ARTICLE IN PRESS

Computers and Electrical Engineering xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

An efficient two-stage region merging method for interactive image segmentation ☆

Chongbo Zhou [a,b,*], Dongmei Wu [a], Wenhua Qin [a], Chuancai Liu [b]

[a] School of Physics and Engineering, Qufu Normal University, Qufu, Shandong 273165, China

[b] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China

ARTICLE INFO

Article history:
Received 30 July 2013
Received in revised form 16 September 2015
Accepted 16 September 2015
Available online xxxx

Keywords:
Image segmentation
Interactive method
Region merging
Nearest neighbor

ABSTRACT

Interactive image segmentation aims to extract user-specified regions from the background. In this paper, an efficient two-stage region merging based method is proposed for interactive image segmentation. An image is first over-segmented into many super-pixels using a bottom-up method. The color histogram is exploited to represent each super-pixel, and the Bhattacharyya coefficient is computed to measure the similarity of two adjacent super-pixels. Then some strokes, denoting the desired object and background, are manually labeled by the user on the over-segmented image. With the labeled seed super-pixels, a merging strategy is designed to realize adaptive region merging. The whole merging process is divided into two stages, which are repeatedly executed until no new merging occurs. In the first stage, some unlabelled super-pixels are merged into the labeled foreground or background super-pixels if the labeled ones are their nearest neighbors. In the second stage, any two unlabelled super-pixels are merged together if one super-pixel is the nearest neighbor of the other. Extensive experiments are conducted to evaluate the performance of the proposed method. The results show that the proposed method can extract the object reliably and quickly from the background.

1. Introduction

Image segmentation is an important research topic in computer vision, which aims to partition an image into a finite number of non-overlapping regions. Interactive segmentation is a semi-automatic method, and it exploits the user inputs to guide the segmentation procedure. In contrast to general segmentation, where algorithms try to find consistent image regions automatically, interactive segmentation is especially suitable for object extraction from the background. Therefore, it is also called foreground/background or figure-ground segmentation.

Some of the earliest work on interactive image segmentation can be traced back to active contours [1] and intelligent scissors [2], which relies on finding good edges in images. Formulating image segmentation as a labeling problem in the optimization framework based on graph cut has become another important method over the last decade. Boykov and Jolly [3] firstly used graph cut based inference for interactive segmentation, and they proposed an energy function allowing for the user interaction. Based on the optimization framework, Rother et al. [4] proposed GrabCut. They used the same energy function as Boykov and Jolly [3], but in addition to optimizing the segmentation labels, they also optimized the color models.

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. F. Sahin.

* Corresponding author at: School of Physics and Engineering, Qufu Normal University, Qufu, Shandong 273165, China. Tel.: +86 13345196367.
E-mail address: zhouchongbo@163.com (C. Zhou).

http://dx.doi.org/10.1016/j.compeleceng.2015.09.013
0045-7906/

Please cite this article in press as: Zhou C et al. An efficient two-stage region merging method for interactive image segmentation. Comput Electr Eng (2015), http://dx.doi.org/10.1016/j.compeleceng.2015.09.013

