Pattern Recognition 48 (2015) 1315–1327

Contents lists available at ScienceDirect

Pattern Recognition

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

Study of visual saliency detection via nonlocal anisotropic diffusion equation

Xiujun Zhang [a], Chen Xu [b,*], Min Li [c], Robert K.F. Teng [a,d]

[a] College of Information Engineering, Shenzhen University, Shenzhen 518060, China
[b] Institute of Intelligent Computing Science, Shenzhen University, Shenzhen 518060, China
[c] College of Mathematics and Computational Science, Shenzhen University, Shenzhen 518060, China
[d] Department of Electrical Engineering, California State University, Long Beach, 90840 CA, USA

ARTICLE INFO

Article history:
Received 19 June 2014
Received in revised form
26 September 2014
Accepted 9 October 2014
Available online 22 October 2014

Keywords:
Saliency detection
Boundary prior
Saliency seeds
Anisotropic diffusion
Nonlocal operators

ABSTRACT

In this paper, a nonlocal anisotropic diffusion equation is constructed to model the evolution of visual saliency. A new multi-directions discretization scheme is adopted to solve the equation. Visual attention diffusion is modeled as a series of diffusion progresses until achieving a stable status. So two stages of diffusion are involved. In the first stage, image boundaries are used to guide the background iterative diffusion process, and the initial saliency seeds are obtained by sorting. In the second stage, an optimization of saliency seeds is executed, and the good saliency seeds are strengthened and updated during each iteration. The final saliency map is obtained through iteratively diffusing saliency scores from the optimal saliency seeds (i.e. the most representative salient elements). Extensive experimental results on two large benchmark databases demonstrate the effectiveness of the proposed method.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The task of visual saliency detection is to find out the most attractive and informative part of a scene. It has been used in many computer vision fields, such as object detection and recognition [1,2], image segmentation [3,4], video compression and coding [5,6], motion detection [7] and video surveillance [8]. According to their mechanisms of representing image saliency, existing methods can be roughly divided into two categories: bottom-up methods (data-driven) and top-down methods (task-driven). The top-down methods require huge prior knowledge about visual content, which is often obtained by training with class labels. While the bottom-up methods are mainly driven by image features. Many research results [9,10] show that the most important feature in bottom-up visual saliency is contrast, including local and global contrasts. In general, local methods rely on pixel/region difference in the vicinity, while global methods rely mainly on color uniqueness in terms of global statistics [11]. In this paper, the bottom-up salient object detection method is focused.

As one of the seminal works, Itti et al. [12] propose a biologically plausible model, which computes a saliency map by combining color, intensity and orientation features at multiple scales. Since then, all kinds of theories have been introduced into this area, and various models have been proposed, including cognitive models [12,13], Bayesian models [14–16], decision theoretic models [17–19], information theoretic models [20,21], graphical models [22–24], spectral analysis models [25–27], pattern classification models [28] and other models [3,29–31]. Although so many models have been proposed, the mathematical or physical analysis of saliency detection problem are still not enough. As known to all, the saliency detection result, called a saliency map, is a probability map. The saliency score of one point in an image is more higher, the probability which point belongs to saliency region is more larger. So probability estimation under the Bayesian framework has been widely used for saliency detection problem. For the purpose of understanding the nature of saliency detection from multiple perspectives, in this paper, a novel and unique mathematical or physical interpretation is presented via diffusion equation. The diffusion equation is an important mathematical approach, and has been successfully applied into many fields, such as image restoration, image smoothing and image inpainting. Because through diffusion equation, all of these problems can be given a reasonable mathematical or physical explanation, the excellent results have been achieved in these fields. But the related studies of saliency detection via this useful method are rare. Inspired by this idea, in this paper, a novel method is proposed through incorporating saliency detection task with anisotropic diffusion process.

Firstly, four image boundaries are observed, and parts of them often belong to the background. According to this cue, the color differences between the superpixel and its neighbors are used to

* Corresponding author. Tel.: +86 18924641918; fax: +86 0755 26538959.
E-mail addresses: zhangxiujunszu@gmail.com (X. Zhang), xuchenszu@gmail.com (C. Xu), limin800@szu.edu.cn (M. Li), robert.teng@csulb.edu (R.K.F. Teng).

http://dx.doi.org/10.1016/j.patcog.2014.10.016
0031-3203/© 2014 Elsevier Ltd. All rights reserved.

guide the nonlocal anisotropic diffusion. The background saliency scores are gradually diffused from boundary prior to the entire image. The iterative diffusion progress is stopped when it reaches a stable status. The first stage saliency map is obtained by this way. The most representative salient elements (i.e. the saliency seeds) can be selected via sorting the saliency scores of all superpixels by descend. Another iteration of diffusion from saliency seeds to the whole image is executed to generate the final visual saliency map. During the iteration, the effects of good seeds are strengthened continuously, which constantly correct the saliency seeds of the previous stage, and ensure the well performance of the proposed model.

