Digital Signal Processing 28 (2014) 120–135

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Maximum similarity thresholding

Yaobin Zou [a,b,c,*], Fangmin Dong [a,c], Bangjun Lei [a,b], Shuifa Sun [a,b], Tingyao Jiang [a,b], Peng Chen [a,b]

[a] *Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang, Hubei 443002, China*
[b] *Collaborative Innovation Center for Geo-Hazards and Eco-Environment in Three Gorges Area, Hubei Province, China*
[c] *Collaborative Innovation Center for Key Technology of Smart Irrigation District in Hubei, China Three Gorges University, China*

ARTICLE INFO

Article history:
Available online 20 February 2014

Keywords:
Image thresholding
Image similarity
Pearson correlation coefficient
Otsu thresholding
Minimum error thresholding

ABSTRACT

Otsu method is one of the most popular image thresholding methods. The segmentation results of Otsu method are in general acceptable for the gray level images with bimodal histogram patterns that can be approximated with mixture Gaussian modal. However, it is difficult for Otsu method to determine the reliable thresholds for the images with mixture non-Gaussian modal, such as mixture Rayleigh modal, mixture extreme value modal, mixture Beta modal, mixture uniform modal, comb-like modal. In order to determine automatically the robust and optimum thresholds for the images with various histogram patterns, this paper proposes a new global thresholding method based on a maximum-image-similarity idea. The idea is inspired by analyzing the relationship between Otsu method and Pearson correlation coefficient (PCC), which provides a novel interpretation of Otsu method from the perspective of maximizing image similarity. It is then natural to construct a maximum similarity thresholding (MST) framework by generalizing Otsu method with the maximum-image-similarity concept. As an example, a novel MST method is directly designed according to this framework, and its robustness and effectiveness are confirmed by the experimental results on 41 synthetic images and 86 real world images with various histogram shapes. Its extension to multilevel thresholding case is also discussed briefly.

1. Introduction

Image segmentation is a critical preprocessing step in computer vision and image understanding. After reviewing various methods for gray level image segmentation, Pal and Pal [1] state that image thresholding is a popular segmentation method because of its simplicity and ease of implementation. There are two types of thresholding methods, i.e., global thresholding and local thresholding. Generally, a local thresholding method better suits poor and unevenly illuminated images [2,3]. However, a global thresholding approach is a more appropriate choice for images in which the object and background can be separated with an optimal threshold [4–6]. Although there are many algorithms to select the segmentation threshold of global thresholding, automatic selection of a robust, optimum threshold remains an interesting and challenging task [7,8]. In this paper, we will focus on global thresholding methods.

Global thresholding methods compare each pixel in a gray level image with a calculated threshold. Thus, how to determine an optimal threshold becomes a core problem of global thresholding segmentation [5]. Many methods have been proposed to automatically determine the segmentation threshold over the past several decades [6–20]. An early review of thresholding approaches was reported in [21]. A comparative performance study of global thresholding techniques was reported by Lee et al. [22]. Another comparative analysis of the performance of eleven histogram-based thresholding methods was carried out by Glasbey [23]. Specially, an exhaustive survey of image thresholding methods was conducted recently by Sezgin and Sankur [24], where the readers can discover more comprehensive information about image thresholding methods.

Among these global thresholding approaches, the two most popular ones are a minimum error thresholding (MET) method [12] and an Otsu method [13]. The two methods are widely used in practice and highly cited in scientific publications [25]. MET method finds the optimum threshold by optimizing the average pixel classification error rate. This method assumes that an image is characterized by a mixture Gaussian distribution. Otsu method utilizes the discriminant analysis to find the maximum separability

* Corresponding author at: Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, 8 University Avenue, Yichang 443002, China.
E-mail address: bwin2010@gmail.com (Y. Zou).

http://dx.doi.org/10.1016/j.dsp.2014.02.008
1051-2004/

of two classes. For every possible gray level, this method evaluates the goodness of this value if it is used as a threshold. Recently, several interesting views of Otsu method have been reported, which enhances understanding of the properties and thresholding performance of Otsu method [15,25,26]. From the perspective of maximum likelihood estimation, Kurita et al. [15], Xue and Zhang [25] show that Otsu method can be viewed as a special case of MET method, in which case equal class sizes, equal class variances, and image histogram composed of two Gaussian distribution are assumed. In addition, Xue and Titterington [26] provide a statistical-hypothesis-testing view of Otsu method by revealing the relationships between t-tests, F-tests, and Otsu method.

MET method is based on the Gaussian assumption for modeling the class distributions in the gray level image. When the assumption of Gaussian distribution is strictly satisfied, MET method will succeed to determine a reasonable threshold. However, its effectiveness is strongly reduced when the prior probabilities of object and background classes are far away from the Gaussian distribution [4]. Otsu method seems not to depend on any special assumption about the class distributions. However, after analyzing the relationship between Student's t-test and Otsu method, Xue and Titterington [26] recently reported that two assumptions should be satisfied if Otsu method wants to obtain an optimal threshold. One assumption is that the prior probabilities of object and background classes follow Gaussian distribution with equal class variances. The other assumption is that object and background classes are of equal size. When the assumptions of equal class sizes and equal class variances are apparently violated, Otsu threshold tends to split the class with a larger size, and to bias towards the class with a larger variance [14,25,26].

