Optics & Laser Technology 57 (2014) 236–242

Contents lists available at ScienceDirect

Optics & Laser Technology

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

Weighted entropy for segmentation evaluation

Jesmin F. Khan*, Sharif M. Bhuiyan

Department of EE, Tuskegee University, Tuskegee, AL 36088, United States

ARTICLE INFO

Available online 21 August 2013

Keywords:
Information theory
Entropy
Image segmentation

ABSTRACT

In many image, video and computer vision systems the image segmentation is an essential part. Significant research has been done in image segmentation and a number of quantitative evaluation methods have already been proposed in the literature. However, often the segmentation evaluation is subjective that means it has been done visually or qualitatively. A segmentation evaluation method based on entropy is proposed in this work which is objective and simple to implement. A weighted self and mutual entropy are proposed to measure the dissimilarity of the pixels among the segmented regions and the similarity within a region. This evaluation technique gives a score that can be used to compare different segmentation algorithms for the same image, or to compare the segmentation results of a given algorithm with different images, or to find the best suited values of the parameters of a segmentation algorithm for a given image. The simulation results show that the proposed method can identify over-segmentation, under-segmentation, and the good segmentation.

1. Introduction

Image segmentation plays an important role in many image processing and computer vision algorithm. It is basically partitioning an image into separate regions corresponding to different real-world objects. A number of image segmentation algorithms have already been proposed in the literature for more accurate and effective results. But being a complex problem segmentation does not have an exact solution. Analyzing segmentation methods usually encounters two problems [1]: (1) inability to compare different segmentations effectively, and (2) inability to determine the best segmentation for an image. Consequently, segmentation evaluation plays an important role in research related to image segmentation. In an effort, considerable work has done in generating hand-labeled segmentations of natural images [2] to compare the performance of different segmentation algorithms. However, it is still a challenge to put a numerical score for the performance of a segmentation algorithm. This is mainly because the image segmentation is an ill-defined problem for which there is no single ground truth segmentation. Moreover, depending on the application the best segmentation can be different for the same image.

The goal of this work is to define a standard quality measure that can be applied to automatically provide a ranking among different segmentation algorithms or to optimally set the parameters of a given algorithm, under a predefined framework. A quantitative measure for an objective image segmentation algorithm evaluation based on entropy in information theory is proposed in this paper. The novelty of this work is that it not only gives a score demonstrating the results if the pixels of the segmented regions are uniform or similar in their own region and dissimilar compared to the other regions but also the given score exhibits the degree of similarity and dissimilarity which has been accomplished by putting a weight in the standard entropy expression.

2. Related work

A number of techniques have already been proposed in the literature for the quantitative evaluation of segmentation methods. There are three main categories or groups for the classification of these techniques [1]. The first category is called the analytic methods in which the evaluation is done by measuring some intrinsic properties of the segmentation algorithm, for example, the stability or the computational complexity. This measure does not evaluate the segmentation accuracy rather in an indirect way. The second group is called empirical discrepancy methods which includes supervised evaluation methods. In these methods, the evaluation is done by comparing the segmentation results to the ground truth or the reference image which is generated by hand. For the recent objective evaluation algorithms the second category is the most commonly used method [3,4]. However, producing a hand-labeled reference image or ground truth is error prone, time consuming and difficult. It varies from person to person and thus one unique hand-generated segmentation cannot be guaranteed

* Corresponding author. Tel.: +1 334 4399105.
E-mail addresses: khanj@mytu.tuskegee.edu, mimi_toha@yahoo.com (J.F. Khan).

0030-3992/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.optlastec.2013.07.012

as a ground truth for most of the images. Therefore, this measure requires comparing with multiple reference images which makes it complex and unreliable. The third group is called empirical goodness methods which includes unsupervised evaluation methods [5] in which the segmentation is considered as a part of a larger system, for example, object recognition, image querying, image reconstruction, etc and evaluated based on the performance of the larger system. As a specific example, for object recognition system the segmentation results are evaluated by judging the quality of the segmented image such as the partitioning of foreground objects from the background. This strategy for comparison becomes unfair, inconsistent and inappropriate when the segmentation algorithm is designed for different applications.

Recently, to overcome the limitations of the existing segmentation evaluation methods, a few segmentation evaluation methods have been proposed based on an information theoretic approach called entropy [6,1,7]. Hao et al. [6] use region entropy and segmentation entropy for the segmentation evaluation method. The segmentation entropy is proposed as the difference between the summation of all the region entropies and the whole image entropy. Zhang et al. [1] use expected entropy, H_{ere} and layout entropy, H_{lay} and the summation of these two entropies, $Q=H_{ere}+H_{lay}$ for the segmentation evaluation. Lai et al. [7] add weights w_1 and w_2 with the expected entropy, H_{ere} and layout entropy, H_{lay} respectively and the summation of these two weighted entropies $Q=w_1H_{ere}+w_2H_{lay}$ are used for the segmentation evaluation. Entropy is a measure of uncertainty of a random variable, hence, it can be used for the measurement of the degree of homogeneity among the pixels within an image. Any good segmentation method tries to gather the homogeneous pixels in a region, whereas to put the non-homogeneous pixels in different regions. Hence, entropy, which has the natural characteristic to measure the disarrangement, should be a good choice to use for segmentation evaluation.

