper a novel method is presented for detecting salient objects from a set of images, known as co-saliency detection. We treat co-saliency detection as a two-stage saliency propagation problem. The first inter-saliency propagation stage utilizes the similarity between a pair of images to discover common properties of the images with the help of a single image saliency map. With the pairwise co-salient foreground cue maps obtained, the second intra-saliency propagation stage refines pairwise saliency detection using a graph-based method combining both foreground and background cues. A new fusion strategy is then used to obtain the co-saliency detection results. Finally an integrated multi-scale scheme is employed to obtain pixel-level co-saliency maps. The proposed method makes use of existing saliency detection models for co-saliency detection and is not overly sensitive to the initial saliency model selected. Extensive experiments on three benchmark databases show the superiority of the proposed co-saliency model against the state-of-the-art methods both subjectively and objectively.

1. Introduction

Visual saliency of the human or machine visual system serves as a filter for selecting a certain subset of visual information for further processing. The results of saliency detection are saliency maps which enhance salient objects while suppressing background objects. Saliency detection has applications in many fields including objects segmentation [1,2], content based image editing [3–5], image retrieval [6], image compression [7] and video summarization [8–10]. Existing saliency detection models however mainly detect salient objects from a single image [11–14], and the information of similar salient objects in a sequence of images is not exploited.

The concept of co-saliency has been proposed to select common salient objects from a sequence of images [15–21]. Co-saliency detection satisfies the following two properties: (1) co-salient regions should be salient regions in the image and (2) all co-salient regions from different images should share similar characteristics. Because co-saliency maps highlight similar foreground objects, they naturally can be used in many applications such as object co-segmentation [22], co-recognition [23] and common pattern discovery [24].

Most of current co-saliency models [16–19] split the co-salient object detection problem into single-image saliency detection and multi-image saliency detection, to discover what the salient object is within each image and how frequently the salient object occurs across the images. Various features are used such as texture descriptors [16], corresponding feature [17], and color histograms [18,19] to solve the co-saliency detection problem. Such low-level features are not enough to describe the properties of co-salient objects, leading to unsatisfactory co-saliency detection results. To overcome the difficulties, recently Li et al. [21] directly use single image saliency maps to find co-salient objects based on the fact that co-salient objects in an image sequence should also be salient in each image. In their method, single image saliency maps are exploited to highlight salient objects through stage-one manifold ranking. Thereafter, the foreground of each image is probed to find similar regions in other images through stage-two manifold ranking. The problem with the method is that it is unable to highlight the co-salient object and suppress the background information simultaneously. The co-saliency detection result is easily affected by inaccurate foreground maps obtained by stage-one manifold ranking. If the foreground and background have similar colors, background regions will be highlighted as well, leading to unsatisfactory detection results.

In this paper we propose a novel saliency propagation framework to fulfill the co-saliency detection task. Under this framework co-salient foreground and background cues are separately acquired to enhance the co-salient object. A preliminary conference version of our work has appeared in [25]. Unlike the

* Corresponding author.
E-mail address: jieyang@sjtu.edu.cn (J. Yang).

