Pattern Recognition 47 (2014) 1330–1348

Contents lists available at ScienceDirect

Pattern Recognition

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

A generalized multiclass histogram thresholding approach based on mixture modelling

Aïssa Boulmerka [a], Mohand Saïd Allili [b,*], Samy Ait-Aoudia [a]

[a] École Nationale Supérieure en Informatique, Algiers, Algeria
[b] Université du Québec en Outaouais, Gatineau, QC, Canada

ARTICLE INFO

Article history:
Received 4 February 2013
Received in revised form
13 July 2013
Accepted 10 September 2013
Available online 19 September 2013

Keywords:
Image segmentation
Multi-modal class thresholding
Mixture of generalized Gaussian distributions (MoGG)

ABSTRACT

This paper presents a new approach to multi-class thresholding-based segmentation. It considerably improves existing thresholding methods by efficiently modeling non-Gaussian and multi-modal class-conditional distributions using mixtures of generalized Gaussian distributions (MoGG). The proposed approach seamlessly: (1) extends the standard Otsu's method to arbitrary numbers of thresholds and (2) extends the Kittler and Illingworth minimum error thresholding to non-Gaussian and multi-modal class-conditional data. MoGGs enable efficient representation of heavy-tailed data and multi-modal histograms with flat or sharply shaped peaks. Experiments on synthetic data and real-world image segmentation show the performance of the proposed approach with comparison to recent state-of-the-art techniques.

1. Introduction

Thresholding-based image segmentation is a well-known technique that is used in a broad range of applications, such as change detection [20], object recognition [3,34] and document image analysis [26], to name a few. Image thresholding aims at building a partition of an image into K classes, $C_1, \ldots, C_K$, which are separated by $K-1$ thresholds $T_1, \ldots, T_{K-1}$. In case of $K=2$, the image is segmented into foreground and background regions. In case of $K>2$, the image is segmented into K distinct regions. In most of existing thresholding methods, the parameter K is generally given and it corresponds to the number of histogram modes [27]. Comparative studies about existing thresholding techniques applied to image segmentation can be found in [10,24,27,32].

Among the most popular methods for image thresholding are the standard Otsu's method [22] and Kittler and Illingworth's method [14]. While the former uses inter-class separability to calculate optimal thresholds between classes, the latter is based on the minimization of Bayes classification error, where each class is modeled by a Gaussian distribution. Both methods, however, assume a uni-modal shape for classes and use sample mean and standard deviation (i.e., the parameters of a Gaussian) to approximate their distributions. In [12,32], the authors established the relationship between the two methods, where these parameters can be obtained in either methods using maximum likelihood estimation of a Gaussian model for each class. *Entropy* and *relative entropy* can also be used to derive good thresholds for image segmentation when the distribution of classes is Gaussian [5,6,25]. For example, Jiulun and Winxin [12] gave a relative-entropy interpretation for the minimum error thresholding (MET) [14,19]. In that work, the Kullback–Leibler divergence [15] is used to measure the discrepancy between histograms of a source image and a mixture of two Gaussians. Recently, Xue and Titterington [30] proposed a thresholding method where class data are modeled by Laplacian distributions. They showed that the obtained thresholds offer better separation of classes when their distributions are skewed, heavy-tailed or contaminated by outliers. Indeed, the location and dispersion parameters of the Laplacian distribution are the median and the absolute deviation from the median, which are more robust to outliers compared to the sample mean and standard deviation, respectively [11].

Previous methods for image thresholding were basically devised to separate classes that are unimodal [27]. Therefore, they are not adapted to multi-modal class segmentation. For instance, in many segmentation problems (e.g., medical images, and remote sensing), one might want to separate the image foreground from a background region, each of which may have a multi-modal distribution. Another limitation for the standard methods [14] and [22], and as pointed out in [30], lies in the assumption that class data are Gaussian. In several image examples, one can find histogram modes that are skewed, sharply peaked or heavy tailed, making the assumption of Gaussian-distributed classes not realistic. Recently, researchers have used other distribution types to

* Corresponding author. Tel.: +1 819 595 3900x1601.
E-mail addresses: a_boulmerka@esi.dz (A. Boulmerka), mohandsaid.allili@uqo.ca (M. Saïd Allili), s_ait_aoudia@esi.dz (S. Ait-Aoudia).

0031-3203/$ - see front matter
http://dx.doi.org/10.1016/j.patcog.2013.09.004

provide better image thresholding methods by modeling histogram classes using, for instance, Poisson [23], generalized Gaussian [2,7,8], skew-normal [31] and Rayleigh [29] distributions. However, these approaches are also built on the assumption that all classes are unimodal. Worth mentioning is the parallel trend of using mixture methods for segmentation (ex. [1,21,35,36]), where data are clustered to classes determined by the components of a learned mixture model. For such works, the number of classes (which correspond to the number of mixture components) can be estimated using information-theoretic criteria such as AIC, BIC, MML, etc. [18]. This paper deals with a different problem which consists of finding thresholds between classes with distributions that can be constituted of arbitrary numbers of (non-Gaussian) histogram modes. Thus, contrary to [1,21], the number of classes K will not necessarily correspond to the number histogram modes.

In this paper, we propose a new thresholding approach that performs segmentation for multi-modal classes with arbitrarily shaped modes. We generalize the aforementioned state-of-art techniques, based on using single probability density functions (pdf's), to mixtures of generalized Gaussian distributions (MoGG's). The Generalized Gaussian Distributions (GGD) is a generalization of the Laplacian and the normal distributions in that it has an additional degree of freedom that controls its kurtosis. Therefore, histogram modes, ranging from sharply peaked to flat ones, can be accurately represented using this model. Furthermore, skewed and multi-modal classes are accurately represented using mixtures of GGDs. We propose an objective function that finds optimal thresholds for multi-modal classes of data. It also extends easily to arbitrary numbers of classes ($K>2$) with reasonable computational time. Experiments on synthetic data, as well as real-world image segmentation, show the performance of the proposed approach.

