Neurocomputing 175 (2016) 336–347

Contents lists available at ScienceDirect

Neurocomputing

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

Robust manifold-preserving diffusion-based saliency detection by adaptive weight construction

Keren Fu [a,b], Irene Y.H. Gu [b], Chen Gong [a], Jie Yang [a,*]

[a] Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai, China
[b] Signal Processing Group, Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 4 June 2015
Received in revised form
13 October 2015
Accepted 19 October 2015
Communicated by Liang Lin
Available online 30 October 2015

Keywords:
Salient object detection
Saliency map
Manifold
Graph-based diffusion

ABSTRACT

Graph-based diffusion techniques have drawn much interest lately for salient object detection. The diffusion performance is heavily dependent on the edge weights in graph representing the similarity between nodes, and are usually set through manually tuning. To improve the diffusion performance, this paper proposes a robust diffusion scheme, referred to as manifold-preserving diffusion (MPD), that is built jointly on two assumptions for preserving the manifold used in saliency detection. The smoothness assumption reflects the conditional random field (CRF) property and the related penalty term enforces similar saliency on similar graph neighbors. The penalty term related to the local reconstruction assumption enforces a local linear mapping from the feature space to saliency values. Graph edge weights in the above two penalties in the proposed MPD method are determined adaptively by minimizing local reconstruction errors in feature space. This enables a better adaption of diffusion on different images. The final diffusion process is then formulated as a regularized optimization problem, taking into account of initial seeds, manifold smoothness and local reconstruction. Consequently, when applied to saliency diffusion, MPD provides a higher performance upper bound than some existing diffusion methods such as manifold ranking. By utilizing MPD, we further introduce a two-stage saliency detection scheme, referred to as manifold-preserving diffusion-based saliency (MPDS), where boundary prior, Harris convex hull, and foci convex hull are employed for deriving initial seeds and a coarse map for MPD. Experiments were conducted on five benchmark datasets and compared with eight existing methods. Our results show that the proposed method is robust in terms of consistently achieving the highest weighted F-measure and lowest mean absolute error, meanwhile maintaining comparable precision–recall curves. Salient objects in different background can be uniformly highlighted in the output final saliency maps.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Salient object, or region detection is an important research topic in computer vision [1,2]. Given an image, the main aim is to detect and uniformly emphasize objects attracting visual attention in the image, meanwhile suppress irrelevant background. Its applications to vision and graphics are numerous, especially in topics requiring object-level priors such as "proto object" detection [3], segmentation [4,5], content-based image editing [6–9], and image retrieval [10]. In the past decade, a variety of models is proposed, including heuristic color contrast-based models [11,12,9,13–16], learning-based models [17–19], segmentation-assisted approaches [20–23], and graph-based saliency modeling [24–27].

Among graph-based saliency modeling, graph-based diffusion [26–28] has recently been studied for saliency detection with good performance. To conduct saliency diffusion, an input image is first represented by a graph, followed by computing the unified formulation as follows:

$$\mathbf{s} = \mathbf{A}^{*}\mathbf{y} \tag{1}$$

where $\mathbf{A}^{*}$ is a global pair-wise propagation matrix, $\mathbf{y}$ is a seed vector that gives a preliminary assessment of saliency level of graph nodes, and $\mathbf{s}$ is the diffused result.

Aiming at improving the diffusion quality and detection performance, this paper proposes a novel and robust diffusion method, referred to as *manifold-preserving diffusion* (MPD). MPD builds jointly upon two assumptions on data manifold, namely the *smoothness* [29–31] and *local reconstruction* [32,33], for better preserving the manifold for saliency detection. The proposed MPD

* Corresponding author.
E-mail addresses: fkrsuper@sjtu.edu.cn (K. Fu), irenegu@chalmer.se (I.Y.H. Gu), goodgongchen@sjtu.edu.cn (C. Gong), jieyang@sjtu.edu.cn (J. Yang).

http://dx.doi.org/10.1016/j.neucom.2015.10.066
0925-2312/© 2015 Elsevier B.V. All rights reserved.

hence is a new way to compute $\mathbf{A}^*$ for the diffusion process in saliency detection. Based on the two assumptions, we introduce two penalties in the diffusion model. As described later, edge weights of graph in the two penalties are determined *adaptively* by solving two optimization problems. This enables better adaption of diffusion on different images. By utilizing MPD, we further introduce a two-stage saliency detection scheme, referred to as manifold-preserving diffusion-based saliency (MPDS), where boundary prior, Harris convex hull, and foci convex hull are employed for deriving initial seeds and a coarse map for MPD. Consequently, better salient object detection can be obtained in various background.

In one of the related studies on diffusion-based methods, Yang et al. [26] propose a manifold ranking-based saliency detector that employs the graph-based manifold ranking [30] to diffuse energy from four image borders. In their work, $\mathbf{A}^*$ has the form $(\mathbf{D}-\alpha\mathbf{W})^{-1}$, where $\mathbf{W}$ denotes the graph affinity matrix with entry w_{ij}, $\mathbf{D}$ is the diagonal degree matrix whose ith diagonal entry is $d_i = \sum_j w_{ij}$, and α is a constant. One can see that $\mathbf{A}^*$ in their case is a deterministic function on $\mathbf{W}$. In [26], manually tuned edge weights of graph are used for $\mathbf{W}$, where the parameter is fixed on all images. Furthermore, only the smoothness assumption is concerned. The proposed MPD differs from [26] by utilizing two assumptions and adaptive weights. Since different images have different contents and color contrast, using manually tuned edge weights is less desirable and can degrade diffusion quality. The proposed method also differs from another work [27], where their diffusion is based on geodesic distance.