The main contributions of the study are summarized in the following three areas:

- A novel nonlocal anisotropic diffusion method is provided to describe the visual attention transfer process.
- Instead of using Harris operator and estimating a convex hull to obtain the background prior, the boundary prior is exploited to diffusion from boundary to the entire image. It still presents a very promising performance in any images of complicated background.
- It is firstly proposed that visual attention spreading is a series of diffusion progresses until achieving a stable status, rather than a one-off action. Then an iteration of diffusion is used to formulate the visual attention spreading progress. Through strengthening the effect of good seeds during the iteration, the background is relatively suppressed, and a model with satisfactory performance is obtained.

The rest of this paper is organized in the following sections. The related works that lead to this proposed study are reviewed and discussed in Section 2. The proposed study of saliency detection method is presented in Section 3. Experimental results and comparisons are shown in Section 4. Finally, the conclusions and future work are drawn in Section 5.

2. Related works

A survey and benchmark study of saliency detection is provided in [32]. Below, some representative models are briefly reviewed. After that, the differences between the proposed model and some similar works are also discussed.

Based on the information theory, Bruce et al. [33,20] express visual saliency by a local likelihood of each image patch decomposed via the filters learned from natural images. Harel et al. [22] introduce the graph theoretic ideas to the saliency detection task, and give high saliency values to the nodes, which are at the center of the image. From the viewpoint of the frequency domain, Hou and Zhang [25] propose the spectrum residual approach based on natural image statistics. Further more, Guo et al. [26] demonstrate that incorporating the phase spectrum of Quaternion Fourier Transform (PQFT) is better than only amplitude transformation in saliency detection task. Instead of local contrast, Cheng et al. [29] consider the global region contrast and spatial relationships with respect to the entire image. Different from aforementioned models, in [3], Goferman et al. propose a context-aware saliency detection method based on four principles of human visual attention. Recently, in [30], Perazzi et al. use high-dimensional Gaussian filters to formulate the saliency estimation. Inspired by recent advances in low rank matrix recovery theory, Shen and Wu [34] assume that a natural image can always be decomposed into the sparse component (salient object) and the low rank component (background). So RPCA [35] is used to complete the decomposition, and get the saliency detection map. But the assumption that background is low rank could be invalid for many natural images, the method may fail in more complex images. To remedy this issue, submodularity is introduced to saliency detection task by Jiang and Davis [36], and saliency detection is formulated as a semi-supervised clustering problem. In a word, the studies of saliency detection are going on, more problems have been proposed and solved, and the study is more closer to the physical nature of the problem. More mathematical or physical interpretations of this problem are needed, and this is also the motivation of this paper.

2.1. Boundary prior for saliency detection

The work in [37,31] has considered image boundary as a background prior. In [37], Wei et al. has proposed that the distance between a pair of background regions is shorter than that of a region from the salient object and a region from the background. Based on this criteria, the saliency detection task is formulated as an energy minimization problem. In [31], the nodes on the image boundary have been used as background seeds to rank the relevances of all the other regions, and the saliency score is measured by the relevances between the saliency seeds and all the other regions.

In this proposed method, the boundary prior is also adopted. However, the method is quite different from the previous works mentioned above. The proposed method is based on the viewpoint of anisotropic diffusion. Color contrast prior is used to guide the diffusion process. The saliency map of the first stage is the result of boundary prior (background prior) diffusion, instead of the relevances between the background seeds and all the other regions.

2.2. PDE system for saliency detection

In [38], by learning a guidance map, high level human prior knowledge is fused into the diffusion system. The saliency seeds can be learnt by optimizing a discrete submodular function, and the final saliency map is achieved through the evolution of PDE system [38]. These results have relied heavily on the guidance map's accuracy. By using Harris points detection to estimate the convex hull as the background prior, the method presented in [38] works well on images with simple backgrounds or high contrast between foreground and background. However, it fails on images with more complex backgrounds. Fig. 1 illustrates the results of using Harris points detection to estimate background prior. The red lines in Fig. 1 are the convex hull of Harris points and the regions outside of the convex hull are estimated background. As shown in Fig. 1, the estimated background result in picture (a) is clear and precise, the results in picture (b) and (c) are not precise due to the more complex backgrounds. The worst case is in picture (d), the convex hull does not even cover the object. These will lead to imprecise final results. In addition, solving optimal submodular function to learn the saliency seeds leads to higher computation cost.

3. Saliency detection via nonlocal anisotropic diffusion

This section discusses the saliency detection method from a viewpoint of the nonlocal anisotropic diffusion equation, and the detailed steps to model the visual attention spreading. Theoretical and numerical analysis is also discussed here.

3.1. Preliminary

First, the nonlocal operators and new definitions related to the saliency detection problem are defined. Hereafter, lowercase bold symbols (e.g., $\mathbf{x}$) are used to represent points and capital blackboard

Download English Version:

https://daneshyari.com/en/article/532036

Download Persian Version:

https://daneshyari.com/article/532036

Daneshyari.com