3. Weighted entropy based evaluation method

Based on the above facts, the existing methods based on standard entropy give a binary decision of “good” or “bad”, but not in what extent. But in the proposed method, the standard entropy definition has been modified by putting a weight relative to the goodness of uniformity of the pixels in a region and non-uniformity across the regions. In the next section it is shown how the final scores of the novel weighted self-entropy and mutual entropy exhibits the degree of likeness of the pixels in a region and unlikeness compared to other regions.

According to information theory, for a random variable X with $X^{(x)}$) as the set of all possible outcomes, the entropy (the measure of uncertainty) denoted by $H(X)$ is defined as

$$H(X)=-\sum_{t\in X^{(x)}} p(t)\log\, p(t). \tag{1}$$

where $p(t)$ is the probability mass function of outcome t. Thus $H(X)$ is the number of bits needed to encode the random variable X.

This can be applied to an image. For a given image, I, if the pixel feature, G is considered as a random variable with $G^{(g)}$ as the set of all possible values, the total entropy for that image associated with that feature can be expressed as

$$H(I)=-\sum_{t\in G^{(g)}} p(t)\log\, p(t)=-\sum_{t\in G^{(g)}}\frac{N(t)}{M}\log\,\frac{N(t)}{M} \tag{2}$$

where $N(t)$ is the number of pixels in the image that have a value of t for the feature G and M is the total number of pixels of that image. That means, $N(t)/M$ represents the probability that a pixel in the image has a feature value of t. Thus $H(I)$ is the number of bits per pixel needed to encode the feature for the image I.

Now it is explained what the entropy of an image actually means by using a simple example. Considering that the intensity is the pixel feature, if all the pixels of an image have the same intensity that means the entropy is zero and there is no uncertainty about that image. Whereas, if all the pixel have different intensity then the magnitude of the entropy is the highest and the image has the most uncertainty. Now the idea can be extended to use entropy for the evaluation of an image segmentation algorithm by measuring the uniformity of the pixels in a segmented region and the non-uniformity among different regions.

Consider a segmented image with a number of homogeneous regions. For any segmented region i of the image, G is defined as one of the features to describe the pixels and $G_i^{(g)}$ as the set of all possible values for the feature G in region i and $n_i^{(g)}$ as the total number of elements in $G_i^{(g)}$. Now, for the segmented region i, t is defined as a value for the feature G in that region, $N_i(t)$ as the number of pixels in the region i that have a value t for feature G (e.g. pixel intensity) and M_i as the total number of pixels in the segmented region i. Then, the self-entropy for the region i can be defined as

$$H_G(i)=-\sum_{t\in G_i^{(g)}}\frac{N_i(t)}{M_i}\log\,\frac{N_i(t)}{M_i}, \tag{3}$$

where, $N_i(t)/m_i$ represents the probability that a pixel in region i has a feature (e.g. pixel intensity or some other feature) value of t. Thus $H_G(i)$ is the number of bits per pixel needed to encode the feature for region i, given that the region i is known.

To measure the homogeneity of the pixels within the segmented regions of an image I, the above described self-entropy can be used. When each segmented region is homogeneous then self-entropy for all regions will be small, for example, when all pixel intensities in a region have the same value, then the entropy for that region will be 0. On the other hand the entropy for a region will be the maximum when all the pixels will have a different value. Since an over-segmented image will have a very small self-entropy for all the regions, mutual-entropy needs to be counted simultaneously to assess the segmentation performance, where for a good segmentation all the mutual-entropy among all the regions should be large.

Following the above definition of self-entropy, for mutual entropy between any two regions j and k of the segmented image, $G_{jk}^{(g)}$ is defined as the set of all possible values associated with feature G in region j and k. $n_{jk}^{(g)}$ is denoted as the total number of elements in $G_{jk}^{(g)}$. Then, for regions j and k of the segmentation and value t of feature G in those regions, $N_{jk}(t)$ is used to denote the number of pixels in regions j and k that have a value of t for feature G and M_{jk} as the total number of pixels in the regions j and k. Then, the mutual-entropy between regions j and k then can be defined as

$$H_G(jk)=-\sum_{t\in G_{jk}^{(g)}}\frac{N_{jk}(t)}{M_{jk}}\log\,\frac{N_{jk}(t)}{M_{jk}}, \tag{4}$$

where $N_{jk}(t)/M_{jk}$ represents the probability that a pixel in regions j and k has a feature (e.g. luminance or other feature) value of t. Thus $H_G(jk)$ is the number of bits per pixel needed to encode the feature for regions j and k, given that the regions j and k are known.

Now the limitation of the standard entropy measure is presented to evaluate segmentation quality successfully. Also a weighted entropy will be proposed to be more effective in segmentation evaluation. As shown in Fig. 1, there are three images that have the same segmentation results. For the three images in Fig. 1(a)–(c); the corresponding histograms are given in Fig. 1(d), (e) and (f), respectively. The same segmented image resulting from those three images is presented in Fig. 1(g). From