The main contributions of this paper are threefold:

(i) We propose an effective graph-based diffusion method: manifold-preserving diffusion (MPD) that jointly exploits the assumptions of smoothness and local reconstruction on the manifold.
(ii) We derive two types of graph edge weights by adaptively minimizing local reconstruction errors in feature space. Hence the method is more suitable to be applied on different images. This is different from previous work where the edge weights of graph are controlled by manually tuned parameter such as bandwidth.
(iii) We introduce a two-stage saliency detection scheme: manifold-preserving diffusion-based saliency (MPDS), that leverages MPD together with boundary prior, Harris convex hull, and foci convex hull. The proposed MPDS achieves better performance than 8 recently published methods on 5 benchmark datasets.

The remainder of the paper is organized as follows. Section 2 reviews related work on salient object detection. Section 3 describes the proposed method in details, including the big picture on the proposed method, graph construction, manifold-preserving diffusion (MPD), and the two-stage manifold-preserving diffusion-based saliency detection (MPDS). Results from experiments and comparisons are given in Section 4. Finally, the conclusion is drawn in Section 5.

2. Related work

We classify existing methods into four categories: heuristic color contrast-based methods, learning-based methods, segmentation-assisted approaches, and graph-based saliency modeling. Methods beyond these four categories fall into the fifth category. For more details, readers are also referred to the comprehensive surveys [1,2].

Heuristic color contrast-based methods: Methods of this category model saliency using local or global color statistics. The underlying assumption is that salient objects are unique in color and present high color contrast to the rest parts of an image. Many methods for computing such contrast-based saliency had been proposed since 2006. Zhai et al. [11] introduce image histograms which only model luminance channel to calculate pixel-level saliency. Achanta et al. [12] provide a saliency approximation by subtracting the average color from low-pass filtered result of an image. This operation of [12] is equivalent to combining center-surround differences of all bandwidth to detect objects of different sizes. Goferman et al. [9] combine local and global features to estimate patch saliency in multi-scales. To consider both local and global factors, they compute saliency of a certain patch as its contrast to the nearest patches in feature space. Under this framework, inner parts of an object are often attenuated due to the edge preference. Cheng et al. [13] extend the method in [11] and incorporate color histograms. A regional contrast saliency measure is proposed as the contrast to other regions. Jiang et al. [14] also use regional contrast to define saliency. Instead, they use only context information from neighborhood of a region. Perazzi et al. [15] propose "saliency filter", which formulates complete contrast and saliency estimation using high dimensional Gaussian filters. Wang et al. [16] compute pixel-wise image saliency by aggregating complementary appearance contrast measures with spatial priors. A more recent method [34] computes contrast-based saliency as dissimilarity/similarity to carefully selected background/foreground seeds. Most of the above contrast-based saliency are straightforward to compute, though the performance is often less satisfactory on images with complex background.

Learning-based methods: Methods in this category estimate image saliency by machine learning. The basic idea is to learn weights of features for saliency computation. Jiang et al. [17] perform pre-segmentation for an input image and extract abundant discriminative features from each region. A random forest regressor trained is adopted to map features to a regional saliency score. Liu et al. [18] segment salient objects by aggregating pixel saliency cues in a conditional random field (CRF). The linear weights for those cues are learned under the maximized likelihood (ML) criteria by tree-reweighted belief propagation. Recently, Wang et al. [35] aim at segmenting objects-of-interest as well but solve the problem in a general joint deep learning framework, where two convolutional neural networks are employed collaboratively to boost the detection and segmentation performance. Mai et al. [19] propose a data-driven approach for aggregating saliency maps output by existing saliency detection models using a CRF. Weights for aggregation are learned in a data-driven way from most similar images retrieved from a pre-defined dataset. Lu et al. [28] learn optimal combination of seeds by maximizing figure-ground segregation. Learning-based methods can achieve good performance in complex scenarios attributed to the learning, however, high computational cost is usually needed due to feature extraction and learning.

Segmentation-assisted methods: Methods in this category aim at generating good segmentation, usually in hierarchy or multi-scale, to facilitate saliency computation. Lu et al. [20] exploit the concavity context in a scene and detect concave arcs from multi-scale segmentation. The detected arcs then contribute to a figure-ground segmentation phase. Yan et al. [21] propose a hierarchical saliency detection method that merges regions according to user-defined scales. Each region in a hierarchy is evaluated by using local contrast and location prior. Cheng et al. [22] measure saliency by hierarchical soft abstraction. They form a 4-layer hierarchical structure including pixel layer, histogram layer, GMM layer and clustering layer with an index table to associate cross-layer relations efficiently. Saliency estimation using color contrast and distribution is conducted on the coarse layers and then propagated to the pixel layer. Jiang et al. [23] find potential salient regions by maximizing a submodular objective function. The problem is solved efficiently by finding a closed-form

Download English Version:

https://daneshyari.com/en/article/407165

Download Persian Version:

https://daneshyari.com/article/407165

Daneshyari.com