Available online at www.sciencedirect.com

ScienceDirect

April 2017, 24(2): 38–47
www.sciencedirect.com/science/journal/10058885

The Journal of China Universities of Posts and Telecommunications

http://jcupt.bupt.edu.cn

GrabCut image segmentation algorithm based on structure tensor

Zhang Yong[1], Yuan Jiazheng[1,2] (✉), Liu Hongzhe[1], Li Qing[1]

1. Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China
2. Beijing High-Tech Innovation Centre of Imaging Technology, Capital Normal University, Beijing 100048, China

Abstract

This paper attempts to present an interactive color natural images segmentation method. This method extracts the feature of images by using the nonlinear compact structure tensor (NCST) and then uses GrabCut method to obtain the segmentation. This method not only realizes the non-parametric fusion of texture information and color information, but also improves the efficiency of the calculation. Then, the improved GrabCut algorithm is used to evaluate the foreground target segmentation. In order to calculate the simplicity and efficiency, this paper also extends the Gaussian mixture model (GMM) constructed base on the GrabCut to the tensor space, and uses the Kullback-Leibler (KL) divergence instead of the usual Riemannian geometry. Lastly, an iteration convergence criterion is proposed to reduce the time of the iteration of GrabCut algorithm dramatically with satisfied segmentation accuracy. After conducting a large number of experiments on synthetic texture images and natural images, the results demonstrate that this method has a more accurate segmentation effect.

Keywords image segmentation, structure tensor, GrabCut, Kullback-Leibler, GMM

1 Introduction

Extracting a foreground object in a complex environment is of great practical importance in computer vision. To extract objects from color images is even more challenging. Color images carry much more information than gray ones [1], and these information can be used to enhance the image analysis process and improve segmentation results. Therefore, color images segmentation has been studied for decades, and recently received much attention for special effects in film, television, publication, photography and a number of desktop applications.

Because the image contains a wealth of information as well as their unpredictability and complexity, manual segmentation will be tedious and time consuming for a large number of image segmentation tasks, also is unrealistic. A general purpose of image segmentation technique should be able to accurately define the desired object boundaries or regions automatically or semi-automatically with minimal user input. Existing image segmentation algorithms can be generally classified into three major categories: feature-based, region-based and boundary-based [2]. Under normal circumstances, each class has its own scope of application to the method. Besides, feature-based method attempts to classify pixel points based upon their position on the feature space, regardless of the similarities between the connectivity. Common features include color, gradients, texture and depth.

The classification based on the gray threshold and the distance information belongs to the feature-based method. This approach has a serious drawback that there are a number of pixel points within the image which does not belong to the same region. However, when they have a high degree of similarity in the feature space, this method will put them into a class. Besides, the region-based method is based on the expansion of the feature method, which not only maintains the similarity in the feature space but also keeps the connectivity of the region. Segmentation algorithms on the basis of the region, including spot color,

Received date: 13-09-2016
Corresponding author: Yuan Jiazheng, E-mail: jiazheng@buu.edu.cn
DOI: 10.1016/S1005-8885(17)60197-3

region growing, region merged, regional segmentation and intelligent coatings. Nevertheless, these methods may produce a large number of discrete areas fragmented. This drawback cannot be found in boundary-based method. The methods are based on the boundary edge through optimizing existing standard curve to approximate the real border, surrounded by the outline of the object is extracted or sub-regions. Existing boundary-based algorithms include: boundary tracing, dynamic programming, active contour algorithms and smart scissors. However, due to localized edge information, if we want to achieve the desired global object, it requires a lot of human interactions.

2 Related work

In recent years, image segmentation methods based on graph theory have revealed a high accuracy. The biggest feature of these algorithms are to combine the segmented energy function and edge regularization in the region. Their common strategy is to construct a weighted graph. Each vertex of the graph corresponds to an image pixel point or region. The weight of the connections between each two vertices indicates the magnitude of the probability that they belong to the same area [3]. Besides, the weight is often associated with the extracted features. Subsequently, the graph is divided into a plurality of components to minimize the energy function.

Graph cuts algorithm is an interactive segmentation which is on the basis of graph theory. At first, Graph cuts was proposed by Boykov et al. [4]. This algorithm requires the user to primarily identify the object needs to be split. Then, through calculating the results obtained in line with the user calibration, the optimal value is selected to achieve the image segmentation. Such an approach enables the user to obtain the satisfactory segmentation results through using a very intuitive interaction. This approach in the field of computer vision and graphics has been widely studied for image restoration, segmentation of two-dimensional (2D) and three-dimensional (3D) images and so on. GrabCut [5] extends Graph cuts to color images and incomplete trimaps. It replaces the monochrome image model based on histograms in Ref. [6] by GMM and iteratively alternates between estimation and parameter learning to solve the min-cut problem until converges. Therefore, the user interaction can be relaxed to simply placing a rectangle or a lasso around the object, followed by a small amount of corrective editing. These developments make GrabCut more convenient to image editing such as foreground extraction.

In Refs. [5–6], the image is divided into target and background process, which is mainly based upon the value of regional statistical image. However, the image value of statistical calculations may not be sufficient to distinguish the various regions. In some cases, the texture information is often more suitable as a discriminating feature. There are many texture feature description methods in image segmentation, including the Markov random fields [7–8], multi-scale solution [9–10], Gabor wavelet filter and so on. In 1991, Bigun et al. [11] proposed structure tensor (ST) to describe the texture. This method only contains three characteristic channels, and each channel can provide more significant features than Gabor wavelet. Based on the structure of the image segmentation algorithm, a variety of image segmentation has also been proposed, including the active contour model, level set model [12–13] and so on. For example, in Ref. [14], the use of each channel Gaussian approximation of nonlinear structure tensor (NST) and strengthen the histogram as a feature, which was brought into active contour model for image segmentation. In addition, in the Ref. [15], Graph cuts is used for diffusion tensor magnetic resonance images, and KL divergence is used as the dissimilarity measure, obtained by calculating the average distance of each sub-tensor and as between the terminal the weight of.

Malcolm et al. [16] extended this tensor-based approach through introducing tensor in Riemannian geometry, for multi-object and background image segmentation. Moreover, the interactive Graph cuts is also applied for multi-modal tensor image segmentation. In order to improve the efficiency of segmentation, a 3×3 contained intensity extended structure tensor (EST) information is constructed, as input data of Graph cuts. However, in the EST, color information and texture information of each ratio is 50%, due to the mixing factor cannot be adaptively adjusted. Therefore, there will be a lot of useless information. Additionally, if the introduction of high dimensional 5×5 (considering all color channels) EST, it means that the energy must be implemented to minimize higher order in space, which would be extremely difficult and very possible to be a plurality of local small results.

To deal with high computational complexity of the classical ST of texture images effectively, a NCST is proposed to describe the texture feature of images and