Pattern Recognition 46 (2013) 1020–1038

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

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

A survey of graph theoretical approaches to image segmentation

Bo Peng [a,b], Lei Zhang [b,*], David Zhang [b]

[a] Department of Software Engineering, Southwest Jiaotong University, Chengdu, China
[b] Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China

ARTICLE INFO

Article history:
Received 18 December 2011
Received in revised form
15 September 2012
Accepted 18 September 2012
Available online 5 October 2012

Keywords:
Image segmentation
Graph theoretical methods
Minimal spanning tree
Graph cut

ABSTRACT

Image segmentation is a fundamental problem in computer vision. Despite many years of research, general purpose image segmentation is still a very challenging task because segmentation is inherently ill-posed. Among different segmentation schemes, graph theoretical ones have several good features in practical applications. It explicitly organizes the image elements into mathematically sound structures, and makes the formulation of the problem more flexible and the computation more efficient. In this paper, we conduct a systematic survey of graph theoretical methods for image segmentation, where the problem is modeled in terms of partitioning a graph into several sub-graphs such that each of them represents a meaningful object of interest in the image. These methods are categorized into five classes under a uniform notation: the minimal spanning tree based methods, graph cut based methods with cost functions, graph cut based methods on Markov random field models, the shortest path based methods and the other methods that do not belong to any of these classes. We present motivations and detailed technical descriptions for each category of methods. The quantitative evaluation is carried by using five indices – Probabilistic Rand (PR) index, Normalized Probabilistic Rand (NPR) index, Variation of Information (VI), Global Consistency Error (GCE) and Boundary Displacement Error (BDE) – on some representative automatic and interactive segmentation methods.

1. Introduction

Image segmentation is a classical and fundamental problem in computer vision. It refers to partitioning an image into several disjoint subsets such that each subset corresponds to a meaningful part of the image. As an integral step of many computer vision problems, the quality of segmentation output largely influences the performance of the whole vision system. A rich amount of literature on image segmentation has been published over the past decades. Some of them have achieved an extraordinary success and become popular in a wide range of applications, such as medical image processing [1–3], object tracking [4,5], recognition [6,7], image reconstruction [8,9] and so on.

Since the very beginning, image segmentation has been closely related to perceptual grouping or data clustering. Such a relationship was clearly pointed out by Wertheimer's gestalt theory [10] in 1938. In this theory, a set of grouping laws such as similarity, proximity and good continuation are identified to explain the particular way by which the human perceptual system groups tokens together. The gestalt theory has inspired many approaches to segmentation, and it is hoped that a good segmentation can capture perceptually important clusters which reflect local and/or global properties of the image. Early edge detection methods such as the Robert edge detector, the Sobel edge detector [11] and the Canny edge detector [12,13] are based on the abrupt changes in image intensity or color. Due to the distinguishable features of the objects and the background, a large number of thresholding based methods [14–16] have been proposed to separate the objects from the background. In the partial differential equations (PDE) based methods [17–21], the segmentation of a given image is calculated by evolving parametric curves in the continuous space such that an energy functional is minimized for a desirable segmentation. Region splitting and merging is another popular category of segmentation methods, where the segmentation is performed in an iterative manner until some uniformity criteria [22,23] are satisfied. The reviews of various segmentation techniques can be found for image thresholding methods [24], medical image segmentation [25,26], statistical level set segmentation [27], 3D image segmentation [28], edge detection techniques [29] and so on.

Among the previous image segmentation techniques, many successful ones benefit from mapping the image elements onto a graph. The segmentation problem is then solved in a spatially discrete space by the efficient tools from graph theory. One of the advantages of formulating the segmentation on a graph is that it might require no discretization by virtue of purely combinatorial operators and thus incur no discretization errors. Despite the

* Corresponding author.
E-mail address: cslzhang@comp.polyu.edu.hk (L. Zhang).

0031-3203/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2012.09.015

large amount of efforts devoted to image segmentation, little work has been done to review the work in this field. In this paper, we conduct a systematic survey of some influential graph theoretic techniques for image segmentation, where the problem is generally modeled in term of partitioning a graph into several sub-graphs.

With a history dating back to 1960s, the earliest graph theoretic methods stress the importance of the gestalt principles of similarity or proximity in capturing perceptual clusters. The graph is then partitioned according to these criteria such that each partition is considered as an object segment in the image. In these methods, fixed thresholds and local measures are usually used for computing the segmentation results, while global properties of segmentation are hard to guarantee. The introduction of graph as a general approach to segmentation with a global cost function was brought by Wu et al. [30] in 1990s. From then on, much research attention was moved to the study of optimization techniques on the graph. It is known that one of the difficulties in image segmentation is its ill-posed nature. Since there are multiple possible interpretations of the image content, it might be difficult to find a single correct answer for segmenting a given image. This suggests that image segmentation should incorporate the mid- and high-level knowledge in order to accurately extract objects of interest. In the late 1990s, a prominent graph technique emerged in the use of a combination of model-specific cues and contextual information. An influential representation is the s/t graph cut algorithm [31]. Its technical framework is closely related to some variational methods [17–21] in terms of a discrete manner. Up to now, s/t graph cut and its variants have been extended for solving many computer vision problems, and eventually acting as an optimization tool in these areas.

