J. Vis. Commun. Image R. 26 (2015) 66–79

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

A human visual system-driven image segmentation algorithm ☆

Renbin Peng [a,*], Pramod K. Varshney [b]

[a] Facebook Inc., 1601 Willow Road, Menlo Park, CA 94025, USA
[b] Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244, USA

ARTICLE INFO

Article history:
Received 14 January 2014
Accepted 4 November 2014
Available online 10 November 2014

Keywords:
Image processing
Image segmentation
Human visual system
HVS-driven segmentation
Just-noticeable difference (JND) model
Markov random fields
Quality metrics
Boundary- and region-based segmentation

ABSTRACT

This paper presents a novel image segmentation algorithm driven by human visual system (HVS) properties. Segmentation quality metrics, based on perceptual properties of HVS with respect to segmentation, are integrated into an energy function. The energy function encodes the HVS properties from both region-based and boundary-based perspectives, where the just-noticeable difference (JND) model is employed when calculating the difference between the image contents. Extensive experiments are carried out to compare the performances of three variations of the presented algorithm and several representative segmentation and clustering algorithms available in the literature. The results show superior performance of our approach.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Image segmentation plays a critical role in image analysis. It subdivides an image into its constituent parts in order to extract information regarding objects of interest, and has an impact on all the subsequent image analysis tasks, such as object classification and scene interpretation [1]. Image segmentation is a challenging problem in computer vision, and a wide variety of solutions have been presented. Surveys of image segmentation techniques can be found in [1–3]. Based on the image information being employed for the segmentation task, image segmentation algorithms can be classified into four categories: region-based segmentation, boundary- or edge-based segmentation, methods combining both region and boundary (edge) information, and thresholding (multi-thresholding) methods.

Region-based segmentation methods aim at exploiting the image contextual information, such as spatial dependency or spatial distribution. The segmented images are expected to consist of regions within which the image content is homogeneous, while the contrast between neighboring regions is high. Typical methods falling into this category include region growing/merging [4,8], watershed, some Markov random fields (MRF)-based methods [5–7,14], mean-shift [9] and the lossy data compression-based approach [10]. A recent state-of-the-art algorithm belonging to this category is based on saliency detection [45], where the bounding boxes that cover most of the salient points are considered as the detected object region. In [45], global contrast and three other measures (compactness, continuity, and center) are employed to construct the saliency map.

Segmentation methods based on the boundary or edge information are designed to exploit the discontinuity of the image features, such as difference in texture or pixel intensity, on the two sides of the boundary. Typical methods in this category include those taking into account the interaction between boundaries (or edges) [11–13,20,23] and the methods derived from physics models [15–17].

There also exist algorithms that combine region-based and boundary-based segmentations in order to benefit from the fusion of these two complementary approaches. There are two types of algorithms that belong to this category. The first type of algorithms carry out region and boundary segmentations sequentially [18,19,24,27], where one segmentation method is employed as the preprocessing or initialization step of the other method. A recent state-of-the-art algorithm falling into this category is a fixation point-based algorithm [43], where the proposed segmentation process is carried out in two separate steps: step 1, a probabilistic boundary edge map is calculated by combining visual cues; step 2, the "optimal" closed contour around a given fixation point is found. The second type of algorithms perform segmentation by considering region and boundary information simultaneously [21,22,29]. A state-of-the-art algorithm belonging to this type is a saliency detection-based algorithm [39]. In this algorithm, the salient regions, which are obtained from the visual

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author. Current address: Facebook Inc., 1601 Willow Road, Menlo Park, CA 94025, USA.

E-mail address: rpeng@syr.edu (R. Peng).

http://dx.doi.org/10.1016/j.jvcir.2014.11.002
1047-3203/© 2014 Elsevier Inc. All rights reserved.

attention-driven saliency detection, are employed as the object seeds for an automatic object segmentation system. The energy function of the segmentation system is composed of both regional and boundary terms.

Thresholding is another well-known method for image segmentation in which the threshold (or multiple thresholds) is determined from the features of an image, e.g., pixel intensity, based on some criterion, which is employed to partition an image into different regions [25,26,28].

Design of a suitable energy function to be used for segmentation is crucial to the performance of image segmentation approaches. Successful segmentation algorithms require an appropriate description of the desired properties of the segmentation result, which, of course, is a very challenging task. In real-world applications, lots of segmentation algorithms only partially incorporate the feature information from region or boundary perspectives. For example, the Markov Chain Monte Carlo (MCMC) approach has been employed [30] to solve the maximum *a posteriori* (MAP)-MRF estimation problem for generative image segmentation, but the proposed energy function does not fully exploit the connectivity property of the neighboring pixels. The normalized cut methods [31–33] can capture salient parts of an image. However, due to the ad hoc approximations introduced when relaxing this NP-hard computational problem, these methods do not exploit the image content information well which, however, is important for segmentation. As a result, the algorithms often perform unsatisfactorily.

