Neurocomputing 159 (2015) 1–8

Contents lists available at ScienceDirect

Neurocomputing

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

Global feature integration based salient region detection

Mingqiang Lin, Chenbin Zhang, Zonghai Chen*

Department of Automation, University of Science and Technology of China, Hefei 230027, PR China

ARTICLE INFO

Article history:
Received 23 September 2014
Received in revised form
29 November 2014
Accepted 4 February 2015
Communicated by Mingli Song
Available online 24 February 2015

Keywords:
Saliency detection
Feature integration
Conditional Random Field
Gaussian model

ABSTRACT

The goal of saliency detection is to locate the regions which are most likely to capture human's attention without prior knowledge of their contents. Visual saliency detection has been widely used in image processing, but it is still a challenging problem in computer vision. In this paper, we propose a salient region detection algorithm by integrating global features namely uniqueness and spatial distribution. Two measures of contrast are computed in pixel and superpixel level respectively. In order to suppress background noise, Low-level features are refined by High-level priors which are computed with the Gaussian model based on salient region. We formulate salient region detection as a binary labeling problem that separates salient region from the background. A Conditional Random Field is learned to effectively combine these refined features for salient region detection. Experimental results on the large benchmark database demonstrate the proposed method performs well when against fifteen state-of-the-art methods in terms of precision and recall.

1. Introduction

Human visual has the ability of quick search for their most interest in a scene. This capability is called visual attention. Visual saliency, being closely related to how we perceive and process visual stimuli, is investigated by multiple disciplines including cognitive psychology [1,2], neurobiology [3,4], and computer vision [5,6]. With the continuous development of cognitive psychology and neurobiology, based on the research of the mechanism of visual, Itti et al. found that human visual on the selectivity of the target in the scene can be divided into two manners [7]: a rapid, bottom-up, saliency-driven, task-independent manner and a slower, top-down, volition-controlled, task-dependent manner. Saliency detection is widely used in many computer vision applications including image segmentation [8], object recognition [9], adaptive compression of images [10], content-aware image editing [11], image retrieval [12], object detection [38], object tracking [39] and image quality assessment [40].

Visual attention is driven by the stimulation from unconscious low-level image features. Therefore, most existing algorithms for saliency detection are based on the bottom-up model. On the basis of the model simplifying, the researchers found that contrast is the most important factor which dominantly influences human visual attention [13,14]. According to the different contrast calculation area, in this paper, the method is broadly divided into two major categories of local and global.

Local methods evaluate saliency of an image region using its contrast with respect to local neighborhoods. Motivated by the neural architecture of the early primate vision system, Itti et al. [5] define image saliency using central-surround differences of image features. Harel et al. [15] use graph algorithms to non-linearly combine local uniqueness maps from different feature channels to achieve efficient saliency computation. Ma and Zhang [13] compute local contrast in a fixed neighborhood pixels for generating saliency maps, which is then extended using a fuzzy growth model, and provide three-level attention analysis. Goferman et al. [6] propose context-aware saliency based on four principles of human visual attention. Such methods using local contrast tend to produce higher saliency values near edges instead of uniformly highlighting salient objects (see Fig. 1).

Global methods evaluate saliency of an image region using its contrast with respect to the entire image. Zhai and Shah [16] define a pixel's contrast to all other pixels. Cheng et al. [14] extend the histogram to 3D color space, and propose the global region contrast with respect to the entire image. Liu et al. [17] through the statistics of the whole image color spatial-distribution, calculate contrast by using the variance of distribution. Perazzi et al. [18] decompose an image into perceptually homogeneous elements and define contrast measure based on the spatial distribution of those elements. Hou et al. [19] extract the spectral residual of an image in spectral domain and propose a fast method to construct the corresponding saliency map in spatial domain. Achanta et al. [20] analyze the spatial frequencies in the original image, and use DOG operator to achieve continuous band pass filters. Margolin et al. [35] analyze the inner statistics of patches in the image, then integrate pattern

* Corresponding author. Tel.: +86 55163606104.
E-mail address: chenzh@ustc.edu.cn (Z. Chen).

http://dx.doi.org/10.1016/j.neucom.2015.02.050
0925-2312/

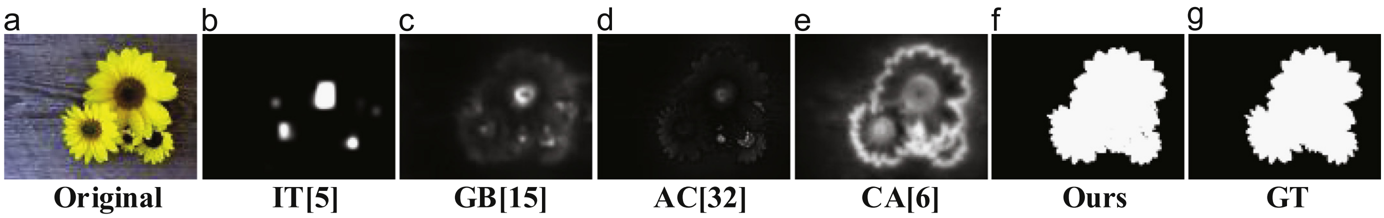

Fig. 1. Saliency maps computed by different state-of-the-art local methods. We compare to visual attention measure (IT [5]), graph-based saliency (GB [15]), salient region detection (AC [32]), context-aware saliency (CA [6]) and ground truth (GT).