This paper is organized as follows: Section 2 presents state-of-the-art theory for thresholding techniques. In Section 3 we outline our proposed approach for image thresholding. Experimental results are given in Section 4. We end the paper with a conclusion and some future work perspectives.

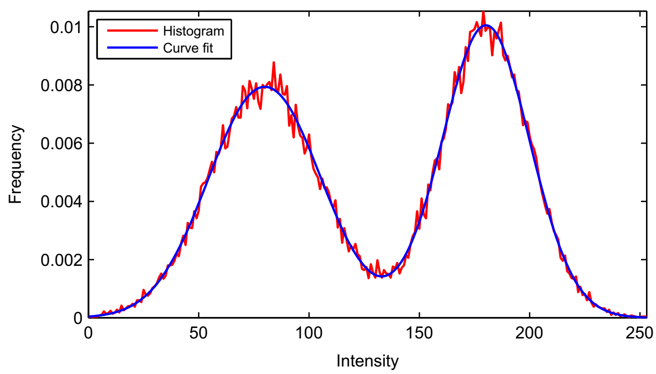

Fig. 1. Bimodal histogram ($K=2$).

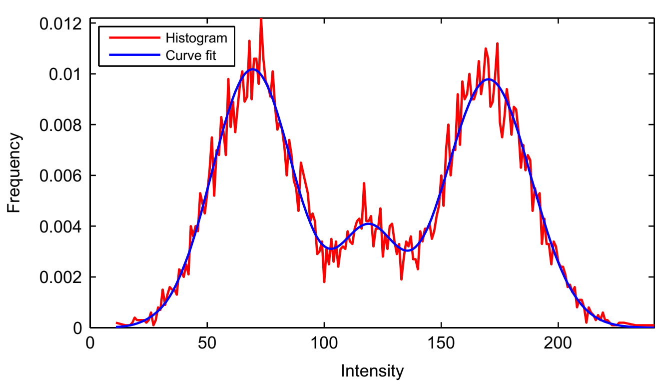

Fig. 2. Multimodal histogram ($K=3$).

2. General formulation of the Otsu's method (case K=2)

Let $\mathcal{X}=\{x_1,x_2,\dots,x_N\}$ be the gray levels of the pixels of an image I of size $N=H\times W$; H and W being the height and the width of the image. Let $\mathbf{t}=(t_1,t_2,\dots,t_{K-1})$ be a set of thresholds that partitions an image into K classes. First we consider the simple case of $K=2$. The most general case of $K>2$ will be elaborated later in this paper. In the case of $K=2$, one threshold t yields two classes $\mathcal{C}_1(t)=\{x:0\le x\le t\}$ and $\mathcal{C}_2(t)=\{x:t+1\le x\le T\}$, where T is the maximum gray level. Finally, we denote by $h(x)$ the histogram frequency of the gray level x, where $\sum_{x=0}^{T}h(x)=1$. The resulting histogram in this case ($K=2$) is bimodal, as shown in Fig. 1. Otsu's method [22] determines the optimal threshold t using discriminant analysis, by maximizing inter-class variation, or equivalently minimizing intra-class variation.

A generalized formula of the Otsu's method for $K=2$ can be defined as follows (see refs. [10,14,22,27,30,32]):

$$\sigma_B^2(t)=\arg\min_t\{\omega_1(t)V_1(t)+\omega_2(t)V_2(t)\}, \tag{1}$$

where, we have

$$\begin{cases}\omega_1(t)=\sum_{x=0}^{t}h(x)\\ \omega_2(t)=\sum_{x=t+1}^{T}h(x)=1-\omega_1(t)\end{cases}, \tag{2}$$

and

$$\begin{cases}V_1(t)=\dfrac{1}{\omega_1(t)}\sum_{x=0}^{t}h(x)\|(x-m_1(t)\|_\beta\\ V_2(t)=\dfrac{1}{\omega_2(t)}\sum_{x=t+1}^{T}h(x)\|(x-m_2(t)\|_\beta\end{cases}, \tag{3}$$

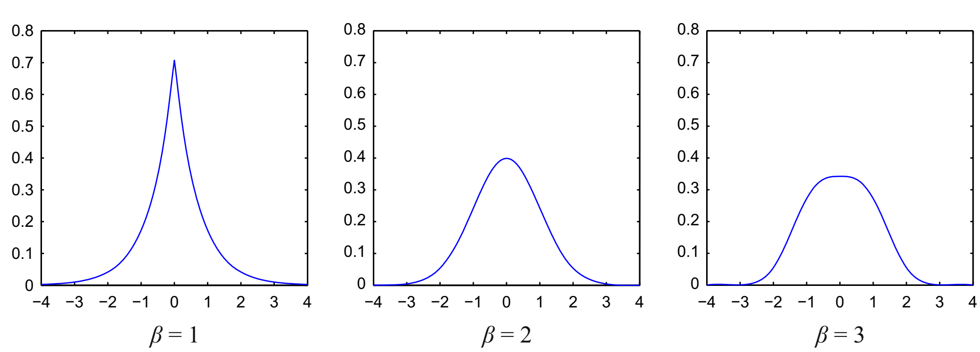

Fig. 3. Different shapes of the GGD distribution as a function of the parameter β (μ=0,σ=1).