Pattern Recognition 42 (2009) 843 – 856

Contents lists available at ScienceDirect

Pattern Recognition

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

# Image thresholding by variational minimax optimization

Baidya Nath Saha, Nilanjan Ray*

*University of Alberta, Edmonton, Alberta, Canada T6G 2E8*

ARTICLE INFO

*Article history:*
Received 23 October 2007
Received in revised form 19 August 2008
Accepted 19 September 2008

*Keywords:*
Minimax optimization
Variational calculus
Image thresholding

ABSTRACT

In this paper we introduce an adaptive image thresholding technique via minimax optimization of a novel energy functional that consists of a non-linear convex combination of an edge sensitive data fidelity term and a regularization term. While the proposed data fidelity term requires the threshold surface to intersect the image surface only at places with large image gradient magnitude, the regularization term enforces smoothness in the threshold surface. To the best of our knowledge, all the previously proposed energy functional-based adaptive image thresholding algorithms rely on manually set weighting parameters to achieve a balance between the data fidelity and the regularization terms. In contrast, we use minimax principle to automatically find this weighting parameter value, as well as the threshold surface. Our conscious choice of the energy functional permits a variational formulation within the minimax principle leading to a globally optimum solution. The proposed variational minimax optimization is carried out by an iterative gradient descent with exact line search technique that we experimentally demonstrate to be computationally far more attractive than the Fibonacci search applied to find the minimax solution. Our method shows promising results to preserve edge/texture structures in different benchmark images over other competing methods. We also demonstrate the efficacy of the proposed method for delineating lung boundaries from magnetic resonance imagery (MRI).



## 1. Introduction

Image thresholding is an image pixel labeling problem. It aims to classify pixels of an image into two classes: foreground and background. Typically, a pixel is classified as foreground if the image intensity at the pixel exceeds a threshold value; otherwise it is classified as a background pixel. Thresholding is a well-known pre-processing step for image segmentation. It is widely used in automatic target recognition, industrial applications of computer visions and medical/biomedical image analysis.

In general there are two types of image thresholding techniques available: global and local. In the global thresholding technique a gray-level image is converted into a binary image based on an image intensity value called global threshold which is fixed in the whole image domain whereas in local thresholding technique, threshold value can vary from one pixel location to next. Global thresholding converts an input image $I$ to a binary image $G$ as follows: $G(i,j)=1$ for $I(i,j)\geqslant T$, or, $G(i,j)=0$ for $I(i,j)<T$, where $T$ is the threshold, $G(i,j)=1$ for foreground and $G(i,j)=0$ for background. Whereas, for a local threshold, the threshold $T$ is a function over the image domain, i.e., $T=T(x, y)$. In addition, if in constructing the threshold value/surface the algorithm adapts itself to the image intensity values, then it is called dynamic or adaptive threshold. Thus, in a general setting, thresholding can be expressed as a test operation that tests against a function $T$ of the form [1]: $T=T[x,y,h,I]$, where, $I$ is the input image and $h$ denotes some local property of this point—for example, the average gray level of a neighborhood centered on $(x, y)$.

Threshold selection depends on the information available in the gray-level histogram of the image. An image function $I(x, y)$ can be expressed as a product of a reflectance function and an illumination function based on a simple image formation model. If the illumination component of the image is uniform then the gray-level histogram of the image is clearly bimodal, because the gray levels of object pixels are significantly different from the gray levels of the background. It indicates that one mode is populated from object pixels and the other mode is populated from background pixels. Then objects could be easily partitioned by placing a single global threshold at the valley at the histogram. However, in reality bimodality in histograms does not occur very often. Consequently, a fixed threshold level based of the information of the gray-level histogram will fail totally to separate objects from the background. In this scenario we turn our attention to adaptive local threshold surface where threshold value changes over the image domain to fit the spatially changing background and lighting conditions.

* Corresponding author.
*E-mail addresses:* baidya@cs.ualberta.ca, nray1@cs.ualberta.ca (N. Ray).

0031-3203/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2008.09.033

Over the years many threshold selection methods have been proposed. Otsu has suggested a global image thresholding technique where the optimal global threshold value is ascertained by maximizing the between-class variance with an exhaustive search [2]. Sahoo et al. [3] claim that Otsu's method is suitable for real world applications with regard to uniformity and shape measures. Though Otsu's method is one of the most popular methods for global thresholding, it does not work well for many real world images where a significant overlap exists in the gray-level histograms between the pixel intensity values of the objects and the background due to uneven and poor illumination. Kittler and Illingworth [4] characterize the image by a Gaussian mixture of foreground and background pixels and address a minimum error Gaussian-density fitting problem. They use either an exhaustive or an iterative search to optimize the average pixel classification error rate.