http://dx.doi.org/10.1016/j.image.2016.03.005
0923-5965/

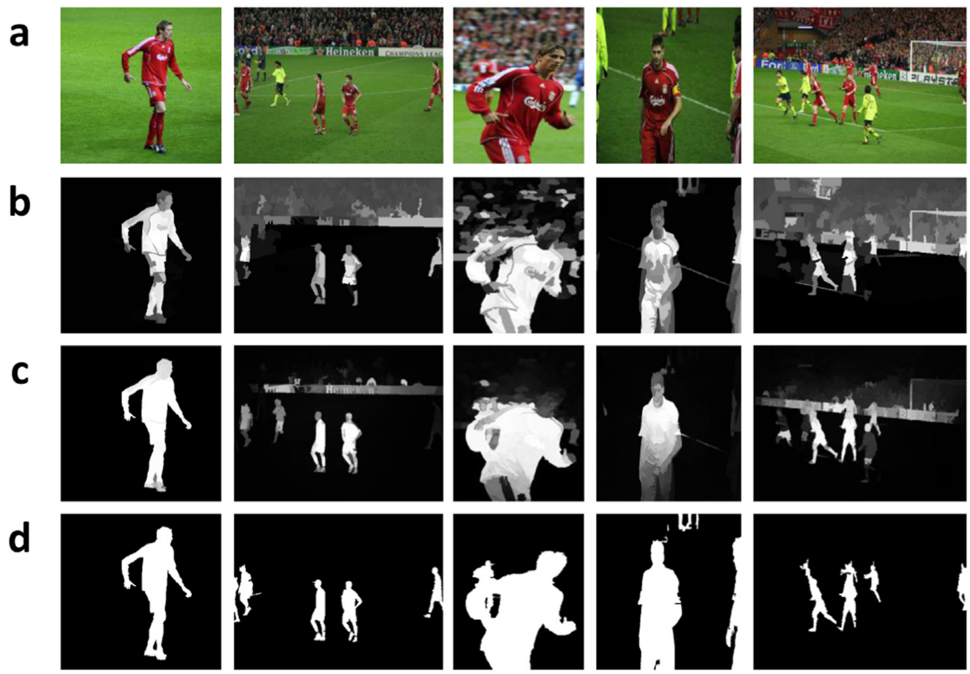

Fig. 1. The performance of [21] and ours. (a) source images; (b) results of [21]; (c) results of our method; (d) ground truth.

method described in [21], the proposed inter-saliency propagation uses superpixel similarity between two images to obtain the foreground cue with the help of a single image saliency map. This is more effective for characterizing objects and preserving edges than pixel-level ranking. Background cue is obtained independently from each image considering both the background connectivity and saliency mask, followed by a graph-based intra-saliency propagation combining both foreground and background cues with a new edge constraint. As shown in Fig. 1, the red soccer players are co-salient objects among all the five images and the proposed model surpasses [21] in both foreground enhancement and background suppression.

The main contributions of this paper are as follows:

- A novel inter-saliency propagation method is proposed to transmit saliency values between two images to find the co-salient foreground cue.
- A new intra-saliency propagation method is presented which simultaneously highlights co-salient objects, suppresses background information and smooths saliency values with edge constraint.
- A new fusion strategy is proposed combining the intra-saliency propagation maps to obtain the final co-saliency map, adaptively weighed by a rough co-saliency approximation.

The paper is organized as follows. We describe the related work in Section 2. The proposed approach is presented in Section 3. Experimental results are shown in Section 4 and finally we conclude in Section 5.

2. Related work

Saliency detection methods can be divided into two major categories: human fixation prediction [26,27] and salient object detection [28–30,11,31–35]. Fixation prediction models usually simulate how human attention focus on an image. This produces a saliency distribution map based on the focus points of the eye. Salient object detection models focus on detecting salient regions of the whole image. The outcome saliency map reflects the saliency probability of each pixel, which can be regarded as a soft segmentation of salient objects.

Saliency detection from a single image has been an active research area for decades. Itti et al. [36] propose a neural network based saliency model integrating three features and multiple scales. Zhai et al. [37] introduce the image histograms in luminance channel to measure saliency. Pixel-level saliency is computed by the luminance contrast. Hou et al. [28] utilize the spectral residual in the amplitude spectrum of the Fourier transform to compute saliency. Guo et al. [38] propose that the phase spectrum of Fourier transform instead of the amplitude transform is the key in obtaining the location of salient areas. Achanta et al. [39] propose to define pixel saliency using color difference from the image average color, which is equivalent to combining center-surround differences from different scales. Goferman et al. [5] use local and global features to estimate the patch saliency in multi-scales. Visual organizational rules and high-level factors are also considered to enhance the saliency map. Cheng et al. [29] propose histogram and region contrast methods to enhance salient objects. In [30], Perrazzi et al. propose to combine the uniqueness of locally constrained regions and the spatial distribution of regions. Fu et al. [11] present a comprehensive salient object detection system taking advantage of both color contrast and color distribution. Most of the methods mentioned above are simple to implement but cannot always highlight all pixels on the salient objects when the background is complex. By considering the connectivity of background regions, Wei et al. [31] use the shortest path towards the boundary to define the saliency value for each region. Such boundary prior is based on the assumption that boundary parts of an image will most likely belong to the background. Later the graph-based manifold ranking [32] is proposed to compute region saliency according to its relevance to boundary patches. Yan et al. [33] present a hierarchical saliency method which merge regions according to user-defined scales to eliminate small-size distracters. A novel learning based method [34] exploits a random forest regressor to map multiple features to the saliency value of the