Neurocomputing 207 (2016) 416–427

Contents lists available at ScienceDirect

Neurocomputing

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

Ensembling over-segmentations: From weak evidence to strong segmentation

Dong Huang [a,b], Jian-Huang Lai [b,c,*], Chang-Dong Wang [b,c], Pong C. Yuen [d]

[a] College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China
[b] School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China
[c] Guangdong Key Laboratory of Information Security Technology, Guangzhou, China
[d] Department of Computer Science, Hong Kong Baptist University, Hong Kong, China

ARTICLE INFO

Article history:
Received 27 September 2015
Received in revised form
26 February 2016
Accepted 5 May 2016
Communicated by Yi Zhe Song
Available online 16 May 2016

Keywords:
Image segmentation
Ensemble clustering
Over-segmentation
Ensemble segmentation
Segmentation fusion

ABSTRACT

Due to the high diversity of image data, image segmentation is still a very challenging problem after decades of development. Each segmentation algorithm has its merits as well as its drawbacks. Instead of segmenting images via conventional techniques, inspired by the idea of the ensemble clustering technique that combines a set of weak clusterers to obtain a strong clusterer, we propose to achieve a consensus segmentation by fusing evidence accumulated from multiple weak segmentations (or over-segmentations). We present a novel image segmentation approach which exploits multiple over-segmentations and achieves segmentation results by hierarchical region merging. The cross-region evidence accumulation (CREA) mechanism is designed for collecting information among over-segmentations. The pixel-pairs across regions are treated as a bag of independent voters and the cumulative votes from multiple over-segmentations are fused to estimate the coherency of adjacent regions. We further integrate the brightness, color, and texture cues for measuring the appearance similarity between regions in an over-segmentation, which, together with the CREA information, are utilized for making the region merging decisions. Experiments are conducted on multiple public datasets, which demonstrate the superiority of our approach in terms of both effectiveness and efficiency when compared to the state-of-the-art.

1. Introduction

Image segmentation is a fundamental yet challenging problem in the field of computer vision. The purpose of image segmentation is to partition an image into a certain number of regions that have coherent properties. There is a large amount of literature on image segmentation in the past few decades [1–19]. However, most of the existing approaches focus on combining low-level visual features and global optimization methods. In this paper, we explore an alternative strategy which aims to accumulate cues from multiple over-segmentations (generated by different segmentation methods or by the same method with different parameters) to obtain a more robust and better segmentation.

Over-segmentation occurs when coherent regions are split into smaller segments, which is generally an easier task than obtaining good segmentations. In one aspect, over-segmenting is usually not preferred for image segmentation. In another aspect, however, lots of object details and pixel-wise relationship are well preserved among the segments in an over-segmentation. An over-segmentation can be viewed as a *weak* segmentation. Similar to the ensemble techniques [20–26] which aim to combine a set of weak clusterers to obtain a strong clusterer, this paper addresses the problem of accumulating evidence from multiple weak segmentations (or over-segmentations) to obtain a strong segmentation (see Fig. 1). We propose a novel image segmentation approach based on over-segmetation fusion and hierarchical region merging. A set of over-segmentations generated by different methods are used, each treated as an independent evidence. The information in the set of over-segmentations, together with the brightness, color and texture cues, are incorporated in the region merging process to construct the final segmentation. The cross-region evidence accumulation (CREA) mechanism is presented for collecting information from multiple over-segmentations via a regional voting strategy. The experiments on three public datasets,

* Corresponding author. Present address: School of Data and Computer Science, Sun Yat-sen University, Guangzhou Higher Education Mega Center, Panyu District, Guangzhou, Guangdong 510006, China. Tel.: +86 13168313819; fax: +86 20 84110175.
E-mail addresses: huangdonghere@gmail.com (D. Huang), stsljh@mail.sysu.edu.cn (J.-H. Lai), changdongwang@hotmail.com (C.-D. Wang), pcyuen@comp.hkbu.edu.hk (P.C. Yuen).