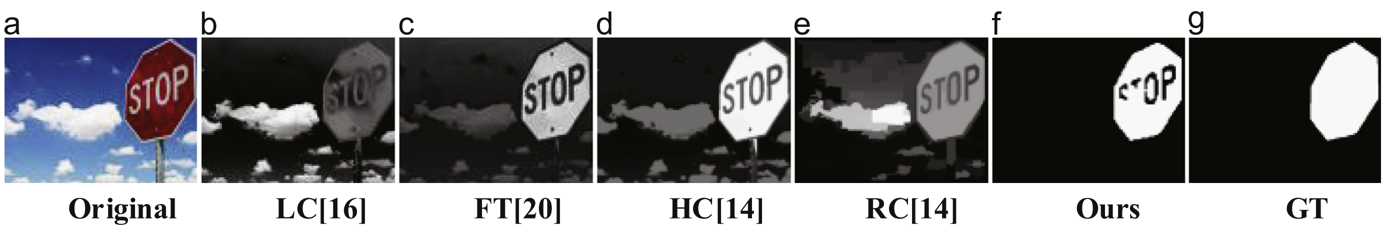

Fig. 2. Saliency maps computed by different state-of-the-art global methods. We compare to spatiotemporal cues (LC [16]), frequency-tuned saliency (FT [20]), global-contrast saliency (HC [14] and RC [14]) and ground truth (GT).

and color distinctness by PCA. These methods based on the global contrast can produce uniformly highlighting salient regions. However, these methods ignore spatial relationships across image parts, and may highlight background regions as salient (see Fig. 2).

Researchers designed the new features. Meanwhile, many feature fusion methods have been introduced. Itti et al. [5] combine colors, intensity and orientations features into a single topographical saliency map. Then, a WTA (Winner-Take-All) competition is employed to select the most conspicuous image locations as attended points. Song et al. [41] integrate both the low-level saliency and the high-level saliency by a linear model, where the weights that correspond to the low-level saliency and the high level saliency are learned from the eye-tracking data set. Perazzi et al. [18] use color and position information to compute two measures of contrast that are the uniqueness and distribution, then fuse these features by two Gaussian filters. The works in [22] propose a saliency detection algorithm by integrating local and global saliency detection in both RGB and Lab color spaces. Inspired by recent advances in machine learning, Judd et al. [21] formulate saliency detection as a binary classification problem. SVM is used to train a bottom-up, top-down model of saliency based on low, mid and high-level image features. Liu et al. [17] propose a set of features including multi-scale contrast, center-surround histogram, and color spatial-distribution. These features are combined for salient object detection by Conditional Random Field. Recently, Jiang et al. [23] regard saliency estimation as a regression problem. The supervised learning approach is used to map the regional feature vector to a saliency score. The works in [24] consider image boundary as a background prior, then reconstruct the entire image by dense and sparse appearance models. The pixel-level saliency is computed by a Bayesian integration of multi-scale reconstruction errors. Considering small-scale high-contrast patterns in an image could be adversely affected in detection accuracy, Yan et al. [25] propose a hierarchical model, to analyze saliency cues from multiple levels of structure, and then integrate them to infer the final saliency map. In [36], salient region detection is achieved by maximizing a submodular objective function which combines High-level priors and Low-level features. Liu et al. [37] generate the initial saliency tree by integrating with three features, and analyze a systematic saliency tree to obtain the pixel-wise saliency map. Li et al. [42] extract shape feature by the image segmentation method. This feature is combined with fixation distribution for salient object detection by random regression forest.

Local methods tend to produce higher saliency values near edges instead of uniformly highlighting salient region. However global methods ignore spatial relationships across image parts, and may highlight background noise regions as salient. In order to obtain uniformly highlighting salient region, and to reduce background noise mistakenly detected, in this paper, we propose a global feature integration based salient region detection algorithm. We calculate two measures of the global contrast which are the uniqueness and the spatial distribution in pixels and superpixels levels. Computing global contrast based on pixels can better retain the boundary information of salient region. Computing global contrast based on superpixels can preserve relevant structure information, and ignore unnecessary texture information. In order to suppress background noise, the object prior map is computed with the Gaussian model based on salient region. Low-level features are refined by High-level priors. We propose a supervised approach for salient region detection, which is formulated as an image segmentation problem using a set of global salient features. A Conditional Random Field is learned to effectively combine these features for salient object detection. Our work differs from learning a model to segment foreground out from background on two points. First we use a set of global salient features. Second we segment salient region out from background for salient region detection. Most of the works in computer vision focus on one of the following two specific tasks of saliency: fixation prediction and salient objects segmentation [42]. Our work can be classified as salient objects segmentation. Salient region detection also can be used in image segmentation. And the work of Judd et al. [21] which formulate saliency detection as a binary classification problem is related to our feature fusion method. We have extensively evaluated our methods on publicly available benchmark data sets, and compared our methods with state-of-the-art saliency methods. The experiments show significant improvements over previous methods both in precision and recall rates.

2. Algorithm overview

In this paper, we formulate salient region detection as a binary labeling problem that separates salient region from the background. We propose a salient region detection algorithm by integrating global features namely uniqueness and spatial distribution. Our algorithm consists of the following steps (see Fig. 3). First calculate the corresponding global feature saliency maps. Second integrate these saliency maps by utilizing CRF, and achieve initial salient regions. Then compute the object prior map by the Gaussian model based on initial salient regions. Low-level saliency maps are refined