This paper provides a systematic survey of graph theoretic techniques and distinguishes them by broadly grouping them into five categories. (1) *Minimal spanning tree based methods*: the clustering or grouping of image pixels are performed on the minimal spanning tree. The connection of graph vertices satisfies the minimal sum on the defined edge weights, and the partition of a graph is achieved by removing edges to form different sub-graphs. (2) *Graph cut with cost functions*: graph cut is a natural description of image segmentation. Using different cut criteria, the global functions for partitioning the graph will be different. Usually, by optimizing these functions, we can get the desirable segmentation. (3) *Graph cut on Markov random field models*: the goal is to combine the high level interactive information with the regularization of the smoothness in the graph cut function. Under the MAP–MRF framework, the optimization of the function is obtained by the classical min-cut/max-flow algorithms or its nearly optimal variants. (4) *The shortest path based methods*: the object boundary is defined on a set of shortest paths between pairs of graph vertices. These methods require user interactions to guide the segmentation. Therefore, the process is more flexible and can provide friendly feedback. (5) *Other methods*: we will refer to several efficient graph theoretic methods that do not belong to any of the above categories, such as random walker [32] and dominant set based method [33].

For each of the above categories, the principle of graph theory will be first introduced, and then the theoretic formulation as well as their segmentation criteria will be reviewed. Performance assessment of some well-known methods will also be given for the sake of completeness and illustration. The outline of the paper is as follows. In Section 2, important notations and definitions in graph theory are introduced. In Section 3, the methodologies of the five categories of methods are reviewed. Explicit explanations are presented on the formulation of the problem and the details of different segmentation criteria. In Section 4, some quantitative metrics of the segmentation quality are described. The performances of some representative automatic and interactive segmentation techniques are analyzed in Sections 5 and 6, respectively. In Section 7, the applications of graph based methods in medical image segmentation are discussed. Section 8 draws the conclusion.

2. Background

In this section we define some terminologies that will be used throughout the paper for explaining the graph based segmentation methods.

Let $G=(V, E)$ be a graph where $V=\{ v_1,\ldots, v_n \}$ is a set of vertices corresponding to the image elements, which might represent pixels or regions in the Euclidean space. E is a set of edges connecting certain pairs of neighboring vertices. Each edge $(v_i,v_j)\in E$ has a corresponding weight $w(v_i,v_j)$ which measures a certain quantity based on the property between the two vertices connected by that edge. For image segmentation, an image is partitioned into mutually exclusive components, such that each component A is a connected graph $G'=(V', E')$, where $V'\subseteq V$, $E'\subseteq E$ and E' contains only edges built from the nodes of V'. In other words, nonempty sets $A_1, \ldots, A_k$ form a partition of the graph G if $A_i\cap A_j=\phi$ $(i,j\in\{1, 2, \ldots, k\}, i\neq j)$ and $A_1\cup\ldots\cup A_k=G$. The well-accepted segmentation criteria [10] require that image elements in each component should have uniform and homogeneous properties in the form of brightness, color, or texture, etc., and elements in different components should be dissimilar.

In graph theoretic definition, the degree of dissimilarity between two components can be computed in the form of a graph cut. A cut is related to a set of edges by which the graph G will be partitioned into two disjoint sets A and B. As a consequence, the segmentation of an image can be interpreted in form of graph cut, and the cut value is usually defined as:

$$\mathrm{cut}(A,B)=\sum_{u\in A, v\in B} w(u,v) \tag{1}$$

where u and v refer to the vertices in the two different components. In image segmentation, noise and other ambiguities bring uncertainties into the understanding of image content. The exact solution to image segmentation is hard to obtain. Therefore, it is more appropriate to solve this problem with optimization methods. The optimization-based approach formulates the problem as a minimization of some established criterion, whereas one can find an exact or approximate solution to the original uncertain visual problem. In this case, the optimal bi-partitioning of a graph can be taken as the one which minimizes the cut value in Eq. (1).

In a large amount of literature, image segmentation is also formulated as a labeling problem, where a set of labels L is assigned to a set of sites in S. In two-class segmentation, for example, the problem can be described as assigning a label f_i from the set $L=\{object, background\}$ to site $i\in S$ where the elements in S are the image pixels or regions. Labeling can be performed separately from image partitioning, while they achieve the same effect on image segmentation. We will see in this survey that many methods perform both partitioning and labeling simultaneously. An example to illustrate the relationship between graph cut and the corresponding vertex labeling is given in Fig. 1, where a graph is segmented by two cuts and thus has 3 labels in the final segmentation.

Methods in image segmentation can be categorized into automatic methods and interactive methods. Automatic segmentation is desirable in many cases for its convenience and generality. However, in many applications such as medical or biomedical imaging, objects of interest are often ill-defined so that even sophisticated automatic segmentation algorithms will fail. Interactive methods can improve the accuracy by incorporating prior knowledge from the user; however, in some practical applications where a large number of images are needed to be handled, they can