Pattern Recognition 60 (2016) 130–144

Contents lists available at ScienceDirect

Pattern Recognition

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

Fine-structured object segmentation via neighborhood propagation

Yongchao Gong, Shiming Xiang *, Chunhong Pan

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 26 January 2016
Received in revised form
11 April 2016
Accepted 21 May 2016
Available online 24 May 2016

Keywords:
Fine-structured object segmentation
Label propagation
Affinity graph
Local and nonlocal neighborhood
Region cost

ABSTRACT

This paper presents a novel method for the challenging task of fine-structured (FS) object segmentation. The task is formulated as a label propagation problem on an affinity graph. The proposed method has mainly three advantages. First, to enhance the completeness and connectivity of FS objects, we introduce a novel neighborhood system combining both local and nonlocal connections, with a robust scheme for edge weight calculation. Second, appearance models are explicitly incorporated into the energy function as a term of region cost. This helps to further preserve the connectivity of the fine parts for which the label information is hard to propagate correctly via neighboring pixels alone. Third, the resulting energy minimization problem has a closed-form solution with global optimum guaranteed, showing an advantage over the FS object segmentation methods that suffer from NP-hardness. To enrich the evaluation of FS object segmentation methods, we created a new challenging data set. It consists of 100 natural images involving diverse FS objects, with accurately hand-labeled ground truth. Extensive experimental results demonstrate that our method is effective in handling FS objects and achieves the state-of-the-art performance.

1. Introduction

Object segmentation is undoubtedly one of the most fundamental tasks in image processing, pattern recognition and computer vision. The goal of object segmentation is to extract a semantic meaningful foreground object from a given image. Due to the requirement of prior knowledge about the expected objects, this goal is often achieved with the help of human–computer interaction, thus interactive techniques have become a trend. The past decades have witnessed the success of a variety of interactive image segmentation approaches [1–13] that can produce satisfactory results in various applications. However, most of them are only able to handle compact objects well, but they do not apply to fine-structured (FS) objects like trees and insects. This drawback leads to an urgent demand of effective solutions to this task in many practical applications dealing with FS objects, e.g., image synthesis [14,13] and plant modeling [15,16]. For this reason, the challenging task of FS object segmentation has naturally attracted increasing attention and become a crucial research branch.

In addition to the intrinsic difficulties in general object segmentation, FS object segmentation also suffers from its own difficulties, which mainly lie in two aspects. (1) The commonly-used boundary length regularization in many approaches [1,2] do not apply to FS objects, because their long boundaries violate this regularization. As a result, the segmented objects often have a bias towards shorter boundaries and the fine parts tend to be suppressed. This is also known as *shrinking bias* [17,5,18,7,9,11]. (2) The label information provided by the user-specified interactions is hard to propagate to the unlabeled pixels along the thin and elongated structures. For this reason, label propagation based methods [3,19,12] often produce disconnected regions around each individual interaction, thus are unsuitable to capture the correct structures of FS objects.

In recent years, there have been some methods focused on overcoming the above difficulties, and they have achieved quite promising performance [5,18,7,9,11]. However, most of these methods have relatively high model and time complexities due to the NP-hardness in their algorithms. Moreover, in some cases only approximate optimal solutions can be found [5,7,11], thus the global optimality can not be guaranteed.

In this work, we focus on addressing the problem of FS object segmentation, and present an effective method based on label propagation on a specially constructed affinity graph. The method is called *local and nonlocal neighborhood propagation* (LNNP). The core idea of LNNP is two-fold. (1) We develop a novel neighborhood system by combining both local and nonlocal connections, with a robust scheme for edge weight calculation. (2) Appearance models, often absent in label propagation based methods [3,19,12], are explicitly incorporated into the energy function as a term of region cost. The segmentation task is finally solved via global optimization with a closed-form solution. Extensive experimental

* Corresponding author.
E-mail addresses: yongchao.gong@nlpr.ia.ac.cn (Y. Gong), smxiang@nlpr.ia.ac.cn (S. Xiang), chpan@nlpr.ia.ac.cn (C. Pan).

http://dx.doi.org/10.1016/j.patcog.2016.05.025
0031-3203/

results demonstrate the effectiveness of LNNP in handling FS objects.

The main contributions of this work are highlighted as follows:

1. We propose to specially construct an affinity graph for label propagation by combining both local and nonlocal connections. The label information is thus encouraged to propagate between both nearby and distant pixels, which largely benefits the propagation along fine structures.
2. The appearance models are novelly incorporated into a label propagation based method to provide reasonable labeling preference for each pixel. This helps to preserve the fine structures for which the label information is hard to propagate correctly via pair-wise connections alone.
3. LNNP has a globally optimal closed-form solution, and linear time complexity with respect to the number of pixels. This exhibits an advantage over the FS object segmentation methods that suffer from NP-hardness [5,20,7,11].

