

Received November 27, 2018, accepted December 9, 2018, date of publication December 20, 2018,
date of current version January 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2889013

A Robust Parameter-Free Thresholding Method for Image Segmentation

XINHUA CAO¹, (Senior Member, IEEE), TAIHAO LI², HONGLI LI³, SHUNREN XIA⁴, (Senior Member, IEEE), FUJI REN⁵, (Senior Member, IEEE), YE SUN⁶, AND XIAOYIN XU^{1D7}, (Senior Member, IEEE)

¹Department of Radiology, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA

²College of Medical Instruments, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China

³Department of Histology and Embryology, Third Military Medical University, Chongqing 400038, China

⁴Key Laboratory of Biomedical Engineering of Ministry of Education, Zhejiang University, Hangzhou 310027, China

⁵Department of Information Science and Intelligent Systems, Tokushima University, Tokushima 770-8501, Japan

⁶Department of Ophthalmology, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA

⁷Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

Corresponding author: Xiaoyin Xu (xxu@bwh.harvard.edu)

The work of X. Cao was supported by NIH under Award R01LM012434. The work of T. Li was supported by the Beijing Advanced Innovation Center for Imaging Technology under Grant BAICIT-2016012. The work of H. Li was supported in part by the National Natural Science Foundation of China under Grant 31271467 and Grant 31471148. The work of S. Xia was supported in part by the National Key Research and Development Program of China under Grant 2016YFC1306600. The work of F. Ren was supported in part by JSPS KAKENHI under Grant 15H01712. The work of Y. Sun was supported by the Boston Children's Hospital's Office of Faculty Development, Basic Translational Executive Committee, Clinical and Translational Research Executive Committee Faculty Career Development Grant. The work of X. Xu was supported by NIH under Award R01LM011415 and Award R01LM012434.

ABSTRACT In this paper, we presented a new parameter-free thresholding method for image segmentation. In separating an image into two classes, the method employs an objective function that not only maximizes the between-class variance but also the distance between the mean of each class and the global mean of the image. The design of the objective function aims to circumvent the challenge that many existing techniques encounter when the underlying two classes have very different sizes or variances. The advantages of the new method are twofold. First, it is parameter-free, meaning that it can generate consistent results. Second, the new method has a simple form that makes it easy to adapt to different applications. We tested and compared the new method with the standard Otsu method, the maximum entropy method, and the 2D Otsu method on the simulated and real biomedical and photographic images and found that the new method can achieve a more accurate and robust performance.

INDEX TERMS Segmentation, parameter-free thresholding, objective function, histogram.

I. INTRODUCTION

Segmentation is a widely encountered task in many aspects of image processing in industrial applications, including computer vision, remote sensing, and automatic object recognition [1]–[3]. Segmentation algorithms can be categorized based on whether they apply a global threshold or localized thresholds in the segmentation process. As many factors affect the performance of segmentation, there is no single technique that suits all types of images, and often one must select a method based on an appropriate set of criteria. For example, Rother *et al.* [4] proposed a graph cut method to iteratively extract an object of interest from an image after a user loosely drew a region around the object. Shi and Malik [5] presented a graph partition technique that considers both the total dissimilarity between the different

groups as well as the total similarity within the groups to delineate objects from the background. Comaniciu and Meer in 2002 presented a mean shift approach to filter an image while preserving its discontinuities, which effectively segments the image into homogeneous regions [6]. Other techniques include weighted variational models [7], random walk [8], minimal surface segmentation [9], active contour model [10], and methods that use convexity as a priori to regularize binary segmentation [11]. One important concern in practice is that a good technique should generate a robust and consistent performance in the sense that different users can apply the technique to obtain the same results.

Among the segmentation algorithms, histogram-based analysis represents a large body of segmentation techniques as many methods have been developed to estimate and extract

the probability density functions of the underlying classes in images. A convex hull thresholding has been developed to find concavities in a histogram as the potential selection of the threshold [12]. A peak and valley detection method has been proposed to convolve a histogram with a kernel to locate the peaks in the histogram [13]. To resolve normal probability distributions from the histogram of images, a series of Gaussians have been used to convolve with the histogram to find the location and direction of zero-crossing in the second derivative [14]. In finding an appropriate threshold from the histogram, entropy has been proposed as a criterion for selecting the appropriate threshold. An algorithm based on computing the entropy of each class to maximize the information between two segmented classes was also developed [15]. A method based on integrating fuzzy set assignment and entropy has been developed to partition a 2D histogram into fuzzy and non-fuzzy regions at first and then compute a threshold that results in the maximization of fuzzy entropy [16]. Histogram-based thresholding is not limited to gray-scale images as there are methods that have been developed for segmenting color images. A hierarchical method for thresholding has been developed for detecting homogeneity regions in a color image and calculating the histogram of the homogeneous regions to identify their peaks and, therefore, the valleys formed between the peaks [17]. In color image segmentation, a fusion procedure based on K-means clustering has been developed to combine segmentation maps in each color channel to obtain a final result [18].

Among the histogram-based thresholding techniques, the Otsu method is a widely used approach [19] because it is fast and optimal in some statistical senses. The Otsu method performs an exhaustive search for a global threshold that maximizes the between-class variance. Equivalently, the threshold also minimizes the intra-class variance. Over the years, many modifications and revisions have been made to the standard Otsu method [20], [21]. In many applications, Otsu method is the first step to segment an image into a foreground and a background for further analysis, thus, its results influence the overall performance of image processing. For instance, Zortea *et al.* [22] proposed a method that detects pigmented skin lesions in macroscopic photographs, in which Otsu method was used to combine independent threshold estimates computed from histograms of different parts of an image, and then post-processing steps were applied to identify the objects of interest. As reviewed by [23], Otsu method is also widely used as the pre-processing step in computer vision and image understanding to segment an image into two classes. Then connected-component analysis is applied to assign all pixels of each connected component to an object in the image. Efficient implementation of the Otsu method has been developed to speed up its computations [24]. For example, Yuan *et al.* [25] proposed a weighted object variance-based Otsu method in which a weight parameter was introduced to the Otsu function to find the valley of the histogram. The weight can then be adjusted to achieve a high detection rate and low false alarm rate in image-based defect

detection. In document image segmentation, Moghaddam and Cheriet proposed an adaptive Otsu method to adaptively estimate the background in an image and extract the characters in the foreground [26].

Also, as a way to improve the performance of the standard Otsu method, two-dimension methods have been proposed to consider both gray levels and local averages of an image [27], [28]. Multilevel thresholding based on the Otsu method has been developed, aiming to segment an image into more than two classes of objects [29]. Correspondingly, fast implementation of multilevel Otsu thresholding has also been presented [30]. A recursive algorithm that divides the histogram of an image into sub-ranges and recursively computes thresholds for fast computations has also been developed [31]. Sha *et al.* [32] presented a 2D Otsu method that searches for the optimal thresholds on two dimensions separately of the 2D histogram of an image. A multi-scale 3D Otsu method has been developed by Feng *et al.* [33] to segment brain images obtained by magnetic resonance imaging.

The above approaches do not effectively address two common challenges encountered in image segmentation. The first challenge is that, for an image with a histogram that is broad or has a flat valley, the segmentation results of many algorithms, including the Otsu method and its modifications, are often sub-optimal. The reason is that these methods assume a bi-modal histogram that has two distinct peaks, and when this condition is not met, the segmentation is often inaccurate. In fact, the pre-condition for a bi-modal histogram is often not met in reality when the two classes of an image have very different sizes. For example, if one class has a very large size and the other class is of a small size, the histogram may have only one prominent peak. The second challenge is that many of the above methods and other existing methods require a careful selection of parameters to perform well as incorrect choices will result in over- or under-segmentation. For example, in 2D Otsu method, a user needs to select the filter size and type to apply to an image to generate the second histogram. To overcome the above challenges, we propose a new method that can successfully segment an image even when its histogram has only one obvious peak, i.e., the two classes of the image are of very different sizes. On the one hand, the new method can effectively correct the imbalance between the classes of different sizes in determining a segmentation threshold. On the other hand, the new method retains the advantage of the Otsu method, i.e., it is parameter-free and, thus, avoids user-based bias in implementation. In our testing, we found that the new method can select a threshold that is more accurate than the standard Otsu method, maximum entropy method, and 2D Otsu method, and the technique is more robust in the presence of various sizes of the classes and the signal-to-noise ratio (SNR). Testing also showed that the new method does not incur significant extra computational time as compared with the standard Otsu method when both methods were implemented in a vectorized manner for efficiency.

The paper is organized as follows. Section II reviews the Otsu method and discusses the cases in which it performs sub-optimally. Section III presents the new method and uses examples to demonstrate and compare its performance. Results of the new method and comparison with existing algorithms are given in Section IV. Section V presents the discussion and conclusions.

II. OTSU METHOD

Assume an image I of size $M \times N$ is represented in L gray levels $[0, 1, \dots, L-1]$, the Otsu method finds a threshold $T \in [0, L-1]$ to segment I into two classes C_0 and C_1 . We denote the mean and variance of C_0 as μ_0 and σ_0^2 , respectively. Similarly, we define μ_1 and σ_1^2 for C_1 . Let p_0 and p_1 be the probability of classes C_0 and C_1 , respectively. Then the Otsu method minimizes the intra-class variance by searching through the histogram of I to find threshold T by

$$\operatorname{argmin}_T p_0(T)\sigma_0^2(T) + p_1(T)\sigma_1^2(T). \quad (1)$$

When T is found, the image is segmented into two classes, C_0 and C_1 ; and their sizes are $p_0 MN$ and $p_1 MN$, respectively. Mathematically, Eq. (1) is equivalent to maximizing the between-class variance

$$\operatorname{argmax}_T \sigma_w^2(T) = p_0(T)p_1(T)(\mu_0(T) - \mu_1(T))^2. \quad (2)$$

For the purpose of our ensuing discussion, we can define

$$A(T) \equiv p_0(T)p_1(T) \quad (3)$$

$$B(T) \equiv (\mu_0(T) - \mu_1(T))^2 \quad (4)$$

then we can rewrite Eq. (2) as

$$\operatorname{argmax}_T \sigma_w^2(T) = A(T)B(T). \quad (5)$$

Here we note that, although term B cannot be determined until image I is given and threshold T is found, we can have an insight into the behavior of term A . Indeed, in term A , when probability $p_0(T)$ changes from 0 to 1.0 as a function of T , $p_1(T)$ changes from 1.0 to 0 in a reverted manner as $p_1(T) = 1 - p_0(T)$. We then have

$$A(T) = p_0(T) - p_0^2(T) \quad (6)$$

where $p_0 \in [0, 1]$. From Eq. (6) we can observe that term A will attain its maximum value of 0.25 when $p_0 = 0.5$. In Figure 1 we plot term A by letting p_0 changes from 0 to 1 at a step size of 0.1. In other words, if we, for the moment, ignore the effect of B on $\sigma_w^2(T)$ and only focus on A , then we expect that the process of finding $\operatorname{argmax} \sigma_w^2(T)$ will tend to stop when $p_0 = 0.5$, meaning that the process will tend to segment the image into two classes of equal size. In other words, the Otsu method tends to penalize the class of smaller size by assigning more pixels to it. In reality, because $\sigma_w^2(T)$ is a function of both A and B , its behavior will not be as simplistic as we assumed; however, as we will next demonstrate, the sizes of each class have an important effect in determining the threshold T . This effect becomes more

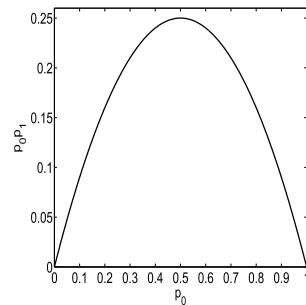


FIGURE 1. Plot of A as p_0 increases from 0 to 1 at a step size of 0.1. Here we note that in practice, p_0 may not increase at a uniform step size, but this plot is still valid for showing how A changes as a function of p_0 .

profound when there is a big difference between the sizes of the two classes.

To illustrate the effect of class sizes on finding threshold T by the Otsu method, we simulated an image of 100×100 pixels. The noise-free image consists of two classes, a bright C_0 with a uniform pixel value of 150 and a dark C_1 with a uniform value of 100, Figure 2(a). In this example, C_0 accounts for 10% of all the pixels and C_1 accounts for the remaining 90% of the pixels. In other words, the ratio of the sizes of the two classes was 1000 to 9000. We then added Gaussian noise $\mathcal{N}(0, \sigma_n^2)$ with a standard deviation of 15 to the image, Figure 2(b). Mathematically, we have

$$X = I + W \quad (7)$$

where X , I , and W are the observation, the noise-free image, and noise, respectively. We define SNR as

$$\text{SNR} = 10 \log_{10} \frac{S^2}{\sigma_n^2} (\text{dB}) \quad (8)$$

where S is the intensity difference between C_0 and C_1 . The SNR of the above simulation is 10 dB. The histograms of the clean and noisy observations are shown in Figures 2(c) and (d), respectively. Though the histograms of C_0 and C_1 were well separated in Figure 2(c), their histograms became mixed in Figure 2(d) but we can observe the two peaks. Because in this simulation we knew *a priori* the pixel values of the two classes and the statistics of the noise, we can determine that the optimal threshold to segment Figure 2(b) should be 125, the average between the mean values of C_0 and C_1 . However, when we applied the Otsu method to the image, the threshold was found to be 114. Its segmentation result is shown in Figure 2(e). The result of using optimal threshold 125 for segmentation is shown in Figure 2(f). To quantitatively compare the results of Figures 2(e) and (f), we counted the number of pixels that were mis-segmented in each class. In Figure 2(e) there were a total of 1566 pixels incorrectly segmented in it, with 13 and 1553 pixels incorrectly segmented in C_0 and C_1 , respectively. In comparison, there were only 408 pixels incorrectly segmented in Figure 2(f), with 79 and 329 pixels incorrectly segmented in C_0 and C_1 , respectively. In another measurement,

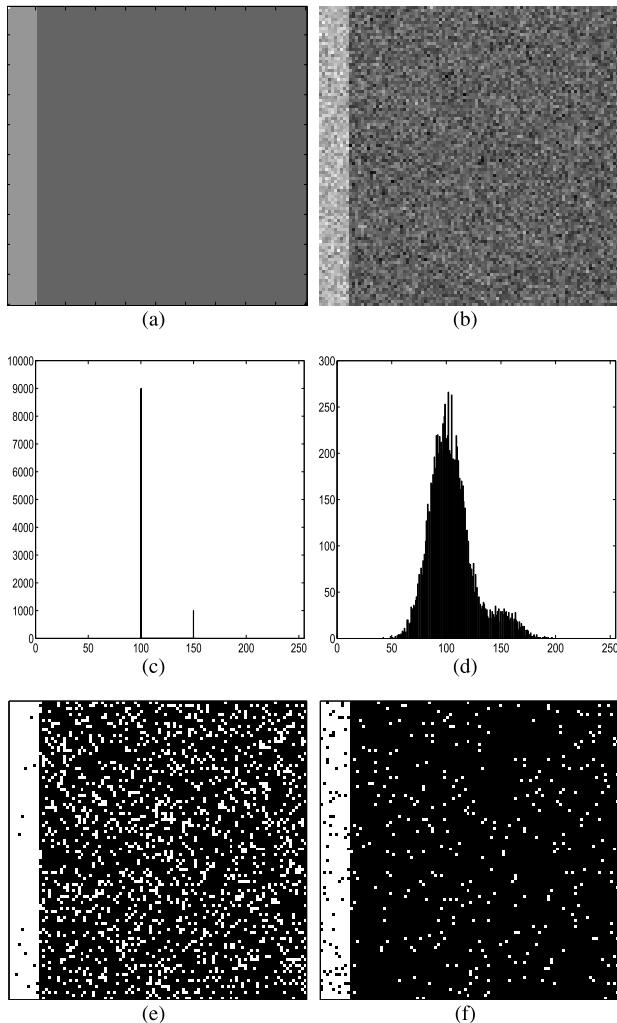


FIGURE 2. A synthetic example. (a) A clean image of 100×100 pixels with two classes. Here C_0 (to the left of the image), accounting for 10% of all the pixels, is bright with an intensity of 150 and C_1 is dark with an intensity of 100, accounting for 90% of the whole image. (b) Gaussian noise with a zero mean and standard deviation of 15 was added to (a). (c) Histogram of (a). (d) Histogram of the noisy observation (b). (e) Segmentation result by the Otsu method whose threshold was 114. We note that there were very few mis-segmented pixels in C_0 but many mis-segmented pixels in C_1 . Total number of mis-segmented pixel was 1566. (f) The result of segmenting (b) by threshold of 125. This result had more mis-segmented pixels in C_0 but many fewer mis-segmented pixels in C_1 . Total number of mis-segmented pixel was 408.

we found that in Figure 2(e) the ratio of the sizes of C_0 to C_1 was 2540 to 7560, while in Figure 2(f) this ratio was 1250 to 8750, which is much closer to the correct ratio of 1000 to 9000 as the two classes were simulated.

To further elucidate the effect of class sizes on determining global threshold, we next simulated an image similar to Figure 2 but changed the ratio of the sizes of C_0 to C_1 to 2000 to 8000, i.e., increasing the number of pixels in C_0 to 2000 and reducing the number of pixels in C_1 to 8000. Figures 3(a) and (b) show the clean image and the noisy observation, with the standard deviation of the noise set at 15. In this example, the threshold given by the Otsu method was 121, as compared to the optimal threshold of 125. The

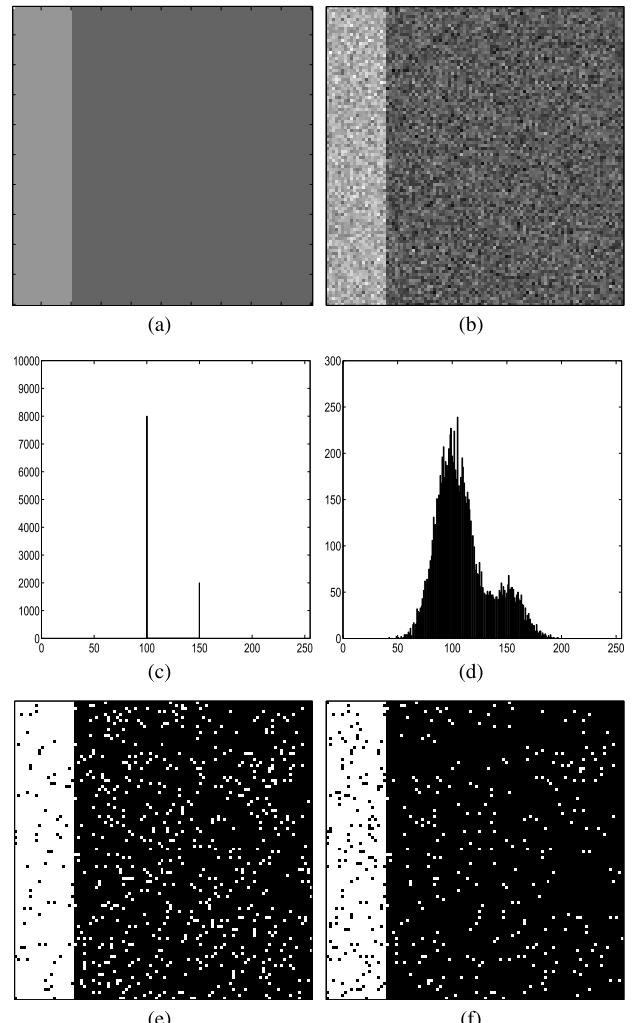


FIGURE 3. A synthetic example where the ratio of the sizes of C_0 to C_1 was 2000 to 8000. (a) The clean image. (b) The noisy observation. (c) Histogram of (a). (d) Histogram of noisy observation (b). (e) Segmentation result by the Otsu method whose threshold was 121. (f) The result of segmenting (b) by the optimal threshold of 125.

histograms of the clean and noisy observations are shown in Figures 3(c) and (d), respectively. We can more easily observe the two peaks in histogram in Figure 3(d) than in the example above, indicating that the Otsu method can be expected to obtain a more accurate threshold. The segmentation results are shown in Figures 3(e) and (f), respectively. We note, based on this example, that as the difference between the sizes of C_0 and C_1 became smaller, the threshold given by the Otsu method improves, though it still does not reach the optimal value of 125. In Figure 3(e) the total number of mis-segmented pixels was 686, with 63 pixels mis-segmented in C_0 and 623 pixels mis-segmented in C_1 . In Figure 3(f), the total number of mis-segmented pixels was 421, with 145 pixels mis-segmented in C_0 and 276 pixels mis-segmented in C_1 . Similarly, we compared the ratio of the sizes of the segmented C_0 to C_1 and found that in Figure 3(e) the ratio was 2560 to 7440 and the ratio in Figure 3(f) was 2131 to 7869, which is closer to the correct ratio of 2000 to 8000.

III. NEW OBJECTIVE FUNCTION

To achieve a higher degree of accuracy and robustness in segmentation, we designed a new objective function

$$\operatorname{argmax}_T \sigma_w^2(T) = p_0(T)p_1(T)[(\mu_0(T) - \mu_1(T))^2 + (\mu_0(T) - \mu)^2 + (\mu_1(T) - \mu)^2] \quad (9)$$

where μ is the mean value of the whole image X and can be expressed as

$$\mu = p_0(T)\mu_0(T) + p_1(T)\mu_1(T). \quad (10)$$

Compared with the standard Otsu method, the new objective function has two new terms measuring the distance between the mean value of the whole image and the mean value of each segmented class. As such the new objective function rewards segmentation results whose classes have mean values that are at some distance from μ . The motivation of introducing the two terms is to correct the tendency of Otsu method to penalize the class of smaller size. The two new terms, measuring the distances from the means of each segmented class to the global mean, have the effect of balancing the weights of each class in searching for the segmentation threshold. In other words, the two new terms lead the objective function to find a threshold such that the two resulting classes are not only well separated from each other but also well separated from the global mean that may be predominantly determined by the large class, hence, reducing the size effect of the large class on finding the threshold. As the first example, we applied the new method to segment the image in Figure 2(b). The threshold was found to be 123; though this is not the same as the optimal threshold of 125, it is much more accurate than the threshold of 114 given by the standard Otsu method. The segmentation result is shown in Figure 4(a). In this result the total number of mis-segmented pixels was 607; as compared to Figure 2(e) which had 1566 mis-segmented pixels. In Figure 4 there were 49 and 558 pixels incorrectly segmented in C_0 and C_1 , respectively. The ratio of the sizes of C_0 to C_1 in Figure 4 was 1509 to 8491, as compared to the true ratio of 1000 to 9000. Here we note that, although the two new terms appear very simple in our method, their effect on the segmentation results is profound as can be seen in the Results section. On the one hand, the two new terms overcome a well-encountered drawback of Otsu method, i.e., its bias toward one class over the other when the two underlying classes have different sizes. On the other hand, our new method remains parameter-free as the standard Otsu method, thus making it a highly objective technique in practice such that its performance does not depend on user intervention.

IV. RESULTS

For a more comprehensive evaluation of the new method in segmenting images, we compared its performance with those of Otsu's method and the entropy-based method by Kapur *et al.* [15]. The entropy-based method seeks to maximize the information between the foreground and the

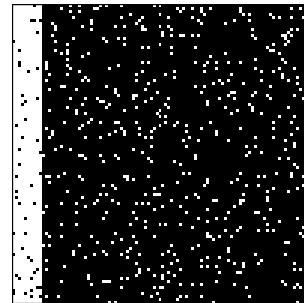


FIGURE 4. The segmentation result obtained by applying the new method to Figure 2(b). Total number of mis-segmented pixel was 607, compared to the 1566 mis-segmented pixels by the Otsu method as shown in Figure 2(e).

background such as

$$\operatorname{argmax}_T H(A; T) + H(B; T) \quad (11)$$

where $H(A)$ and $H(B)$ are the entropy of the foreground and background, respectively, when the image is segmented by threshold T . Kapur's method searches through the histogram of an image to find T that gives the largest total entropies of the two classes. The reason for choosing the standard Otsu and the entropy-based methods is that these two methods are parameter-free, i.e., a user does not need to select a parameter in applying the methods, thus they have the advantage of generating consistent results in practice. The dataset we used for evaluating the performance of the new method include ten photographs of still objects, people, natural scene, animals, 10 microscopic images of cells and neurons, 80 microscopic images of retinal blood vessels, 30 microscopic images of muscle fibers, and 170 finger prints. In total there were 300 images in the dataset. As such, the test images have different amount of details and structures to provide a comprehensive basis for evaluation and comparison. In comparison, we simulated different SNRs and sizes of C_0 and C_1 while setting the mean values of C_0 and C_1 at 150 and 100, respectively. For each case of simulation we repeat the tests 50 times to derive the mean and standard deviation of the results. In the simulation we incremented the percentage of C_0 in the whole image from 10% to 90%; correspondingly, the percentage of C_1 decreased from 90% to 10%. We evaluated how the new method obtained the threshold and compared it with the Otsu method and the method of maximizing the total entropy by Kapur *et al.* [15]. In Figure 5 we compared the thresholds obtained by three methods. It can be observed that in each case the new method outperformed the Otsu method in obtaining a threshold closer to 125, the optimal value. From the figure we can make four observations. First, at relatively high SNRs of 10 and 8 dB, the new method obtained thresholds closer to the optimal threshold than the Otsu method did. The new method and the maximum entropy method obtained thresholds that were of similar distance from the optimal threshold, though their thresholds tend to locate on the opposite side of the optimal threshold. Second, at lower SNRs of 6 and 3 dB, the new

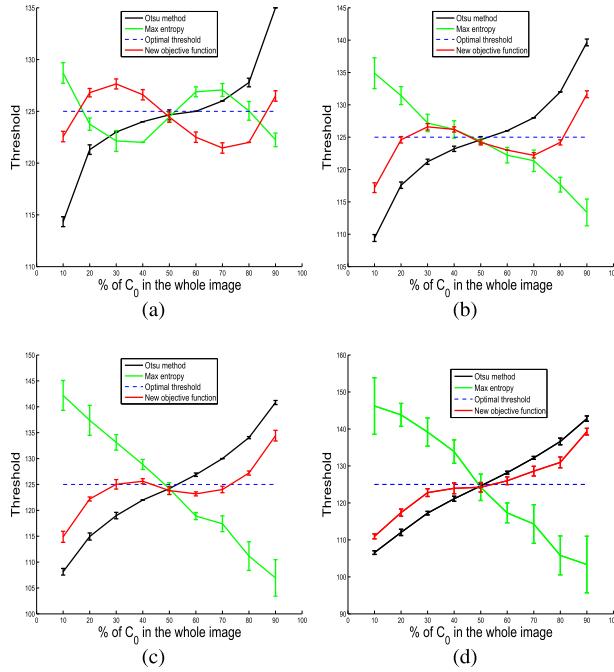


FIGURE 5. We simulated different ratios of the sizes of C_0 to C_1 by increasing the percentage of C_0 in the whole image from 10% to 90% and decreasing the percentage of C_1 from 90% to 10%, correspondingly. (a-d) The threshold given by the Otsu method, the maximum entropy method, and the new method at an SNR of 10, 8, 6, and 3 dB, respectively.

method performed better than both the Otsu method and the maximum entropy method. As the SNR decreased to 3 dB, Figure 5(d) showed that the new method was still able to obtain a threshold within a narrow range of 125 when C_0 accounted for from 20% to 80% of the whole image. Third, at all SNRs, the new method and the Otsu method had more consistent results relative to the maximum entropy method. Fourth, interestingly, the Otsu method and the maximum entropy method tended to obtain thresholds at the opposite sides of the optimal threshold. For example, at SNR of 6 and 3 dB and when C_0 accounted for 10% to 40% of all the pixels, while the Otsu method obtained thresholds that favored C_0 , i.e., assigned more pixels to C_0 than the true number of pixels in C_0 , the maximum entropy method obtained thresholds that favored C_1 . When C_0 accounted for 60% to 90% of all the pixels, the Otsu method obtained thresholds favoring C_1 while the maximum entropy method's results favored C_0 . Next, we evaluated how this improvement in thresholding translates into more accurate segmentation results in terms of the total number of mis-segmented pixels and the mean rate of error (MRE). The total number of mis-segmented pixels N_{total} is defined as

$$N_{total} = N_{fg} + N_{bg} \quad (12)$$

where N_{fg} and N_{bg} are the numbers of mis-segmented pixels in the foreground and background, respectively. The MRE is defined as

$$MRE(\%) = \frac{1}{2} \left[\frac{N_{fg}}{S_{fg}} + \frac{N_{bg}}{S_{bg}} \right] (\%) \quad (13)$$

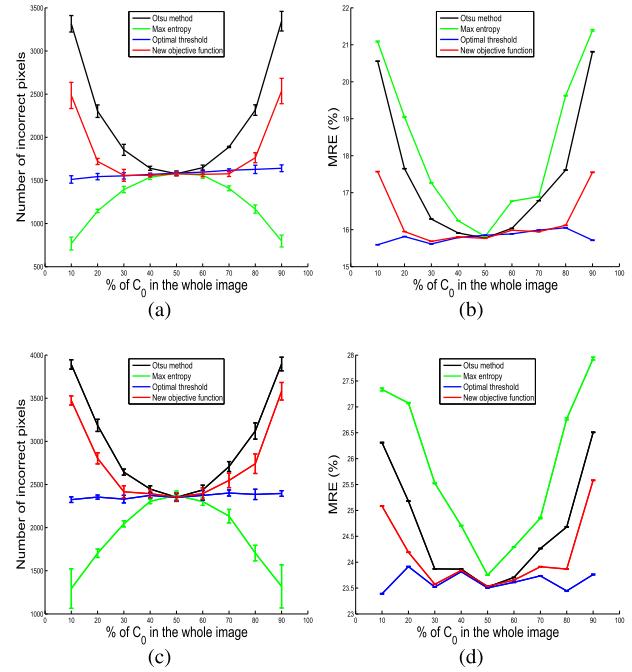


FIGURE 6. (a, c) show N_{total} of each method of C_0 in the whole image increased from 10% to 90% for SNR of 6 and 3 dB, respectively. (b, d) show the corresponding MRE of (a, c).

where S_{fg} and S_{bg} are the sizes of the foreground and background, respectively. As we will see next, together, N_{total} and MRE can provide a balanced evaluation of the performance of the segmentation methods. For the different SNRs of Figure 5, the comparison is shown in Figure 6. We included the segmentation results given by the optimal threshold of 125 in the figure.

At an SNR of 6 dB, the new method obtained clearly better results than the Otsu method and the maximum entropy method, Figures 6(a) and (b). The new method obtained results that were similar to those of the optimal threshold over a large range, from C_0 accounting for 20% of the whole image to 80%.

As the SNR decreased to 3 dB, Figures 6(c) and (d) show that the new method still outperformed the Otsu method and the maximum entropy method. We note that when C_0 accounted for from 20% to 80% of the whole image, the MRE of the new method was not much different from the MRE given by the optimal threshold, Figure 6(d). However, in these cases, the MREs obtained by the Otsu method and, in particular, the maximum entropy method were much higher than that given by the optimal threshold.

By examining Figures 5 and 6 together, we can also observe that the Otsu method and the maximum entropy method are sensitive to the difference in sizes of C_0 and C_1 . First, when C_0 and C_1 each accounts for 50% of all the pixels, both the Otsu method and the maximum entropy method were able to obtain the threshold of 125 as the SNR decreased from 10 dB to 3 dB and the results of the two methods had approximately the same N_{total} and MRE that were very

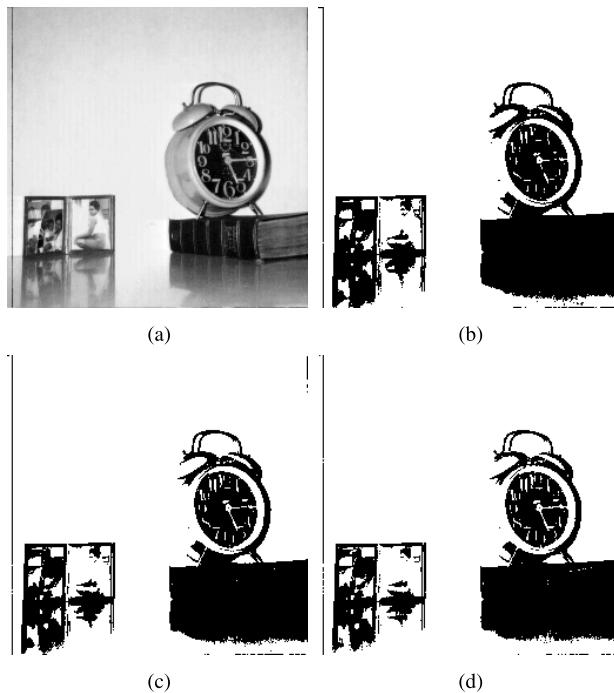


FIGURE 7. (a) Original picture of test image “clock”. (b) Result given by Otsu method with a threshold of 157. (c) Result given by the maximum entropy method with a threshold of 151. (d) Result given by the new method with a threshold of 146.

close to those given by the optimal threshold of 125, meaning that both methods are relatively robust to the changes in SNR when C_0 and C_1 have the same sizes. Second, for a fixed SNR, we can observe that when there is a difference in sizes of C_0 and C_1 , both the Otsu method and the maximum entropy method lose their robustness quickly. For example, at an SNR of 6 dB, when C_0 accounted for 40% of all the pixels (Figure 5(c)), the mean thresholds given by the Otsu method and the maximum entropy method were 122 and 128, compared to the optimal threshold of 125. When C_0 accounted for 30% of all the pixels, the mean thresholds given by the Otsu method and the maximum entropy method were 118 and 132 (Figure 5(c)), respectively, demonstrating the high sensitivity of the Otsu method and the maximum entropy method to the size difference of the two underlying classes. In comparison, as can be observed in Figures 5, the threshold given by the new method was more robust to the differences in sizes of C_0 and C_1 .

We applied the new method to real images, including photographs, microscopic images, and finger prints, to evaluate its performance. For each test, we compared the segmentation performance of the Otsu method, the maximum entropy method, and our new method.

For the first type of tests, we compared the three methods on photographs. As the first example, we applied the three methods on the image “clock”. The original image and the results are shown in Figure 7. Careful comparison of the results shows that the new method created the best segmentation in terms of preserving fine image details. For example,

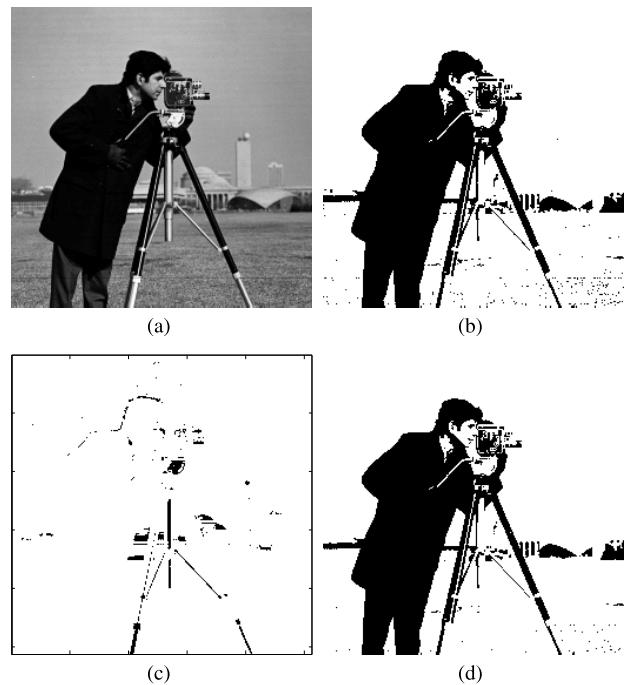


FIGURE 8. (a) Original picture of test image “cameraman”. (b) Result given by Otsu method with a threshold of 88. (c) Result given by the maximum entropy method with a threshold of 192. (d) Result given by the new method with a threshold of 80.

the arms and the numbers on the clock were better preserved in Figure 7(d) than in Figure 7(b) and (c). As the second example shown in Figure 8, we tested the three methods on the “cameraman” picture. We note that, although both the Otsu method and the proposed method segmented the main character, the new method avoided segmenting many isolated pixels at the bottom of the picture. The maximum entropy method, however, did not create a satisfactory threshold in this case.

As the third example of photographic tests, we compared the three methods on segmenting Figure 9(a). From the results given by the three methods, we can observe that the Otsu method created a result containing small patches of over-segmentation near the top of the image, and the maximum entropy method severely over-segmented the image, while the new method correctly segmented the image and had the smallest amount of over-segmentation.

For the second type of tests, we compared the three methods on microscopic images. As the first example, we tested the new method on some microscopic images of cultured cells. Figure 10(a) shows a microscopic image of two mature oligodendrocytes, which are star-shaped cells of the nervous system, including their soma and web-like protrusion processes but also some background interference seen at the top and bottom of the image. Its segmentation results given by the Otsu method, the maximum entropy method, and the new method are shown in Figures 10(b-d), respectively. Compared with the result of the Otsu method (Figure 10(b)), the new method generated a result that largely eliminated the

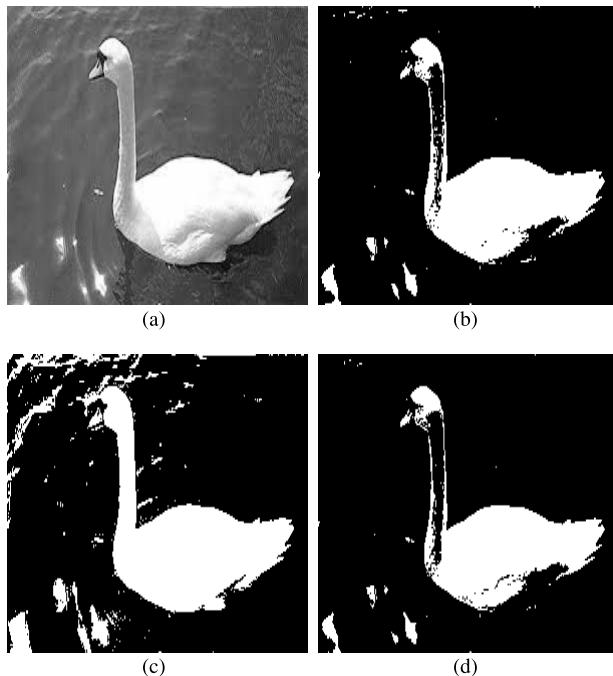


FIGURE 9. (a) Original picture. (b-d) Results given by the Otsu method, the maximum entropy method, and the new method, respectively.

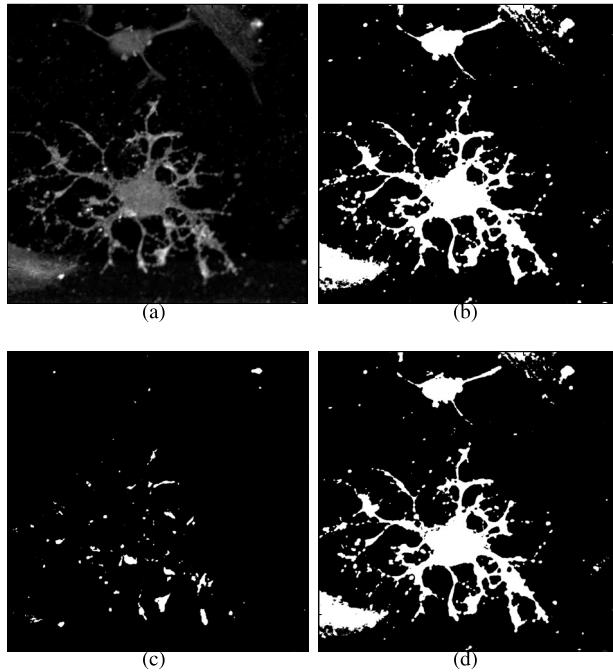


FIGURE 10. (a) Original picture of a microscopic image of two mature oligodendrocytes. (b) Result given by the Otsu method with a threshold of 48. (c) Result given by the maximum entropy method with a threshold of 131. (d) Result given by the new method with a threshold of 55.

background interference at the top of the image. The maximum entropy method did not generate a satisfactory result as the majority of the cellular structure was not well segmented. As the second example, we applied the Otsu method

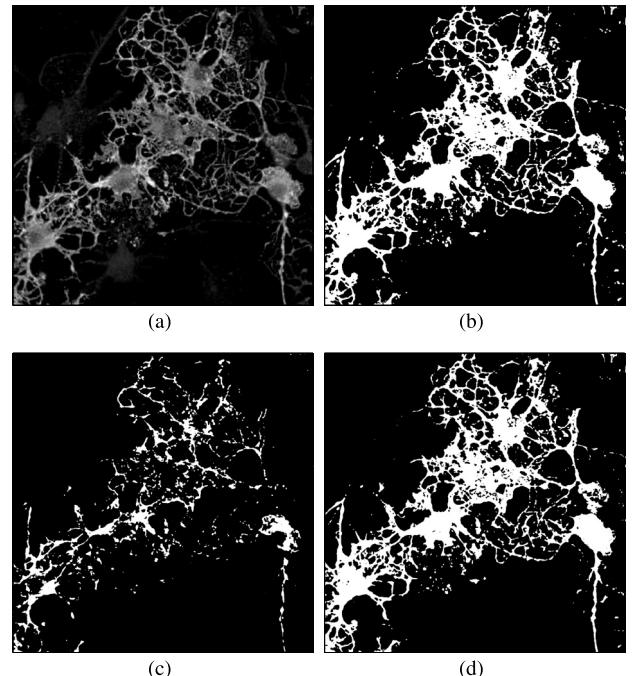


FIGURE 11. (a) Original picture of a microscopic image of multiple oligodendrocytes. (b) Result given by the Otsu method with a threshold of 57. (c) Result given by the maximum entropy method with a threshold of 120. (d) Result given by the new method with a threshold of 66.

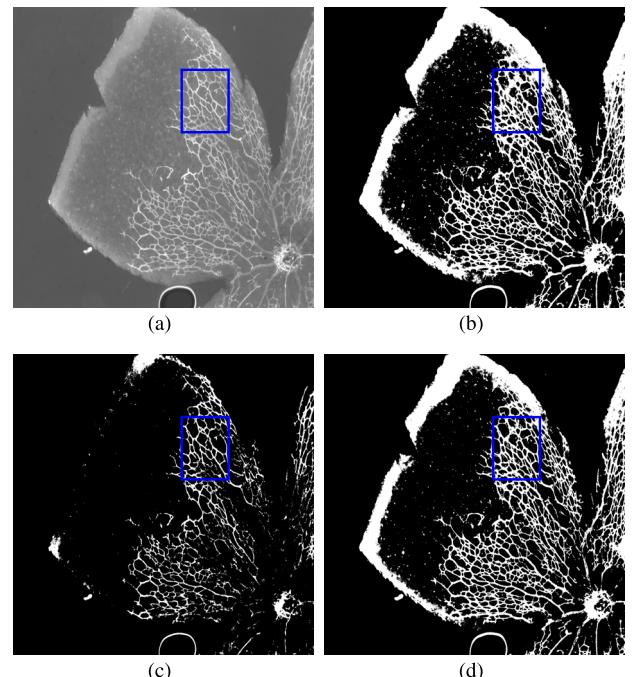
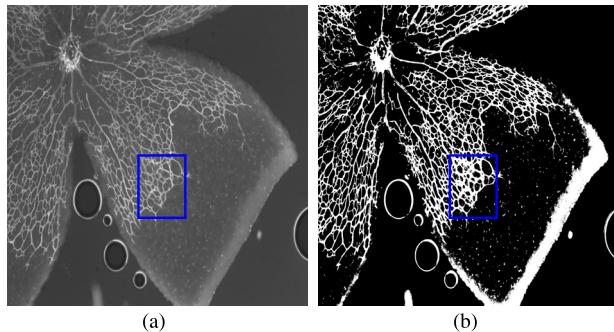


FIGURE 12. (a) Original picture of a microscopic image of mouse retinal vasculature. (b-d) Results given by the Otsu method, the maximum entropy method, and the new method, respectively.

and the new method to Figure 11(a), which has multiple oligodendrocytes. The segmentation results given by the three methods are shown in Figures 11(b-d). From the comparison



(a)

(b)

(c)

(d)

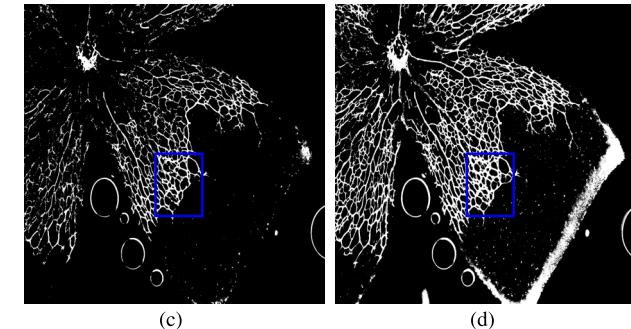
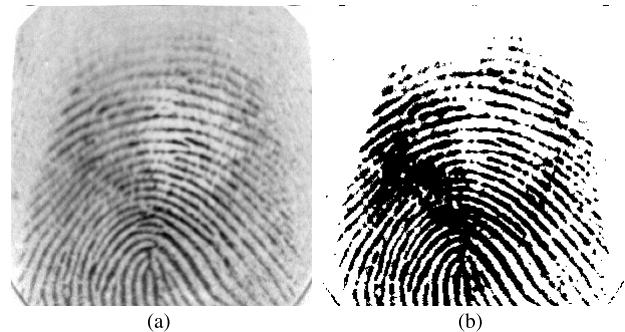


FIGURE 13. (a) Original picture of a microscopic image of mouse retinal vasculature. (b-d) Results given by the Otsu method, the maximum entropy method, and the new method, respectively.

results, we observed that the new method provided a cleaner and more intact delineation of the cells, particularly around their protrusion processes, relative to the other two methods. As the third example, we tested the three methods on a different microscopic image type. The segmentation results are shown in Figure 12. Figure 12(a) is the original image. Figures 12(b-d) are the results of the Otsu method, the maximum entropy method, and the new method, respectively. We can note that the maximum entropy method created an under-segmented result, as compared to the other two methods. Between the Otsu method and the new method we note that the Otsu method over-segmented the images, as can be seen in the area marked by the blue rectangle. As the fourth example, we tested the methods on another microscopic image, Figure 13. Similarly we observed that, on the one hand, the new method can delineate fine details better than the Otsu method, as emphasized by the blue rectangles in the images, and, on the other hand, the new method can preserve the whole structure better than the maximum entropy method.

Lastly, we compared the three methods on segmenting finger prints. The first example is shown in Figure 14. From the comparison we observed that the new method, Figure 14(d), segmented the finger print without the loss of the valid structure as seen in the result of maximum entropy, Figure 14(c). Also, the result of the new method did not generate a large over-segmented area as seen in the upper left side of the Otsu method, Figure 14(b). Another finger print segmentation example is shown in Figure 15. Comparing the results given by the three methods, we again observed that the new method had the best performance.

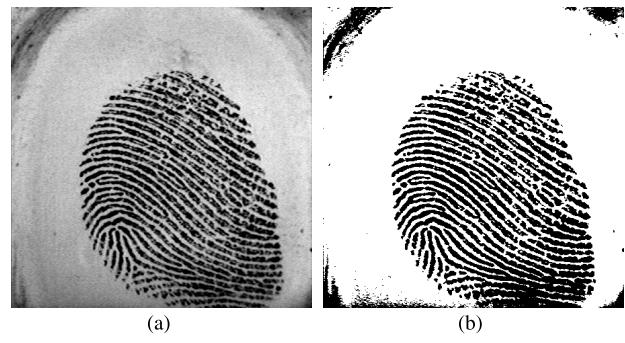


(a)

(b)



FIGURE 14. (a) Original picture of a finger print. (b-d) Results given by the Otsu method, the maximum entropy method, and the new method, respectively.



(a)

(b)



(c)

(d)

FIGURE 15. (a) Original picture of a finger print. (b-d) Results given by the Otsu method, the maximum entropy method, and the new method with thresholds of 123, 147, and 109, respectively.

More comparisons are shown in the supplementary materials, demonstrating the superior performance of the new method. In supplementary materials, in addition to the standard Otsu method and the maximum entropy method, we

compared our method versus four different global thresholding algorithms, namely, the 2D Otsu method, the minimum error method, the maximum Renyi entropy method, and the Shanbhag method [34]–[36]. We also compared our method versus two local thresholding algorithms, the Bernsen method and the Sauvola method [37], [38]. The comparisons show that the new method can achieve better results with a stable performance. We evaluated the computational speed of the new method versus the standard Otsu method and found they have statistically equivalent computational times. In other words, the new method does not require more computational time. The results are shown in the supplementary materials.

V. DISCUSSION AND CONCLUSIONS

In practical applications, often an image has two classes that are of considerably different sizes. Many existing methods do not generate satisfactory results in such cases as these methods are often biased toward one class. In this work we proposed the design of a new algorithm to segment an image based on its histogram analysis. The new method can handle the cases when the two underlying classes of an image are of significantly different sizes, making it well suited for a large number of practical applications. The new method was shown to have a more robust and accurate performance than the Otsu method, which maximizes the total variation of the two segmented classes, the maximum entropy method, which maximizes the total entropy of the two segmented classes, and the 2D Otsu method. A known limitation of the Otsu method is its sensitivity to the class size, as discussed in [39]. The maximum entropy method has the similar drawback, as can be seen from Figures 5 and 6. The new method overcomes the above drawbacks, particularly when there is considerable differences in sizes of the two classes in an image. We note that there are other methods that can be used to segment images shown in this study, however, many existing methods require appropriate settings of parameters and, therefore, have the risk of introducing a user-based bias as each user may choose the parameters differently. The new terms in the objective function have the effect of improving the robustness of threshold selection when the two underlying classes have different sizes and there are no obvious two peaks in the histograms of images. As shown in our examples, even when the difference in sizes of the two classes is significant, e.g., one class accounts for 10% and the other accounts for 90% of the whole image, the new method can still achieve satisfactory results, making it widely applicable to many scenarios. We note that there are other methods that have been developed for image segmentation but such methods are typically not parameter-free, meaning that their performance depends on user interventions. One advantage of the proposed method is that it is parameter-free, meaning that its segmentation result is consistent across different users and there are no inter-user variations. This advantage is critical for practical applications because users can always obtain the same results and can set the method for automatically processing hundreds to thousands of images without the need for tuning any parameters

across the images. Another advantage of the new method is that it has a simple form that makes it amenable to further improvement. For example, the new function can be extended to two-dimensional histogram modeling and analysis like the 2D Otsu method.

One potential limitation of the proposed method is that its performance may be affected by the level of noise in images. It is conceivable that noise reduction may be performed as a pre-processing step in some cases. However, as noise reduction usually results in a smoothed version of the original image, fine details may be lost in the process. In terms of computational speed, our tests showed that the new method did not incur more time than the standard Otsu method. As shown in the supplementary materials, paired t-test of the computational times of the new method and the Otsu method indicated that there were no statistically significant differences between the two methods. In other words, the new method can achieve superior performance without incurring additional computational time. In this work, the two methods were implemented in vectorial manner as an exhaustive search scheme in MATLAB and applied to process images on a computer with CPU only. It is possible that an implementation in C/C++ or other compiled languages and utilization of GPU may further improve the computational speed of the new method.

ACKNOWLEDGMENT

T. Li was with Capital Normal University, Beijing 100048, China.

REFERENCES

- [1] P. F. Felzenszwalb and D. P. Huttenlocher, “Efficient graph-based image segmentation,” *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.
- [2] P. Mitra, B. U. Shankar, and S. K. Pal, “Segmentation of multispectral remote sensing images using active support vector machines,” *Pattern Recognit. Lett.*, vol. 25, no. 9, pp. 1067–1074, 2004.
- [3] Y. Liu, M. Yu, L. Cui, G.-Y. Jiang, Y.-G. Wang, and S.-L. Fan, “An automatic segmentation method for printed circuit board welding component under stereo optical microscope,” *Sensors Transducers*, vol. 196, no. 1, pp. 75–81, 2016.
- [4] C. Rother, V. Kolmogorov, and A. Blake, “grabcut: Interactive foreground extraction using iterated graph cuts,” *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004.
- [5] J. Shi and J. Malik, “Normalized cuts and image segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.
- [6] D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
- [7] C. Liu, M. K.-P. Ng, and T. Zeng, “Weighted variational model for selective image segmentation with application to medical images,” *Pattern Recognit.*, vol. 76, pp. 367–379, Apr. 2018.
- [8] C. G. Bampis, P. Maragos, and A. C. Bovik, “Graph-driven diffusion and random walk schemes for image segmentation,” *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 35–50, Jan. 2016.
- [9] D. Tenbrink and X. Jiang, “Image segmentation with arbitrary noise models by solving minimal surface problems,” *Pattern Recognit.*, vol. 48, no. 11, pp. 3293–3309, 2015.
- [10] S. Niu, Q. Chen, L. de Sisternes, Z. Ji, Z. Zhou, and D. L. Rubin, “Robust noise region-based active contour model via local similarity factor for image segmentation,” *Pattern Recognit.*, vol. 61, pp. 104–119, Jan. 2017.
- [11] L. Gorelick, O. Veksler, Y. Boykov, and C. Nieuwenhuis, “Convexity shape prior for binary segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 2, pp. 258–271, Feb. 2017.

- [12] A. Rosenfeld and P. De La Torre, "Histogram concavity analysis as an aid in threshold selection," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, no. 2, pp. 231–235, Mar./Apr. 1983.
- [13] M. I. Sezan, "A peak detection algorithm and its application to histogram-based image data reduction," *Comput. Vis., Graph., Image Process.*, vol. 49, no. 1, pp. 36–51, 1990.
- [14] M. J. Carlotto, "Histogram analysis using a scale-space approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 1, pp. 121–129, Jan. 1985.
- [15] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Comput. Vis., Graph., Image Process.*, vol. 29, no. 3, pp. 273–285, 1985.
- [16] H. D. Cheng, Y. H. Chen, and X. H. Jiang, "Thresholding using two-dimensional histogram and fuzzy entropy principle," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 732–735, Apr. 2000.
- [17] H.-D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2071–2082, Dec. 2000.
- [18] M. Mignotte, "Segmentation by fusion of histogram-based K -means clusters in different color spaces," *IEEE Trans. Image Process.*, vol. 17, no. 5, pp. 780–787, May 2008.
- [19] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
- [20] H. Cai, Z. Yang, X. Cao, W. Xia, and X. Xu, "A new iterative tri-class thresholding technique in image segmentation," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1038–1046, Mar. 2014.
- [21] W. A. Hussein, S. Sahran, and S. N. H. S. Abdullah, "A fast scheme for multilevel thresholding based on a modified bees algorithm," *Knowl.-Based Syst.*, vol. 101, pp. 114–134, Jun. 2016.
- [22] M. Zortea, E. Flores, and J. Scharcanski, "A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images," *Pattern Recognit.*, vol. 64, pp. 92–104, Apr. 2017.
- [23] L. He, X. Ren, Q. Gao, X. Zhao, B. Yao, and Y. Chao, "The connected-component labeling problem: A review of state-of-the-art algorithms," *Pattern Recognit.*, vol. 70, pp. 25–43, Oct. 2017.
- [24] L. J. Dong, G. Yu, P. Ogunbona, and W. Q. Li, "An efficient iterative algorithm for image thresholding," *Pattern Recognit. Lett.*, vol. 29, no. 9, pp. 1311–1316, 2008.
- [25] X.-C. Yuan, L.-S. Wu, and Q. Peng, "An improved Otsu method using the weighted object variance for defect detection," *Appl. Surf. Sci.*, vol. 349, pp. 472–484, Sep. 2015.
- [26] R. F. Moghaddam and M. Cheriet, "AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization," *Pattern Recognit.*, vol. 45, no. 6, pp. 2419–2431, 2012.
- [27] J. Liu and W. Li, "The automatic thresholding of gray-level pictures via two-dimensional Otsu method," *Acta Autom. Sinica*, vol. 19, no. 1, pp. 101–105, 1993.
- [28] J. Gong, L. Li, and W. Chen, "Fast recursive algorithms for two-dimensional thresholding," *Pattern Recognit.*, vol. 31, no. 3, pp. 295–300, 1998.
- [29] D.-Y. Huang and C.-H. Wang, "Optimal multi-level thresholding using a two-stage Otsu optimization approach," *Pattern Recognit. Lett.*, vol. 30, no. 3, pp. 275–284, 2009.
- [30] P.-S. Liao, T.-S. Chen, and P.-C. Chung, "A fast algorithm for multilevel thresholding," *J. Inf. Sci. Eng.*, vol. 17, no. 5, pp. 713–727, 2001.
- [31] S. Arora, J. Acharya, A. Verma, and P. K. Panigrahi, "Multilevel thresholding for image segmentation through a fast statistical recursive algorithm," *Pattern Recognit. Lett.*, vol. 29, no. 2, pp. 119–125, 2008.
- [32] C. Sha, J. Hou, and H. Cui, "A robust 2D Otsu's thresholding method in image segmentation," *J. Vis. Communun. Image Represent.*, vol. 41, no. 11, pp. 339–351, 2016.
- [33] Y. Feng, H. Zhao, X. Li, X. Zhang, and H. Li, "A multi-scale 3D Otsu thresholding algorithm for medical image segmentation," *Digit. Signal Process.*, vol. 60, no. 1, pp. 186–199, 2017.
- [34] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, 1986.
- [35] P. Sahoo, C. Wilkins, and J. Yeager, "Threshold selection using Renyi's entropy," *Pattern Recognit.*, vol. 30, no. 1, pp. 71–84, 1997.
- [36] A. G. Shanbhag, "Utilization of information measure as a means of image thresholding," *CVGIP, Graph. Models Image Process.*, vol. 56, no. 5, pp. 414–419, 1994.
- [37] J. Bernsen, "Dynamic thresholding of gray-level images," in *Proc. 8th Int. Conf. Pattern Recognit.*, Paris, France, 1986, pp. 1251–1255.
- [38] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern Recognit.*, vol. 33, no. 2, pp. 225–236, 2000.
- [39] J. Wu, H. Xiong, J. Chen, and W. Zhou, "A generalization of proximity functions for k -means," in *Proc. 7th IEEE Int. Conf. Data Mining*, Oct. 2007, pp. 361–370.

XINHUA CAO received the Ph.D. degree. He is currently an Instructor of radiology at Harvard Medical School and the Nuclear Medicine Information System Application Manager of the Boston Children's Hospital. His interests include research and development of advanced algorithms and computational tools to improve the efficiency and productivity of clinical practice and research in nuclear medicine and radiology.

TAIHAO LI received the Ph.D. degree in information system engineering from the University of Tokushima, Japan, in 2006. From 2006 to 2011, he was a Post-Doctoral Researcher with Harvard University. He is currently a Professor with the College of Medical Instruments, Shanghai University of Medicine and Health Sciences, China. His research interests include affective computing, image processing, and artificial intelligence.

HONGLI LI is currently a Professor with the Third Military Medical University, Chongqing, China. Her research interests are in quantitative analysis of biomedical images.

SHUNREN XIA received the Ph.D. degree from Southeast University, in 1992. He is currently a Professor with the Department of Biomedical Engineering, Zhejiang University, Hangzhou, China. He has authored or co-authored more than 160 papers in medical imaging and image processing. His research interests include machine learning, image processing, medical imaging, and computer-aided diagnosis on medical images.

FUJI REN received the Ph.D. degree from the Faculty of Engineering, Hokkaido University, Sapporo, Japan, in 1991. He is currently a Professor with the Department of Information Science and Intelligent Systems, Tokushima University, Tokushima, Japan. His current research interests include natural language processing, machine translation, artificial intelligence, language understanding and communication, robust methods for dialogue understanding, and affective information processing and knowledge engineering.

YE SUN received the M.D. and Ph.D. degrees. She is currently an Instructor of ophthalmology at Harvard Medical School and the Boston Children's Hospital. Her research interests focus on the role of immune-vascular crosstalk, neuro-inflammation and neuro-vascular interaction in the development of vascular eye disorders, including age-related macular degeneration, retinopathy of prematurity, and diabetic retinopathy, and quantitative imaging.

XIAOYIN XU received the Ph.D. degree. He is currently an Associate Professor with the Department of Radiology, Brigham and Women's Hospital, Harvard Medical School. His research interests are in computational science and image processing.