http://dx.doi.org/10.1016/j.neucom.2016.05.028
0925-2312/

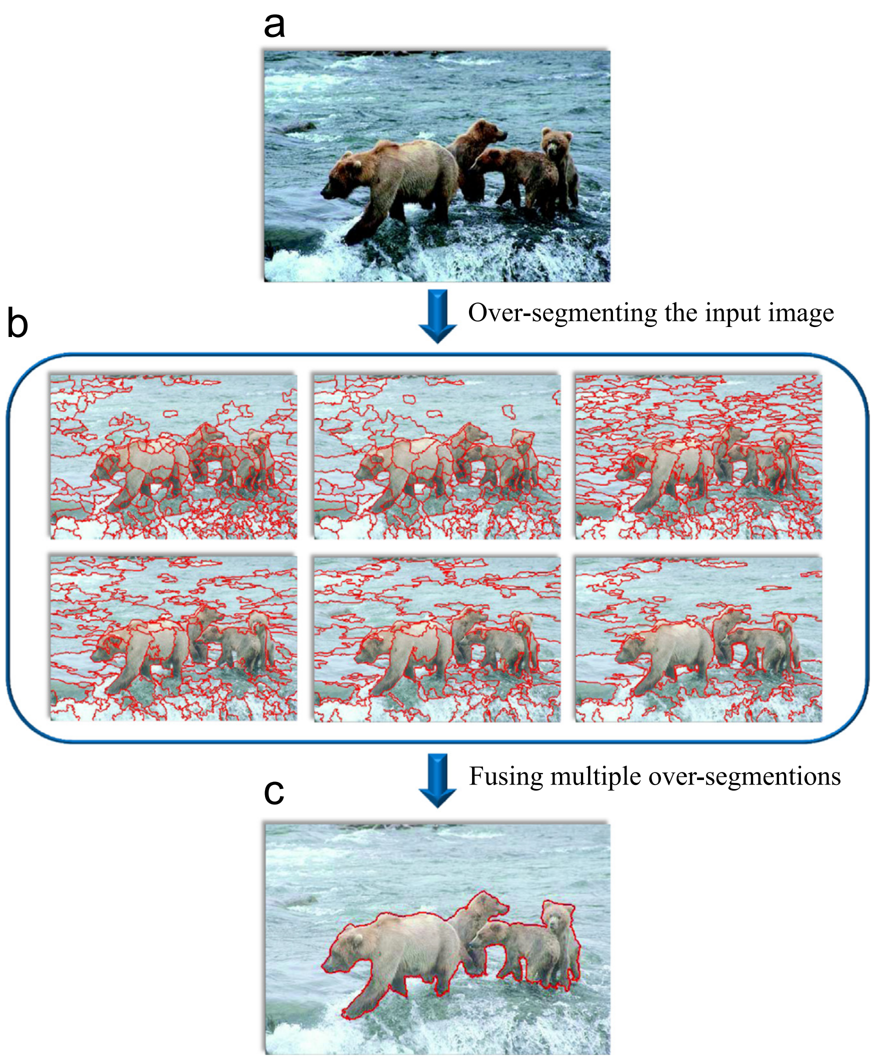

Fig. 1. Image segmentation using multiple over-segmentations. (a) The original image. (b) Multiple over-segmentations produced by the mean-shift method [3] and the F–H method [4]. (c) After building a hierarchy of segmentations by fusing evidence from multiple over-segmentations, the final segmentation is obtained by choosing a segmentation number (or level) for the hierarchy. Here, the PRI-optimized segmentation number is adopted.

i.e., BSDS300, BSDS500, and MSRC, show that the proposed approach produces significantly better segmentation results than the state-of-the-art techniques with a low computational cost in execution time and memory usage.

The remainder of this paper is organized as follows. We review the related work of image segmentation and ensemble clustering in Section 2. The proposed segmentation approach with cross-region evidence accumulation and cue integration is introduced in Section 3. The experimental results are reported in Section 4. We conclude this paper in Section 5.

2. Related work

2.1. Image segmentation

During the past few decades, many image segmentation approaches have been developed by exploiting a wide variety of techniques, such as mode seeking [3,11], graph partitioning [2,6,7,13,27], region merging [4,5,9], fuzzy clustering [28,29], variational methods [12,18,30], Markov random fields [8,17], and level set methods [14,19,31]. Among these categories, the mode seeking, graph partitioning, and region merging methods are three of the most popular techniques for segmenting natural images.

The mode seeking methods [3,11] provide image segmentation with a versatile tool via feature space analysis. The mean-shift method [3] is one of the most widely used mode seeking methods. Local density maxima (or modes) in the feature space are detected and the feature space is partitioned by several clusters with the modes of the density being cluster centroids. The mean-shift method is capable of generating clusters of arbitrary shapes. Though the image details are well respected, the mean-shift method tends to produce artifacts by splitting a coherent region into pieces. In practical applications the mean-shift method is often exploited as a pre-processing step to generate a set of primitive segments [32,33], which are also called superpixels.

The graph partitioning based methods [2,6,7,13] offer a way to incorporate global information into the image segmentation process. Typically, a graph is constructed with the image pixels mapped onto the graph vertices and the relationship between pixels onto the weighted graph links. Spectral clustering [34,35] is often utilized to partition the graph into a certain number of non-overlapping subgraphs and achieve the segmentation of the image. Shi and Malik [2] for the first time, to our knowledge, introduced spectral clustering into the field of computer vision and developed the normalized cuts (Ncuts) for graph partitioning and image segmentation. In the naive implementation of the Ncuts [2], only the links connecting pixels within a small spatial distance