A shorter version of this paper appeared in [21]. Apart from extensive qualitative and quantitative evaluations, the main extensions in the current work are:

1. More comprehensive descriptions and analyses about the algorithm of LNNP are provided.
2. The parameter sensitivity of LNNP is analyzed and verified.
3. Comparative experiments on ten variants of LNNP are conducted to further support the effectiveness of LNNP.
4. A new data set is created and introduced for all the experiments.

The remainder of this paper is organized as follows. In Section 2 we make a brief review on previous works. The graph model construction of LNNP is introduced in Section 3. Section 4 presents the energy function and optimization. The experimental results are reported in Section 5 and finally some conclusions and discussions are given in Section 6.

2. Previous works

We now provide a brief review on several representative methods dealing with FS object segmentation and also some methods closely related to this work.

2.1. Representative methods for FS object segmentation

Representative methods for FS object segmentation can be roughly categorized into three classes, namely methods based on cooperative cuts [7,9], methods based on topological constraints [5,20,22,11], and methods based on curvature regularization [18,23,24].

Due to the shrinking bias, the powerful technique Graph Cuts (GC) [1,25] does not apply to FS objects segmentation [17,5,7,9]. To address this issue, Jegelka and Bilmes [7] proposed cooperative cuts. The core idea is to reduce the weights of the edges cut by the actual object boundaries. In this way, the total cost of a correct segmentation is reduced, thus the shrinking bias is mitigated. Unfortunately, the resulting model suffers from NP-hardness, and only an approximate optimal solution can be found. Later, the same optimization problem was studied again by Kohli et al. [9]. They reformulated the model in [7] as a higher-order Markov random field [26], and developed a globally optimal algorithm. This algorithm proves to the effective, but is quite time-consuming.

Vicente et al. [5] analyzed the shrinking bias of GC and novelly proposed to explicitly incorporate connectivity priors as topological constraints into this framework. The purpose is to force the fine parts to be connected to the main body. Later, a variety of other connectivity prior based methods [20,22,11] are also proposed to address the shrinking bias, and this problem is solved to different extent correspondingly. However, an inherent drawback existing in most of these methods is the NP-hardness in optimization [5,20,11]. Moreover, when tackling natural images, the methods in [5] and [11] require an extra interaction for each fine part on a previously segmented object. This is a tedious and even impossible task for users when there are too many fine parts in an image or there are too many images to deal with. For this reason, these two methods [5,11] are unsuitable to be widely applied in practical applications, despite their satisfactory performance in terms of connectivity.

Curvature regularization proves to be more suitable than the commonly-used length regularization for preserving long object boundaries and fine structures, but leads to complicated optimization problems [18,23,24]. El-Zehiry and Grady [18] gave a simple formulation of a curvature regularizer and developed a fast globally optimal algorithm. One potential limitation of this method lies in that it only considers specific angular resolutions when modeling the curves. Later, Strandmark et al. [23] attempted to find fine structures by minimizing curvature using shortest paths. Recently, an efficient algorithm is developed for curvature calculation [24]. However, these two methods [23,24] have only been shown effective for either medical image or binary and gray image segmentation, while their effectiveness on natural images have not been verified yet.

2.2. Label propagation and nonlocal principle

This work is closely related to the methods based on label propagation [3,12,19] and nonlocal principle [27,28]. In the following, we briefly discuss them.

Methods based on label propagation assume that each pixel receives the label information from its neighbors according to the affinities between them. Typically, Random Walks [3] interprets the propagation as labeling an unlabeled pixel based on the probability of a random walker starting from it to reach a labeled pixel. Recently, Casaca et al. [12] developed a mathematically simple method called Laplacian Coordinates. It minimizes the label deviation of each pixel from the weighted average of its neighbors. However, in these methods, direct label propagation is only allowed between locally neighboring pixels. Connections between distant pixels inherently belonging to the same object, which are beneficial for making label propagation more effective, are not considered. As a result, these methods are prone to fail in FS object segmentation.

We find that, in the image matting community, this limitation has already been studied by Chen et al. [19]. They built connections between nearby pixels in feature space, which might be faraway on the image grid, and got satisfactory results. However, a lack of local connections makes it hard to capture the correct local structures of the objects. Later, Chen et al. [29] extended this work by integrating local cues into their nonlocal model and improved the performance. In these two methods, the nonlocal neighbors of each pixel are found by K-nearest neighbors (KNN) according to the pixel-wise features. Unfortunately, KNN might wrongly assign two pixel-wise similar pixels belonging to different objects as neighbors, thus the connection between them is likely to mislead the propagation. This issue will be verified in our subsequent experiments.

In several tasks such as image denoising and matting, there have been some studies benefitting from the nonlocal principle