Li et al. [5] put forward lazy snapping and they also used the same energy function, but accelerated the segmentation by pre-processing the image into super-pixels. Grady [6] proposed another energy formulation which was motivated by the idea of random walks. Liu et al. [7] presented paint selection for interactive image segmentation, which enforced the unary terms in the energy function by using localized color models based on recent brush strokes. The aforementioned methods extract the object from the background by solving an optimization problem. Besides the optimization based methods, lots of other schemes are put forward [8–11]. Ding and Yilmaz [8] presented an interactive framework for segmenting images using probabilistic hypergraphs which modeled the spatial and appearance relations among image pixels. The probabilistic hypergraph provides a means to pose image segmentation as a machine learning problem. Noma et al. [9] proposed a model-based graph matching approach for interactive image segmentation. It starts from an over-segmentation of the input image, and exploits color and spatial information among regions to propagate the labels from the regions marked by the user-provided seeds to the entire image. Protiere and Sapiro [10] proposed an interactive algorithm for soft segmentation of natural images. The segmentation result is obtained via fast, linear complexity computation of weighted distances to the user-provided scribbles. Zhang and Ji [11] proposed a new Bayesian network model for both automatic and interactive image segmentation. They constructed a multilayer Bayesian network from an over-segmentation to model the statistical dependencies among super-pixels, edge segments, vertices, and their measurements. Region growing or region merging is another class of direct and efficient method for image perception [12–15]. In these methods, the user can design ingenious merging strategies to realize region growing from seed regions. Recently, Ning et al. [15] applied seeded region merging to interactive segmentation and proposed the efficient MSRM (Maximal Similarity Region Merging) method. The conventional region merging methods merge two adjacent regions whose similarity is above a preset threshold [16], but these methods have difficulties in adaptive threshold selection. A big threshold will lead to incomplete merging of the regions belonging to the object, while a small threshold can easily cause over-merging. Moreover, it is difficult to judge when the region merging process should stop. The MSRM method overcomes the shortcomings of conventional methods by designing an adaptive merging process to merge image regions according to the defined maximal similarity rule. Zhou et al. [17] proposed a novel interactive segmentation method based on conditional random field models to use the location and color information provided by the user input. The conditional random field is configured with the optimal weights between two features, which are the color Gaussian mixture model and probability model of location information. Hichri et al. [18] proposed a new mechanism of user-machine interaction for interactive segmentation methods to solve the change detection problem in multitemporal remote-sensing images. The user needs to input markers related to change and no-change classes in the difference image. Then, the pixels under these markers are used by the support vector machine classifier to generate a spectral-change map.

In this paper, region merging is still considered and a more efficient merging strategy is proposed. An image is first over-segmented into super-pixels using a bottom-up method. The color histogram is used to represent each super-pixel, and the Bhattacharyya coefficient is computed to measure the similarity of two adjacent super-pixels. Then some strokes denoting the desired object and background are labeled on the over-segmented image manually by the user. With the labeled seed super-pixels, a merging strategy is designed to realize adaptive region merging. The merging process is repeatedly executed until all the super-pixels are labeled. The whole merging process includes two stages. In the first stage, some unlabelled super-pixels are merged into the labeled foreground or background super-pixels if the labeled ones are their nearest neighbors. In the second stage, any two unlabelled super-pixels are merged together if one super-pixel is the nearest neighbor of the other. The advantage of the proposed method is its simplicity and efficiency. Compared with the MSRM method [15], which outperforms the graph cut based algorithm [19–21] under the same user input strokes, the proposed method overall takes less time to produce more accurate results with fewer strokes.

The rest of this paper is organized as follows. In Section 2, feature representation and similarity measurement are introduced. In Section 3, the proposed merging process is detailed. In Section 4, extensive experiments are performed to evaluate the performance of the proposed method. Finally, concluding remarks are given in Section 5.

2. Feature representation and similarity measurement

2.1. Feature representation

Effective features extracted from the super-pixels are crucial to the following classification task. The proposed method focuses on merging super-pixels instead of pixels mainly because the feature of a super-pixel is more robust than that of a pixel in pattern classification. Color histogram is a class of distinctive feature which is exploited in pattern recognition, object tracking, and image segmentation [15,22,23]. Ning et al. [15] used single-channel indexed color histograms of super-pixels as feature vectors. They applied indexed histograms in RGB color space to interactive image segmentation and obtained fine performance. Ning et al. [15] also demonstrated the effectiveness of indexed histograms in other color spaces. Though super-pixels obtained using the mean shift method [24] are different in size, and irregular in shape, indexed color histograms are discriminative enough to characterize the super-pixels.

We adopt the strategy of using super pixels instead of pixels, because the feature of a super-pixel is more robust than that of a pixel, and furthermore operations on super-pixels demand less computational cost. Any existing low-level over-segmentation methods, such as the graph cut based method [21], the mean shift method [24] and the watershed method [25], can be used in this pre-processing step. In this paper, we adopt the mean shift method proposed in [24] for initial segmentation because it can well preserve the object boundaries which contribute to excellent results.

Please cite this article in press as: Zhou C et al. An efficient two-stage region merging method for interactive image segmentation. Comput Electr Eng (2015), http://dx.doi.org/10.1016/j.compeleceng.2015.09.013