Signal Processing: *Image Communication* 38 (2015) 115–126

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Depth-aware salient object detection using anisotropic center-surround difference

Ran Ju, Yang Liu, Tongwei Ren, Ling Ge, Gangshan Wu *

State Key Laboratory For Novel Software Technology, Nanjing University, China

ARTICLE INFO

Available online 11 July 2015

Keywords:
Salient object detection
Depth map
Center-surround difference

ABSTRACT

Most previous works on salient object detection concentrate on 2D images. In this paper, we propose to explore the power of depth cue for predicting salient regions. Our basic assumption is that a salient object tends to stand out from its surroundings in 3D space. To measure the object-to-surrounding contrast, we propose a novel depth feature which works on a single depth map. Besides, we integrate the 3D spatial prior into our method for saliency refinement. By sparse sampling and representing the image using superpixels, our method works very fast, whose complexity is linear to the image resolution. To segment the salient object, we also develop a saliency based method using adaptive thresholding and GrabCut. The proposed method is evaluated on two large datasets designed for depth-aware salient object detection. The results compared with several state-of-the-art 2D and depth-aware methods show that our method has the most satisfactory overall performance.

1. Introduction

An inherent and powerful ability of human eye is visual attention, which quickly captures the most conspicuous regions from a scene, and passes them to high level visual cortexes. The attention selection reduces the complexity of visual analysis and thus makes human visual system considerably efficient in complex environments. Computational saliency models, which follow the attention mechanism of human eye, occupy an important place in image processing and computer vision society. By saliency analysis, vision tasks are concentrated on a few regions of interests instead of entire images, which benefit many applications in both efficiency and effectiveness, e.g. image classification [1,2], object segmentation [3–6], image retargeting [7–10], adaptive image compression [11,12], content-based image retrieval [13,14] and quality assessment [15,16].

There are generally two categories of visual saliency models: fixation prediction [17] and salient object detection [18]. The former model aims at predicting the gaze points, i.e. where people look. The prediction results usually highlight a few spots people are most likely to pay attention to. In contrast, the goal of salient object detection is to detect the entire object that appears most distinctive, which could be more useful for high level processing. In this paper, we focus on the latter model, salient object detection. We give an example in Fig. 1. Given a color image with its depth map (1st row), our goal is to produce a high resolution saliency map, and extract the object of the greatest interest (2nd row), which is Statue of David in this example.

Most previous works on salient object detection focus on 2D color images. However, human perceives the world with not only color but also abundant 3D spatial information. While the saliency researches on 2D color images have been studied a lot and remarkable achievements have been made, little attention is paid to the effect of scene depth for saliency analysis. One reason is that depth

* Corresponding author.
E-mail address: gswu@nju.edu.cn (G. Wu).

http://dx.doi.org/10.1016/j.image.2015.07.002
0923-5965/

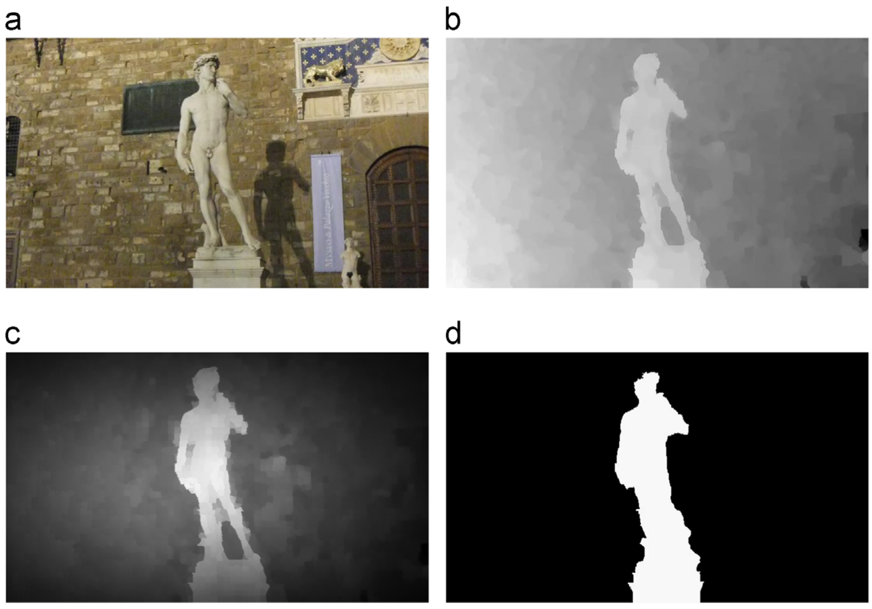

Fig. 1. Salient object detection. (a) Original color image. (b) Depth image of (a). Pixels appear brighter are nearer, vice versa. (c) Saliency map generated by the proposed method. Higher brightness indicates the location is more conspicuous. (d) Salient object mask produced by our method. White indicates object and black indicates background.

data is hard to capture as the devices are expensive in the past. Fortunately, more and more cheap and convenient 3D devices have been developed in recent years, which makes it easier to obtain sufficient depth data to support the research on depth-aware salient object detection. The depth data, no matter captured by stereo [19], ToF (Time-of-Flight) [20] or structured light [21], is usually recorded as a depth map like that shown in Fig. 1(b). Recently, a few tentative researches have shown that depth could be powerful in saliency analysis. For example, Lang et al. [22] integrated depth prior into existing 2D methods and achieved a 6–7% increase in predictive power. Niu et al. [23] combined depth contrast and stereo photography knowledge for salient object detection. The result achieved the best evaluation score in their stereo saliency dataset. Inspired from all the above, we further explore the characteristics of depth information for salient object detection according to the following considerations [24]:

Good detection. A good salient object detection method should both miss real salient regions and falsely detect background as salient regions in a low probability. To this end, we propose to measure pixel-wise saliency according to a basic observation: a salient object usually outstands from its surroundings or background from a global view. This assumption is common for natural scenes and daily life, since generally objects in clutters or occluded by background are unlikely to be the most attention-grabbing. To get the contrast between a pixel and its surroundings, we search for background samples in several directions, and measure the center-surround difference as its saliency value. The background samples are estimated using the depth prior. We also set a few weighting parameters to highlight the contrasts in certain directions. Besides, as is well known and usually used in previous works, salient regions are usually nearer to viewers in 3D space and tend to locate at the center of an image. We employ the 3D spatial prior in a probabilistic manner to improve the detection results.

High resolution. Salient object detection methods should produce full resolution saliency maps and retain clear object boundaries. To this end we perform saliency computation in the superpixel granularity, which could both retain the object boundaries and avoid the noises of depth images.

Computational efficiency. Salient object detection is usually regarded as an early process for vision tasks. High efficiency is required to allow large scale, highly complex computation of following process. Owing to the simple but effective feature we used, our method works very fast and the computational complexity is linear to the image resolution.

For evaluation we compare our method with 8 state-of-the-art 2D methods and 5 representative depth-aware methods. We first built a benchmark including 2000 stereo images with computer generated depth maps and manually labeled groundtruth, which is the largest at present to our knowledge. For a comprehensive evaluation we also compare our method with the others on another large dataset [25] including 1000 RGBD images and groundtruth. In both datasets our method shows superior overall performance to the others. We have also developed an object segmentation method using adaptive thresholding and GrabCut [26]. The segmentation method produces highly accurate binary maps, which shows that our method is competent for salient object detection task.

The contribution of this paper can be briefly stated as follows:

- We proposed a new depth feature for salient object detection, which combines both depth based background