Optik 168 (2018) 446–457

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

A novel Gini index based evaluation criterion for image segmentation

Maryam Habba, Mustapha Ameur, Younes Jabrane*

GECOS Lab, Cadi Ayyad University, Marrakesh, Morocco

ARTICLE INFO

Article history:
Received 20 June 2017
Received in revised form 5 February 2018
Accepted 10 April 2018

Keywords:
Image segmentation
Objective evaluation
Gini index
Entropy
Multilevel thresholding
Particle swarm optimization

ABSTRACT

Image segmentation is the most important operation in image processing that changes the representation of an image to be useful in much computer vision applications. During the last years, several image segmentation algorithms have been developed, consequently, a great number of techniques for evaluating segmentation results have been proposed. Unfortunately, the majority of these methods is subjective and cannot be used to judge the performance of different segmentation algorithms. In this paper, we propose a novel evaluation criterion based on the Gini index and the entropy calculation. This new method permits evaluating the conjunction of the regions arrangement in the segmented image, and measuring the pixels homogeneity within each region. In experiments, benchmark images segmented by the multilevel thresholding technique based on particle swarm optimization are used to conclude the strength, the effectiveness and the rapidity of the proposed evaluation criterion.

1. Introduction

Image segmentation plays an important role in the image processing applications. It allows having proper and comprehensive images which are easier to analyze [1,2]. Image segmentation is either used to separate an image into distinct regions, where each one contains homogeneous pixels, or to extract contours. A correct interpretation of the image depends always on the goodness of the segmentation results.

In the last few years, hundreds of segmentation algorithms have been proposed in the literature [3]. Unfortunately, there is no perfect and exact solution for each segmentation process, the raison behind is that the images are very diverse; they can be degraded, textured or uniform. Thus, no specific segmentation technique is suitable for all images and for all applications. Moreover, just a little number of these techniques provides a good segmentation results. Since this is the way, an effective evaluation of the segmentation results is highly required. However, there is no standard method to design an evaluation function and its conception is a hard task. A most part of the evaluation studies are subjective as they are based only on visual and qualitative criteria, and the results which they provide are biased, time consuming and sometimes wrong. To overcome all these problems, various functions and criteria have been developed such as: the probabilistic rand coefficient [4], the Rosenberger criterion [5], and the contrast of Zeboudj criterion [6]. These functions and criteria are considered as reliable, trustworthy and objective as they are based on a quantitative measurement of the segmented image. Objective evaluation methods are splitted into two categories; analytical objective methods and empirical objective methods. The first category

* Corresponding author.
E-mail address: y.jabrane@uca.ma (Y. Jabrane).

https://doi.org/10.1016/j.ijleo.2018.04.045
0030-4026/

requires previous knowledge and grounding to evaluate segmented images. The second category performs evaluation on the images either based on discrepancy measurements or goodness measurements, this category is widely used because it can evaluate automatically and quickly a numerous set of segmented images [7].

In this paper, we propose a novel evaluation criterion based on the Gini index and the entropy calculation. The multilevel thresholding technique based on particle swarm optimization (PSO) is used to segment a given image into sets of regions. Each one of the regions must be inwardly homogeneous and outwardly distinguished from its neighbor. In this method, the Gini index is used to measure the homogeneity of the pixels within a region, and the entropy calculation is used to evaluate the conjunction of the regions arrangement. The proposed method proved, over the experiments, its rapidity, accuracy and reliability.

The remainder of this paper is organized as follows: a state of the art of the empirical evaluation methods is presented in Section 2. In Section 3, the proposed Gini-entropy evaluation criterion is described. Experimental results and comparisons are discussed in Section 4. Finally, conclusions are drawn in Section 5.

2. State of the art

Empirical objective evaluation methods are classified into two categories, discrepancy and goodness methods; they are also named supervised and unsupervised methods, respectively.