A local thresholding method is superior to the global ones for poorly and unevenly illuminated images. Niblack proposes a local thresholding technique based on the local mean and local standard deviation [5]. The drawback of this algorithm is to determine the size of the neighborhood that is set by the user and it depends on the information available in the images. The window size should be small enough to preserve the local details and at the same time, it should be large enough to suppress noise. One of the well-known local thresholding methods is to fit a plane or biquadratic function to match the background gray-level variations [6] for unevenly illuminated images. A more advanced way is to generate a threshold surface where the threshold level changes dynamically over the image pixel to pixel [7]. Milgram et al. use gradient or edge information to segment images and assumed that different objects may have different thresholds, but each object has a fixed threshold with respect to its background [8].

Yanowitz and Bruckstein [9] have incorporated the concepts of threshold surface and suggested an algorithm where the threshold surface is determined by the interpolation of the intensity values at high magnitudes of image intensity gradient. The computational complexity of the successive over-relaxation method used in Yanowitz and Bruckstein's algorithm is prohibitively expensive: O ($N^3$) for an $N \times N$ image. Moreover, the technique uses a mean filter in the processing stage to eliminate the noise that reduces the contrast of the image and consequently affects the segmentation results. Also, there is no good algorithm to choose the value for the gradient magnitude threshold. Choosing an improper gradient threshold leads to false object boundary points adversely affecting the subsequent steps of the algorithm.

Yan et al. [10] propose a multistage adaptive thresholding where they consider two global thresholds: pixels having gray values lower than the low threshold value are considered as the background, pixels with intensity greater than the high threshold value are considered as objects and the pixels with gray values in between the two threshold values are considered based on local image statistics of mean and variance within a variable neighborhood. The values of two global thresholds are determined using Otsu's multilevel threshold with exhaustive search technique or percentile statistics. The window size of the neighborhood is determined by ad hoc trial and error method.

Chan et al. [11] propose a variational formulation of adaptive thresholding technique based on the principle advocated by Yanowitz and Bruckstein [9]. Whereas Yanowitz and Bruckstein's method involves pre- and post-processing steps followed by solving a dense-matrix inversion, Chan et al. [11] reduce the solution by solving a Poisson equation with only one user supplied scalar parameter.

The interpolation in Yanowitz and Bruckstein's algorithm does not consider the relationship between the threshold surface and the image surface which causes either black stains on white background or vice versa. Liu et al. [12] propose an active surface-based adaptive thresholding algorithm by a repulsive external force. In this model, a Gaussian external repulsive force is designed to keep the active surface away from the image surface while the smoothness constraint and the boundary condition on the threshold surface restrain it from moving far away. Shen and Ip [13] use a Hopfield neural network for an active surface paradigm like Liu et al. [12]. Li et al. compare different thresholding algorithms for segmenting aurora oval boundary from spacecraft UV imagery and they show that local adaptive thresholding algorithm show better results than other thresholding techniques [14]. Sezgin and Sankur have made a recent survey on image thresholding [15].

To the best of our knowledge, all the available local thresholding techniques, including the ones using energy functional minimization, have manually tuned parameters that need to be adjusted for differently illuminated images by trial and error method. The values of these parameters vary significantly for different images. To mitigate the effort of tuning parameters, in this paper, we introduce an adaptive local image thresholding method based on energy functional minimization, where the weighting parameters for the data and the regularization terms do not need to be supplied by a user. The proposed energy functional consists of a non-linear convex combination of a data term and a regularization term. The data term encourages the threshold surface to intersect the image surface at high gradient location and the regularization term imposes smoothness on the threshold surface. In order to find out the solution, i.e., the desired threshold surface, we propose a variational minimax (VM) optimization. The VM algorithm consists of two interleaved iterative steps: maximization with respect to the weighting parameter and variational minimization with respect to the threshold surface. We demonstrate that our conscious choice of the energy functional creates a unique saddle point and the VM algorithm finds this saddle point containing the desired solution. We guarantee this unique saddle point by making the energy functional strictly concave with respect to the weighting parameter and strictly convex with respect to the threshold surface. A short version of this work has been communicated in Ref. [16].

It is worth mentioning that Gennert and Yuille have proposed a generic minimax solution for non-convex energy functional by repeated Fibonacci search [17]. Their method consists of multiple minimization computations for different values of the weighting parameter dictated by a Fibonacci sequence. In contrast, the proposed VM method avoids multiple minimization computations by virtue of a deliberate choice of concavo-convex energy functional. As a result, we compare VM method favorably against the Fibonacci search technique with respect to computations (see Section 4.4 for comparisons).

The outline of this paper is as follows. In Section 2 our proposed threshold algorithm is described. In Section 3 numerical implementation is furnished in details. Experimental results are presented in Section 4. Section 5 contains directions for future work and conclusions.

## 2. Proposed method

To describe the mathematical principle behind the proposed method, let $I(x, y)$ and $T(x, y)$, respectively denote the image and the threshold functions. Our proposed energy functional consists of two terms: a data term and a regularization term as follows:

$$E(T;\alpha) = \sqrt{1-\alpha^2}E_1(T) + \alpha E_2(T), \tag{1}$$