Another weakness of many existing segmentation algorithms is that they neglect the fact that the human is the best and usually the ultimate evaluator of the segmentation result. That is, these algorithms do not consider the impact of the human visual system (HVS) on object interpretation and information extraction. As a result, the outputs of many segmentation algorithms are inconsistent with the preferences of human vision. There do exist efforts to incorporate HVS information into image segmentation, e.g., [34,35,38,40–42,44], but these works either lack tractable computational models or do not incorporate HVS properties to a sufficient extent when designing segmentation energy functions. The performance of several of the above algorithms will be discussed later in this paper.

Our work aims at designing an image segmentation algorithm based on HVS properties. More specifically, based on the quality metrics for region-based and boundary-based segmentation evaluations, we integrate region label estimation for each pixel with boundary localization for each region. These metrics attempt to mimic the preferences of human vision to good segmentation and thus make the segmentation HVS-driven. Segmentation is carried out by optimizing the energy function which reflects the desired properties of segmentation from both global and local perspectives.

This paper is structured as follows. Section 2 introduces the idea of HVS-driven image segmentation, where the properties of good segmentation results are described from region and boundary perspectives respectively, based on which a segmentation energy function is introduced. Also, the JND-weighting scheme is described which is used throughout the development of the energy functions in this paper. The energy function corresponding to the region-based segmentation is discussed in Section 3. In Section 4, the energy function corresponding to boundary-based segmentation is developed via the introduction of a novel concept, called boundary element in this paper, which describes the interaction between pixel labels, boundary configuration and image content. The integrated energy function that includes both region and boundary information is described in Section 5, where the optimization methods and the three variations of the HVS-driven segmentation algorithm are discussed. Experimental results and performance comparisons between the presented algorithms and other representative segmentation and clustering algorithms are presented in Section 6. Concluding remarks are provided in Section 7.

2. HVS-driven image segmentation and JND-weighted contrast measure

2.1. HVS-driven image segmentation

HVS-driven segmentation, considered in this paper, is motivated by the fact that humans are the ultimate judges of the quality of a segmentation result in most circumstances. So a segmentation algorithm is likely to yield more satisfactory results if the energy function is designed by including HVS preferences within the context of segmentation. In this section, several HVS-based segmentation quality evaluation metrics are first introduced, based on which the segmentation energy function is developed.

A human often evaluates the segmentation result in both global and local manners, that is, the fitness of a segmentation result to the entire image content and the local image region is considered simultaneously. At the same time, we note that region and boundary are two mutually complementary factors when segmenting an image. Therefore, we summarize the desirable properties of good segmentation [1,46–48] as follows.

Region-based properties:

(R1) The contrast of pixel intensities between neighboring regions, i.e., inter-region contrast, should be large;
(R2) The contrast of pixel intensities within a region, i.e., intra-region contrast, should be small;
(R3) Neighboring pixels with the same label are preferable, i.e., homogeneity of neighboring pixels.

Properties (R1) and (R2) represent the global properties of a good segmentation, and property (R3) is a local property which indicates that segmentation should yield large-sized regions.

Boundary-based properties:

(B1) Region boundary should be smooth and of as small a length as possible. In other words, the boundary should avoid containing too many sharp angles or turns;
(B2) The intensity contrast of a neighboring pixel pair on the two sides of the boundary should be large, while the contrast of the neighboring pixels on the same side of the boundary should be small;
(B3) The pixels lying on the boundary curve should be connected.

Property (B1) represents a property which is characterized by the image content in both global and local manners. Properties (B2) and (B3) are properties of a good segmentation in small regions, and can be measured locally.

By considering both region-based and boundary-based properties simultaneously, we formulate the image segmentation problem as an optimization problem as follows.

$$(\hat{L}, \hat{B}) = \underset{L\in\Omega_L, B\in\Omega_B}{\arg\min}\, U(L(Y), B(Y)) \tag{1}$$

where the energy function $U(L(Y), B(Y))$ consists of two factors, corresponding to region-based and boundary-based segmentation properties described above. L and B in (1) denote pixel labels and boundary elements, respectively. Both of them depend on the observed image data Y, so they are functions of Y. The boundary elements will be defined in Section 4. Since a segmented image region is composed of the pixels with the same label, label-based

Download English Version:

https://daneshyari.com/en/article/6938542

Download Persian Version:

https://daneshyari.com/article/6938542

Daneshyari.com