Discrepancy or supervised evaluation methods [8], also known as relative evaluation methods [9], evaluate a segmented image on the basis of a reference image or a ground truth image predetermined and hand drawn by experts. These methods measure the success of segmentation by computing the similarities and the dissimilarities between the two set of images. Supervised methods are the most widely used objective evaluation functions, the reason behind is that the comparison between the ground truth image and the segmented image assures a finer resolution of the evaluation process. However, these methods can be subjective when it is difficult to generate a reliable and unique hand made reference image.

In the literature, a variety of empirical supervised methods have been proposed. Everinghan et al. [10] used the Pareto front to develop a new objective evaluation method; in a multi-dimensional search space, they used multiple supervised functions to evaluate the segmentation results instead of using one supervised function in one dimensional discrepancy space. Receiver operating characteristics (ROC) curve [11] is one of the commonly used relative methods. On the basis of the confusing matrix, ROC compares the segmented image and the reference image. In medical imagery, the dice co-efficient [12] is used to judge the accuracy of output images that have ground truth data sets. Jaccard co-efficient [13], similar to the dice co-efficient, computes the similarities and the dissimilarities between two images. In the case when the number of the objects in the segmented image and the output image are not the same, some discrepancy methods have been developed to accommodate this problem [14]. Other group of supervised methods evaluates the efficiency of the segmentations by comparing the differences between the feature values of the segmentation results and the reference image [15].

Goodness or unsupervised methods evaluate a segmented image without any previous grounding or references. They allow the quantification of the accuracy of a segmentation result by computing some statistics (uniformity, entropy, variance) in each region of the segmented image on the basis of its characteristics such as luminance or intensity, and by measuring the disparity of the regions. Unsupervised methods are very useful when there is no information about the ground truth images.

Among the most frequent unsupervised methods that have been used in several segmentation cases is the entropy calculation. The goal is to measure the disorder and the disturbance of a segmented image by computing the grey level values probabilities of the pixels in each region. Lower value of the entropy means that the image is well segmented and vise versa. Zhang et al. [16] used the concept of the entropy to propose a novel objective evaluation function. It is based on the measurement of the pixels uniformity in each region in the segmented image, and on the calculation of the layout entropy of the segmentation results. Nie et al. [17] developed a new entropy based bilevel thresholding method based on the Masi entropy [18]. In this method, the uniformity of each region in the segmented image is evaluated by measuring the non-extensive property of the information existing in the pixels of each region. Based on the same concept of the entropy calculation, Singh and Zwiggelaar [19] suggested another evaluation criterion based on the Gini index. In this approach, the Gini index of an image is calculated at a pixel value, resulting in a "G" image where zero G value occurs within homogeneous regions, and high G value occurs at contrasting boundaries. The optimal segmentation result is obtained by using the multilevel thresholding technique based on the aferementioned "G" image, and the region growing segmentation method using 8-neighborhood pixels connectivity. In [20] a novel multilevel thresholding method based on multiobjective particle swarm optimization (MOPSO) is proposed; this method is based on the simultaneous minimization of the layout entropy and the entropy of the color errors of the segmented image. De Albuquerque et al. [21] suggested another multilevel thresholding selection based on Tsallis entropy. Khairuzzaman et al. [22] developed a new multilevel thresholding method based on Grey wolf optimizer using the Otsu's between class variance and the Kapur entropy. In [23], a new two dimensional image segmentation based on the genetic algorithm (GA) and the flexible representation of Tsallis and Renyi's entropies was proposed. Indeed, the entropy calculation is used to measure the amount of information that exists in the two dimensional histogram of the image, and the GA is used to maximize the said entropy.

Always in order to evaluate the reliability of a segmentation result, metrics and criteria based on the intuition have been proposed in the literature. Based on the principle of the intra-region uniformity in a segmented image [24], Levine and Nazif [25] defined their own criterion by which they calculate the variance of the pixels feature (luminance for example)