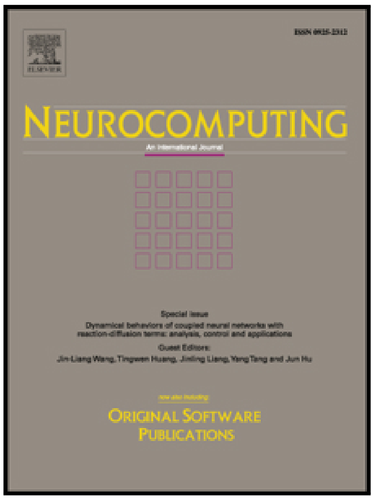

Attention Region Detection Based on Closure Prior in Layered Bit Planes

Jiazhong Chen , Bingpeng Ma , Hua Cao , Jie Chen , Yebin Fan , Tao Xia , Rong Li

PII: S0925-2312(17)30690-2
DOI: 10.1016/j.neucom.2017.04.020
Reference: NEUCOM 18353

To appear in: *Neurocomputing*

Received date: 28 January 2016
Revised date: 11 February 2017
Accepted date: 4 April 2017

Please cite this article as: Jiazhong Chen , Bingpeng Ma , Hua Cao , Jie Chen , Yebin Fan , Tao Xia , Rong Li , Attention Region Detection Based on Closure Prior in Layered Bit Planes, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.04.020

Attention Region Detection Based on Closure Prior in Layered Bit Planes

Jiazhong Chen[1*], Bingpeng Ma[2], Hua Cao[1], Jie Chen[3], Yebin Fan[1], Tao Xia[1], and Rong Li[1]

1 School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, 430074, China
2 School of Computer and Control Engineering, University of Chinese Academy Science, Beijing, 100190, China
3 International School, Jinan University, Guangzhou, 510660 China

Abstract

An attention region usually possesses a closed contour and a significant color contrast against its surroundings. Inspired by this observation, a novel prior so called closure prior is introduced as an important high-level saliency cue in this paper. Specifically, firstly, the connectivity analysis is used in layered bit planes to extract closed regions based on closure prior. Secondly, the closed regions touching image boundaries are removed based on the assumption that most photographers will not crop attention regions along the view frame. Thirdly, based on the hypothesis that an image region should have a big contrast against its surroundings if it has more chances in all bit planes to be a closed region, the contrast contributions of a closed region in all bit planes are accumulated to obtain color contrast. Meanwhile, by taking account for the characteristics of human visual system according to the perception of small attention region, and the mechanism of human visual resource allocation between attention regions, several morphological filtering technologies are applied to the steps of color contrast calculating. Finally, the saliency map aiming for attention region detection is generated from the contrast map. The experimental results show the presented detection method achieves acceptable performance compared with twelve state-of-the-art models.

Keywords:
Attention region; Closure prior; Saliency; Contrast; Bit plane

1. Introduction

Visual attention is a dominant characteristic of human vision system (HVS). Visual saliency measures low-level stimuli that the human vision system perceives as the visual attention [1]. Most of the previous methods can be classified into two basic categories depending on the definition way of saliency cue: contrast prior and background prior [2]. Itti *et al.* extract center-surround contrast at multiple spatial scales to find prominent regions [3]. Itti's model contains two parts: feature extraction and feature fusion. The feature extraction often considers what features will attract human's attention. Generally, intensity, color, orientation, structure, and motion are considered as the most attractive features. The feature fusion computes feature maps of those features using a center-surround operator across different scales, and performs normalization and summation to generate the saliency map. The approach proposed by Achanta focuses on global features [4]. Evaluated by the distance between the average feature vectors of the pixels of an image sub-region with the average feature vector of the pixels of its neighborhood, the local contrast of an image region with respect to its neighborhood at various scales is used to generate saliency maps, and these individual maps are combined to generate the final saliency map. Similarly, the authors of [5] propose an algorithm for context-aware saliency detection. The idea of their algorithm is that attention regions are distinctive with respect to both their local and global surroundings.

Some models are established in frequency domain to measure the contrast. Schauerte *et al.* introduce the eigenaxes and eigenangles for spectral saliency models that are based on the quaternion Fourier transform [6]. Li *et al.* propose a frequency domain paradigm which permits the full use of global information by defining the concept of non-saliency [7]. Unlike the approaches which model saliency as a global phenomenon, Hou *et al.* [8] provide an approach to a figure-ground separation problem using a binary, holistic image descriptor, which is defined as the sign function of the Discrete Cosine Transform of an image.

García-Díaz *et al.* argue that a natural approximation to image pixel distinctiveness can be done by computing the covariance matrix but not the variances in a representative coordinate system. Thus they propose a model grounded on classical hierarchical decomposition of images and measure the contrast by the statistical properties of images [9]. Specifically, their model first separates chromatic components, and next performs on each of components a multiscale and multioriented decomposition. Then an initial whitening is applied on chromatic components by the covariance matrix in coordinates, and a further whitening is imposed to groups of oriented scales for each whitened chromatic component to provide a retinotopic measure of local feature contrast. Erdem *et al.* incorporate the first-order statistics of images to their saliency model by using the covariance matrices of local features to represent image regions [10]. Then they exploit the distance between their covariance matrices in saliency estimation.

The contrast also can be measured by learning approaches. Judd *et al.* extract a set of low, mid and high-level features to define salient locations and adopt a linear support vector machine to train a saliency model [11]. Since natural images can be sparsely represented by a set of localized and oriented filters, Borji *et al.* propose a framework that measures patch rarities in each color space by local contrast and global contrast to estimate saliency on a patch-by-patch basis [12], and Riche *et al.* propose a saliency prediction model, which selects information worthy of attention based on multi-scale spatial rarity [13]. Recently, Wang *et al.* fuse the state-of-the-art saliency models at the score level in a para-boosting learning fashion to construct a saliency model [14], while Xia *et al.* construct a deep autoencoder-based network

* Corresponding author. Tel.: +86-27-87541764; fax: +86-27-87793022; e-mail: chenjz70@163.com, jzchen@hust.edu.cn (J. Chen). This work was supported in part by the Natural Science Foundation of China under Grant NSFCU1536203 and NSFC61300140.