The histograms of gray level images show various patterns, such as mixture Gaussian modal (see Figs. 2(d)–(e) and Figs. 4(c)–(d)), mixture Rayleigh modal (see Fig. 6(g)), mixture extreme value modal (see Fig. 7(g)), mixture Beta modal (see Fig. 8(g)), mixture uniform modal (see Fig. 9(g)), comb-like modal (see Fig. 10(g)). Although the segmentation results of MET and Otsu methods are in general acceptable for the gray level images with mixture Gaussian or approximate Gaussian modal, these assumptions mentioned above will limit their segmentation accuracy and their application areas, since these assumptions often do not hold in real world images. Some improved methods have been proposed, such as recursive Otsu method [9], 2-dimensional Otsu method [16], two-stage Otsu method [17], expectation-maximization method [4], valley-emphasis method [18]. However, it is still difficult for these improved methods to determine the reliable thresholds for the images with mixture non-Gaussian modal.

In order to determine automatically the robust and optimum thresholds for the images with various histogram patterns, this paper proposes a new global thresholding method based on a maximum-image-similarity idea. The idea is inspired by analyzing the relationship between Otsu method and Pearson correlation coefficient (PCC) [27,28], which provides a novel interpretation of Otsu method from the perspective of maximizing image similarity, i.e. Otsu method is to search for a binary image as the output that is most similar to the input gray level image, and PCC is used for the similarity measure. It is then natural to construct a maximum similarity thresholding (MST) framework by generalizing Otsu method with the maximum-image-similarity concept. The MST framework transforms the optimum threshold selection issue to the computation of image similarity, which facilitates us to develop new image thresholding methods with the image similarity theories. According to this framework, a novel thresholding method called MST is directly designed, and its robustness and effectiveness are confirmed by the experimental results on the synthetic and real world images.

The rest of the paper is organized as follows. Section 2 describes some symbol specifications. Section 3 analyzes the relationship between PCC measure and Otsu method in detail. Section 4 proposes the MST framework, and Section 5 introduces the proposed MST method. In Section 6, 41 synthetic images and 86 real world images are used to verify the robustness and effectiveness of the proposed MST method. Section 7 discusses the multilevel thresholding extension of MST method. Finally, some conclusions and future works are described in Section 8.

2. Symbol specifications

Let χ denote a gray level image with N pixels, x_i represent the gray level of the ith pixel, and L is the largest gray level in this image. A threshold t classifies the image χ into the background class $C_1(t)$ and the object class $C_2(t)$, where $C_1(t) = \{i \mid 0 \leqslant x_i \leqslant t, 1 \leqslant i \leqslant N\}$ and $C_2(t) = \{i \mid t < x_i \leqslant L, 1 \leqslant i \leqslant N\}$. In addition, let the gray level histogram of the image χ be denoted with $h(x)$ that subjects to the following constraint:

$$\sum_{x=0}^{L} h(x) = 1 \tag{1}$$

A binary image obtained by thresholding the image χ with t is denoted with $\gamma(t)$, which is defined as:

$$y_i(t) = \begin{cases} 0 & \text{if } i \in C_1(t) \\ 1 & \text{if } i \in C_2(t) \end{cases} \tag{2}$$

where $y_i(t)$ represents the ith pixel value in the binary image $\gamma(t)$.

Let $\omega_1(t)$ and $\omega_2(t)$ denote the proportions of pixels representing classes $C_1(t)$ and $C_2(t)$, respectively. $\omega_1(t)$ and $\omega_2(t)$ are defined as:

$$\omega_1(t) = \sum_{x=0}^{t} h(x) \tag{3}$$

$$\omega_2(t) = \sum_{x=t+1}^{L} h(x) \tag{4}$$

Let $\mu_1(t)$ and $\mu_2(t)$ denote the average gray levels of pixels representing classes $C_1(t)$ and $C_2(t)$, respectively. $\mu_1(t)$ and $\mu_2(t)$ are given by:

$$\mu_1(t) = \sum_{x=0}^{t} xh(x)/\omega_1(t) \tag{5}$$

$$\mu_2(t) = \sum_{x=t+1}^{L} xh(x)/\omega_2(t) \tag{6}$$

Let μ_χ and σ_χ^2 denote the mean and the variance of the image χ, respectively. μ_χ and σ_χ^2 are defined as:

$$\mu_\chi = \sum_{x=0}^{L} xh(x) \tag{7}$$

$$\sigma_\chi^2 = \sum_{x=0}^{L} (x - \mu_\chi)^2 h(x) = \sum_{i=1}^{N} (x_i - \mu_\chi)^2/N \tag{8}$$

3. Relationship between PCC and Otsu method

PCC measure is widely used in statistical analysis, pattern recognition, and image processing. For two gray level images χ and ψ with the same size, PCC measure is defined as [27,28]: