



# A novel relative homogeneity thresholding method with optimization strategy

Hong Zhang<sup>1</sup> · Yi-Jui Chiu<sup>2</sup> · Jiulun Fan<sup>1</sup>

Received: 22 November 2018 / Accepted: 28 June 2019 / Published online: 11 July 2019  
© Springer-Verlag London Ltd., part of Springer Nature 2019

## Abstract

Determining thresholds by measuring class variance is highly effective for image segmentation. Otsu's method and its derivatives are common approaches that are both simple and adaptable. In spite of these methods' excellent segmentation performance, images with particular gray distributions cause a thresholding bias that limits their usefulness. We explore the limitations of Otsu's method and apply other evaluation criteria. In particular, we determine the relative homogeneity between the object and the background and then use it as a classification criterion along with a new binary thresholding method. Our method employs a histogram-smoothing method to improve valley-point selection, establishes a uniformity measure to identify the region with the best homogeneity, and identifies an optimization function for obtaining the best values for the adjustable parameters and threshold value. We also introduce a multilevel thresholding criterion based on a binary thresholding approach. Experiments using real and ground truth test images confirm the validity of our proposed method. Our method also offers a denoising ability when configured to use neighborhood information.

**Keywords** Thresholding segmentation · Relative homogeneity · Optimization strategy · Multilevel thresholding

## 1 Introduction

Image segmentation is a fundamental and crucial processing step in image analysis and computer vision [1]. Segmentation consists of partitioning a digital image into meaningful parts [2] or several connected regions based on a similarity criterion such as probabilistic features [3], color features [4], or shape features [5]. Medical image segmentation techniques play an essential role in clinical diagnosis, surgical planning, and treatment evaluation [6], as well as in tissue labeling according to organ intensity and shape [7]. Among image segmentation methods, thresholding is one of the most widely used techniques due

to its simplicity [8]. Furthermore, Otsu's thresholding discriminant [9] is a popular criterion because of its excellent segmentation performance on real images [10].

Otsu's work spawned additional research as well. Xue et al. [11, 12] researched the close relationships between Otsu's method and other popular approaches. Vala and Baxi [13] studied various algorithms related to Otsu's method. Zou et al. [14] analyzed the relationship between Otsu's method and the Pearson correlation coefficient and provided a novel interpretation of Otsu's method from the perspective of maximizing image similarity. For images with a large variance between the object and background, Fan [15] proposed a modified valley-emphasis method using a weight term to adjust threshold bias. Yuan [16] applied a parameter weighted on the object variance to force the threshold value to reside at the valley points between peaks or the left bottom rim of a single peak histogram. Harb et al. [17] presented an improved algorithm to preserve high-frequency components of an image. By combining a grid-based model with the estimated background map, Moghaddam and Cheriet [18] proposed an adaptive and parameterless generalized method. By combining Otsu's circular feature and other color and

✉ Yi-Jui Chiu  
chiuyijui@xmut.edu.cn

Hong Zhang  
zhmlsa@xupt.edu.cn

Jiulun Fan  
jiulunf@xupt.edu.cn

<sup>1</sup> Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi, China

<sup>2</sup> Xiamen University of Technology, Xiamen, Fujian, China

texture features, Lai and Rosin [19] developed an efficient algorithm to solve the two-class thresholding problem. Cai et al. [20] searched an iterative three-class method, achieving better performance when identifying weak objects and revealing the structures of complex objects. In addition, using the 3D Otsu and multiscale image representations, Feng et al. [21] advanced a novel thresholding algorithm for medical image segmentation. He and Schomaker [22] presented an iterative deep learning framework, using iterative learning of enhanced images and Otsu's global thresholding techniques to obtain binary images efficiently. For multilevel image segmentation, the authors [23–25] also researched some methods to search for optimal thresholds efficiently and accurately.

All of these methods improved the segmentation performance of Otsu's algorithm by adding multidimensional information, parameter-optimization strategies, weight terms, and intelligent learning. However, none of these have changed Otsu's principle and structure. Therefore, for images with a large distribution-variance difference between objects and the background, the Otsu's and related methods provide a biased threshold value, making it difficult to extract objects integrally. To solve the problem, Chen [26] analyzed the limitations of Otsu's algorithm and applied homogeneity information from the foreground and background to develop a new binary method. Differing from those improved methods, Chen's method focuses on measuring the homogeneity of objects, reducing the negative effects from unequal distribution variances. Experimental results validate its success. Based on Chen's method, we have researched further, introducing gray probability-distribution information and presenting a new criterion function. Furthermore, considering the uniformity-distribution information of both the foreground and background, we have explored a relative homogeneity thresholding criterion between classes [27].

In this paper, we propose a new relative homogeneity criterion using region-probability features. We use a new probability-weighted item to adjust the proportion of relative homogeneity and apply a smoothing method to filter valley points. We then apply a uniformity measure to determine the threshold value region having better homogeneity. We design the characteristic function to select parameters for improving accuracy and adaptability. Thresholding results with real images show that our new method achieves better segmentation accuracy for images with very unequal distribution variances between classes, including skewed and heavy-tailed distributions. We also deduce a multilevel thresholding criterion based on the binary discriminant.

## 2 The criterion of focusing on objects

In Otsu's criterion for binary thresholding, object and background pixels are viewed as having a uniform or homogeneous gray distribution. However, it is considered only partially. For some images, one class may have greater distribution uniformity or homogeneity than other classes, which means that a biased threshold estimation may result using Otsu's method.

As a remedy, Chen [26] defined an alternative discriminant criterion, which assumes an object has a homogeneous gray distribution and focuses primarily on information from the segmented object. This method then selects an optimal gray level  $t^* \in \{0, 1, 2, \dots, L-1\}$  of an image to minimize the criterion  $J_{LC}(t)$ :

$$J_{LC}(t) = \frac{\left(\frac{P_1(t)}{P_2(t)}\right)^\alpha}{\frac{\sum_{(x,y) \in O} [\lambda(g(x,y) - m)^2 + (1 - \lambda)(\bar{g}(x,y) - m)^2]}{\sum_{(x,y) \notin O} [\lambda(g(x,y) - m)^2 + (1 - \lambda)(\bar{g}(x,y) - m)^2]}}, \quad (1)$$

where  $t \in \{0, 1, 2, \dots, L-1\}$  is the gray level of image,  $O$  denotes the pixel set belonging to the object,  $P_1(t) = \frac{1}{|O|}$ ,  $P_2(t) = \frac{1}{N-|O|}$ ,  $N$  is the total number of pixels in an image, and  $m = \frac{1}{|O|} \sum_{(x,y) \in O} g(x,y)$ .  $g(x,y)$  is the gray level of pixel point  $(x,y)$ , with  $\bar{g}(x,y) = \frac{1}{|W|} \sum_{(r,l) \in W} g(r,l)$  as the neighboring average gray of  $(x,y)$ .  $|W|$  is the size of window centered at  $(x,y)$  and usually taken as  $3 \times 3$  or  $5 \times 5$  pixels. Additionally,  $\lambda$  and  $\alpha$  are adjustable parameters, with  $\lambda(0 \leq \lambda \leq 1)$  being the proportion between gray and its neighborhood average and  $\alpha(\alpha \geq 0)$  being used to adjust  $(P_1(t)/P_2(t))^\alpha$ .

In Eq. (1), the variance of the background to an object is counted, with the numerator term measuring the similarity within an object class. The more similar pixels in an object class, the smaller the numerator value is. The denominator measures the dissimilarity of the background class to the object class, with larger values indicating that the two classes are more scattered. This criterion emphasizes the similarity of the object classes themselves and, more importantly, the dissimilarity of the background to the object. The threshold value is selected based on the homogeneity of the object and the heterogeneity of the background, to reduce the effect of a large variance difference between the classes.

By introducing the probability distributions of a pixel and its neighborhood average gray in the calculation of Eq. (1), we obtain a more detailed description. Thus, we determine a new thresholding criterion using the gray distribution information:

$$J_O(t) = \left(\frac{P_1(t)}{P_2(t)}\right)^\alpha \frac{\sum_{i \in O} \left[ \lambda \cdot p(i)(i - \mu(t))^2 + (1 - \lambda) \cdot p(i)(i - \bar{\mu}(t))^2 \right]}{\sum_{i \notin O} \left[ \lambda \cdot p(i)(i - \mu(t))^2 + (1 - \lambda) \cdot p(i) \cdot (i - \bar{\mu}(t))^2 \right]} \tag{2}$$

In this discriminant,  $O$  is the pixel set belonging to the object, with  $P_1(t) = 1 / \sum_{i \in O} p(i)$  and  $P_2(t) = 1 / \sum_{i \notin O} p(i)$ , where  $p(i)$  is the probability of gray level  $i$ .  $\mu(t) = \sum_{i \in O} ip(i) / \sum_{i \in O} p(i)$  denotes the mean value of the object area in the original image, and  $\bar{\mu}(t)$  denotes the object area mean value. We use a size of the neighboring window as  $3 \times 3$  pixels.

### 3 Relative homogeneity measure criterion between classes

In Sect. 2, we described the thresholding criterion  $J_O(t)$  to focus on the object’s gray distribution information. However, there are two limitations in  $J_O(t)$ .

The first is that  $J_O(t)$  only includes the consistency degree of the object (background) with respect to the other pixels. It is very likely that fewer pixels have better internal uniformity. That is, a larger number of pixels decrease the uniformity. Similarly,  $J_O(t)$  is a monotonic function with  $t$ . The quantity  $(P_1(t)/P_2(t))^\alpha$  provides a trade off, but it is difficult to obtain an effective adjustment. Thus, we cannot obtain an optimal thresholding value by using  $J_O(t)$ .

Figure 1 shows a sample image of red blood cells with  $\lambda = 1$  and  $\alpha = 0.5$ . Figure 1a is an original image. Figure 1b is the plot of  $J_O(t)$ , with the abscissa (x-axis) representing the gray level and the ordinate (y-axis) representing the value of  $J_O(t)$ . Figure 1b shows that  $J_O(t)$

increases monotonically with  $t$ . Removing the 0 values, the threshold value is 49 according to criterion  $J_O(t)$ . The segmentation result is shown in Fig. 1c, which shows some object pixels segmented to the background.

The second is that the image homogeneity region may be located in a region of low or high gray-level values. In this case, prior to applying criterion  $J_O(t)$ , we must determine the region with better homogeneity (i.e., the object region). This makes it inconvenient for determining thresholds.

In view of the above, combining the gray level and probability information of the object and background, we modify the discriminant  $J_O(t)$  to construct new criterion functions. For binary segmentation, there are two position cases with better homogeneity. If the region with better homogeneity is located at a low gray level, we build the criterion as

$$J_{OL1}(t) = (P_F(t))^\alpha \cdot \frac{\sum_{i=0}^{t-1} \left[ \lambda \cdot p(i) \cdot (i - \mu_1(t))^2 + (1 - \lambda) \cdot p(i) \cdot (i - \bar{\mu}_1(t))^2 \right]}{\sum_{j=t}^{L-1} \left[ \lambda \cdot p(j) \cdot (j - \mu_1(t))^2 + (1 - \lambda) \cdot p(j) \cdot (j - \bar{\mu}_1(t))^2 \right]} \tag{3}$$

and

$$J_{OL2}(t) = (1 - (P_F(t))^\alpha) \cdot \frac{\sum_{i=t}^{L-1} \left[ \lambda \cdot p(i) \cdot (i - \mu_2(t))^2 + (1 - \lambda) \cdot p(i) \cdot (i - \bar{\mu}_2(t))^2 \right]}{\sum_{j=0}^t \left[ \lambda \cdot p(j) \cdot (j - \mu_2(t))^2 + (1 - \lambda) \cdot p(j) \cdot (j - \bar{\mu}_2(t))^2 \right]} \tag{4}$$

Here,  $J_{OL1}(t)$  describes the relative information when the homogeneity foreground is in a low gray region. In addition, the homogeneity between the foreground and background should also be considered, described as

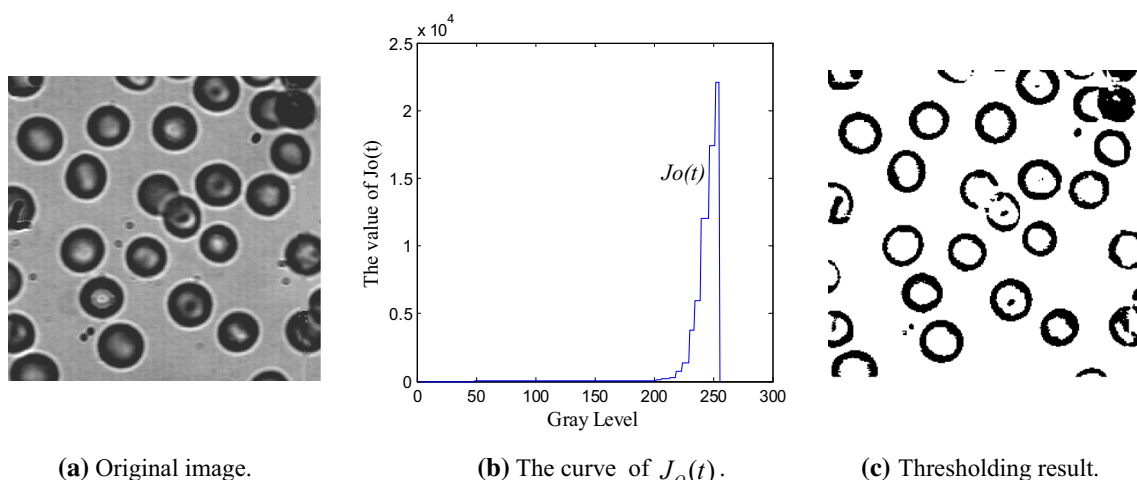


Fig. 1 Red blood cells image

$J_{OL2}(t)$ .  $p(i)$  and  $p(j)$  are the corresponding gray probabilities,  $P_F(t) = \sum_{i=0}^{t-1} p(i)$ ,  $\mu_1(t) = \sum_{i=0}^{t-1} ip(i) / \sum_{i=0}^{t-1} p(i)$ ,  $\mu_2(t) = \sum_{i=t}^{L-1} ip(i) / \sum_{i=t}^{L-1} p(i)$ , and  $\bar{\mu}_1(t)$  and  $\bar{\mu}_2(t)$  are the mean values of the local average images in the object and background areas.

$J_{OL1}(t)$  and  $J_{OL2}(t)$  differ from  $J_O(t)$  in that we apply  $(P_F(t))^\alpha$  to measure the relativity ratio from the homogeneous foreground to the total image and use  $1 - (P_F(t))^\alpha$  to measure the other part to make the relativity measurement more consistent with the real distribution.

Similarly, when the region with better homogeneity has a high gray range, the criteria should be

$$J_{OH1}(t) = (1 - P_B(t))^\alpha \cdot \frac{\sum_{i=0}^{t-1} [\lambda \cdot p(i) \cdot (i - \mu_1(t))^2 + (1 - \lambda) \cdot p(i) \cdot (i - \bar{\mu}_1(t))^2]}{\sum_{j=t}^{L-1} [\lambda \cdot p(j) \cdot (j - \mu_1(t))^2 + (1 - \lambda) \cdot p(j) \cdot (j - \bar{\mu}_1(t))^2]} \tag{5}$$

and

$$J_{OH2}(t) = (P_B(t))^\alpha \cdot \frac{\sum_{i=t}^{L-1} [\lambda \cdot p(i) \cdot (i - \mu_2(t))^2 + (1 - \lambda) \cdot p(i) \cdot (i - \bar{\mu}_2(t))^2]}{\sum_{j=0}^{t-1} [\lambda \cdot p(j) \cdot (j - \mu_2(t))^2 + (1 - \lambda) \cdot p(j) \cdot (j - \bar{\mu}_2(t))^2]} \tag{6}$$

Here,  $P_B(t) = \sum_{i=t}^{L-1} p(i)$  is the probability sum of high gray-level region.

In image thresholding, the homogeneity of foreground or background relative to the other should be considered simultaneously. Therefore, when the low gray-level region has better homogeneity, we combine Eqs. (3) and (4) to construct the thresholding-criterion function  $J_{OBL}(t)$ . In the same way, when region with better homogeneity is in the high gray-level range, we build  $J_{OBH}(t)$ . The two methods are:

$$J_{OBL}(t) = J_{OL1}(t) + J_{OL2}(t) \tag{7}$$

and

$$J_{OBH}(t) = J_{OH1}(t) + J_{OH2}(t). \tag{8}$$

The above criteria include the relative homogeneity information of both the foreground and the background, which is more reasonable for image-distribution descriptions than the  $J_O(t)$  method. The relative measurements of the object and background also enable us to determine the bias caused by the class distribution difference.

## 4 Optimization of threshold point and criterion

### 4.1 Smoothing of 1d gray histogram

In some thresholding algorithms, the threshold point is selected by traversing all gray levels. In fact, the optimal threshold value is found among valley points in histogram-based methods. Therefore, instead of calculating all gray points, it is more efficient to select the optimal threshold point from among the valley points of the histogram. In real images, there are many valley points, as with the rice grains image histogram shown in Fig. 2. However, some of them are only local valley points and not globally optimal threshold points. Hence, to improve the search process, we apply a regression function to smooth the one-dimensional gray histogram (1d\_histogram) curve. We choose a locally weighted scatter plot smoothing (LOESS) function to eliminate scattered or isolated points that interfere with the overall distribution. From this, we obtain a smoothed histogram that is good for finding the optimal threshold points as shown in Fig. 3. From our experiment, we set the width of the LOESS bandwidth to 10.

Figures 2 and 3 show the original 1d\_histogram and its smoothed histogram curve. The abscissa (x-axis) is the gray level, and the vertical axis (y-axis) is the gray probability. The valley points of the 1d\_histogram are marked with \*.

Comparing Fig. 2 with Fig. 3, the LOESS function smoothed the curve’s burrs, producing fewer valley points and removing false valleys and stochastic noise. The reduced number of valley points speeds and simplifies in finding the optimal threshold value.

### 4.2 Application of uniformity measure

In Sect. 3, we proposed two thresholding functions  $J_{OBL}(t)$  and  $J_{OBH}(t)$  selected according to the homogeneity region.

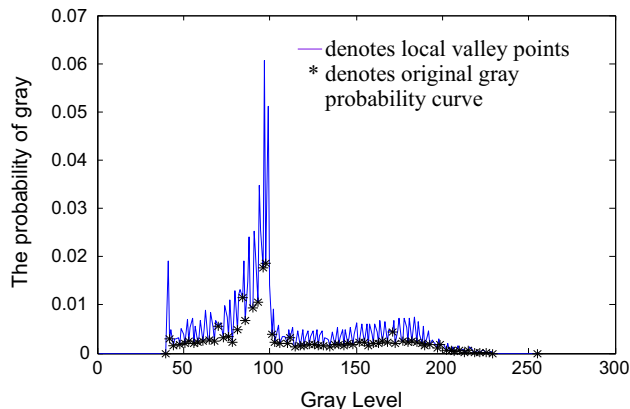


Fig. 2 Original 1d\_histogram

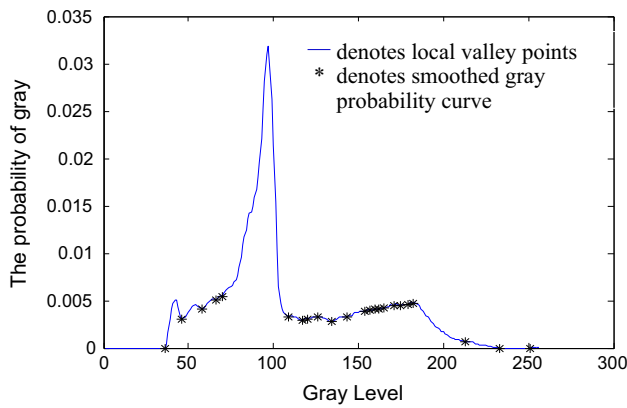


Fig. 3 LOESS smoothed histogram (bandwidth = 10)

To determine the homogeneity position automatically, we offer a method based on the uniformity measurement,  $U$ . In view of its performance-evaluation ability, the uniformity measure is a popular metric [21, 23]. The uniformity measure is computed as

$$U = 1 - 2 * (K - 1) * \frac{\sum_{j=0}^{K-1} \sum_{i \in R_j} (f_i - \mu_j)^2}{M * N * (f_{\max} - f_{\min})^2}, \quad (9)$$

where  $K$  is the classification number,  $R_j$  is the region segmented by the threshold value  $t_{(K-1)}$ ,  $f_i$  is the gray level of pixel  $i$ ,  $\mu_j$  is the mean gray level in  $j$  the region,  $f_{\min}$  and  $f_{\max}$  are the minimum and maximum gray levels, and  $M * N$  is the total number of image pixels. In formula (9), larger  $U$  indicates better uniformity.

The above metric usually measures the uniformity of segmented images, but it is also useful to determine the initial threshold value. For most images, the threshold value is on the side with better homogeneity.

Figures 4 and 5 show synthetic aperture radar (SAR) and lymph images illustrating this behavior. Figures 4a and 5a are original images, and Figs. 4b and 5b are the 1d-gray histogram curves of the two images. Although there are threshold deviations caused by large differences in

between-class variance, the threshold position indicates the region with better homogeneity. In Figs. 4b and 5b, we use a solid line to denote the optimal uniformity threshold value and a dashed line to denote the middle of the gray-level range.

If the optimal uniformity threshold value  $U$  is in a low gray-level range, as in Fig. 4, we choose  $J_{OBL}(t)$  as the criterion function. If the threshold value is in a high gray-level range, as in Fig. 5, we use  $J_{OBH}(t)$ . In this way, we choose the criterion function on the basis of the gray distribution automatically. We also highlight that we reduce the number of calculations by considering only the valley points of the smoothed curve.

### 5 Relative homogeneity criterion with optimized parameters

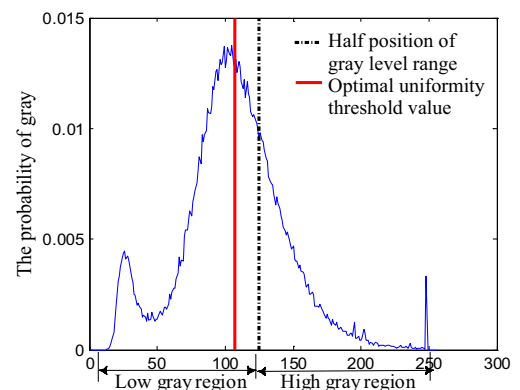
There are two adjustable parameters  $\lambda$  and  $\alpha$  in the criteria functions  $J_{OBL}(t)$  and  $J_{OBH}(t)$ .  $\lambda$  is the weighted coefficient of gray and neighboring average gray. If  $\lambda = 1$ , neighboring average information is not considered. When  $\lambda = 0$ , the neighboring values are considered to obtain better denoising performance.

The parameter  $\alpha$  adjusts the relative homogeneity proportion of the object and background and ranges between 0 and 1. When  $\alpha = 1$ , the weight coefficients are the accumulated probabilities of the foreground and the background. Thus,  $(P_F(t))^\alpha = P_F(t)$  and  $1 - (P_F(t))^\alpha = P_B(t)$  in  $J_{OBL}(t)$ , and  $\alpha = 0$  and  $(P_B(t))^\alpha = P_B(t)$  and  $1 - (P_B(t))^\alpha = P_F(t)$  in  $J_{OBH}(t)$ . When  $\alpha = 0$ , only one part of the relative homogeneity is considered, similar to using only  $J_O(t)$ . Thus,  $(P_F(t))^0 = 1$ ,  $1 - (P_F(t))^0 = 0$  and  $(P_B(t))^0 = 1$ ,  $1 - (P_B(t))^0 = 0$ . Therefore,  $\alpha$  is an important parameter in  $J_{OBL}(t)$  or  $J_{OBH}(t)$  and is chosen to combine the image distributions.

Fig. 4 SAR image

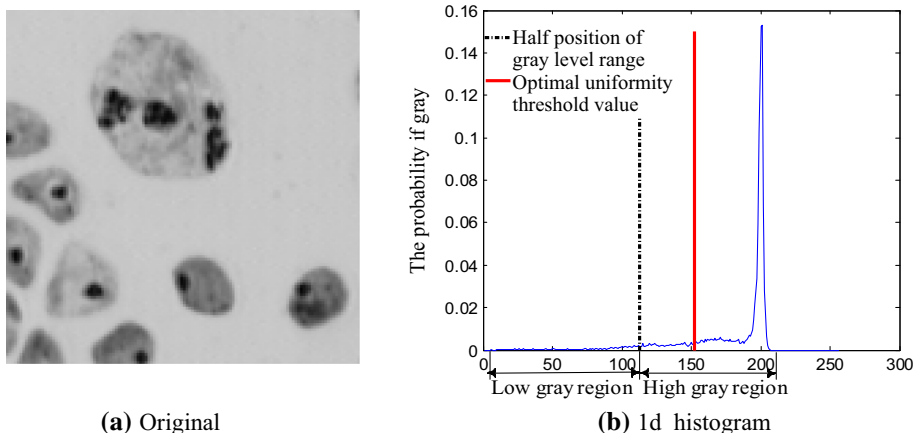


(a) Original



(b) 1d\_histogram

Fig. 5 Lymph image



### 5.1 Parameter-optimization criterion

During thresholding segmentation, the optimal threshold value should be located at the valley point of the gray histogram. In order to emphasize the characteristics of valley points and eliminate some non-threshold special points, we also consider information from points in the neighborhood of the valley points.

Assuming the original image histogram in  $t$  is denoted as  $f(t)$ , the valley-emphasis [15] function  $\bar{f}(t)$  is constructed as  $\bar{f}(t) = f(t - m) + f(t - m + 1) + \dots + f(t) + \dots + f(t + m - 1) + f(t + m)$ .

$$(10)$$

Since the optimal threshold value will be located at the valley point of the histogram, its gray and neighborhood gray probability are both minimum values, and  $\bar{f}(t)$  is also a minimum. We simplify Eq. (10) as

$$\bar{f}(t) = \sum_{i=-m}^m f(t + i), \tag{11}$$

where  $m$  is a positive integer. Greater  $m$  values lengthen the filter.

In the case of some gray distribution histograms, we found that  $\bar{f}(t)$  is only a condition for ideal threshold point, and the gray-level variation is smaller near the optimal threshold point. In view of this, we propose another constraint function  $\Delta f(t)$  to represent the sum of gray absolute differences near the threshold  $t$ :

$$\Delta f(t) = |f(t) - f(t - m)| + |f(t) - f(t - m + 1)| + \dots + |f(t) - f(t + m - 1)| + |f(t) - f(t + m)|. \tag{12}$$

We simplify this as:

$$\Delta f(t) = \sum_{i=-m}^m |f(t) - f(t + i)|, \tag{13}$$

where  $m$  is a positive integer.

In sum, we introduce a new compound function  $F(t)$  to take the minimum valley point and gray-level deviation into account:

$$F(t) = \bar{f}(t) \cdot \Delta f(t). \tag{14}$$

In our use of  $F(t)$ , we set  $m = 3$ . At the optimal threshold point, the value of  $F(t)$  should be a minimum. We choose parameter  $\alpha$  so that it minimizes  $F(t)$ :

$$\alpha^* = \text{Arg} \min_{0 < t < L-1, 0 \leq \alpha \leq 1} \{F[t(\alpha)]\}. \tag{15}$$

After determining parameter  $\alpha^*$ , we obtain the optimal threshold value  $t^*$  according to Eq. (7) or (8), i.e.,

$$t^* = \text{Arg} \min_{0 < t < L-1} \{J_{OBL}[t(\alpha^*)]\} \tag{16}$$

or

$$t^* = \text{Arg} \min_{0 < t < L-1} \{J_{OBH}[t(\alpha^*)]\}. \tag{17}$$

### 5.2 Implementation of relative homogeneity method

Incorporating all of the preceding, we perform the following steps in our algorithm.

**Step 1** Smooth the image 1d gray histogram by applying the LOESS function to eliminate local valley points using a bandwidth of 10.

**Step 2** Identify the valley points of the smoothed histogram.

**Step 3** Determine threshold value region having better homogeneity using the uniformity measure function.

**Step 4** According to the label, select the appropriate relative homogeneity thresholding criterion  $J_{OBL}$  or  $J_{OBH}$ .

**Step 5** Use the synthetic valley-emphasis function  $F(t)$  to select the optimal exponent parameter  $\alpha^*$ .

**Step 6** Search all valley points to find the optimal threshold value  $t^*$  using criterion function  $J_{OBL}$  or  $J_{OBH}$ .

### 5.3 Multilevel thresholding

When there are multiple classifications in an image, the uniformity measure  $U$  has classification number  $K > 2$ . We can construct a multilevel thresholding criterion using our proposed binary algorithm. Using Eq. (9), we identify the class with the lowest homogeneity. We disregard denoising effects by setting  $\lambda = 1$  and obtain the criterion function of the lowest homogeneity region as

$$\begin{aligned}
 J_{Ol}(t_1^*, t_2^*, \dots, t_{K-1}^*) &= \left( 1 - \sum_{\substack{q=1 \\ q \neq l}}^K (P_q(t_{q-1}))^\alpha \right) \cdot \\
 &\frac{\sum_{i=l}^{t_l} [p(i) \cdot (i - \mu_l(t_{l-1}))^2]}{\sum_{\substack{q=1 \\ q \neq l}}^K \left[ \sum_{j=t_{q-1}}^{t_q} [p(j) \cdot (j - \mu_l(t_{l-1}))^2] \right]} \\
 &= \left( 1 - \sum_{\substack{q=1 \\ q \neq l}}^K (P_q(t_{q-1}))^\alpha \right) \cdot \frac{P_l(t_{l-1}) \cdot \sigma_l^2(t_{l-1})}{\sum_{\substack{q=1 \\ q \neq l}}^K [P_q(t_{q-1}) \cdot \sigma_{ql}^2(t_{l-1})]}, \tag{18}
 \end{aligned}$$

where  $t_1^*, t_2^*, \dots, t_{K-1}^*$  are the  $K - 1$  thresholding values,  $l$  is the region with lowest homogeneity,  $q$  indicates the number of other regions, and  $P_q(t_{q-1})$  is the probability sum excluding the  $l$ th region. For the  $q$  regions, the criterion is:

$$\begin{aligned}
 J_{Oq}(t_1^*, t_2^*, \dots, t_{K-1}^*) &= (P_q(t_{q-1}))^\alpha \cdot \\
 &\frac{\sum_{i=l}^{t_q} [p(i) \cdot (i - \mu_q(t_{q-1}))^2]}{\sum_{\substack{r=1 \\ r \neq q}}^K \left[ \sum_{j=t_r}^{t_r} [p(j) \cdot (j - \mu_q(t_{q-1}))^2] \right]} \\
 &= (P_q(t_{q-1}))^\alpha \cdot \frac{P_q(t_{q-1}) \cdot \sigma_q^2(t_{q-1})}{\sum_{\substack{r=1 \\ r \neq q}}^K [P_r(t_{r-1}) \cdot \sigma_{rq}^2(t_{q-1})]}, \tag{19}
 \end{aligned}$$

where  $\mu_l(t_{l-1})$ , and  $\mu_q(t_{q-1})$ ,  $\sigma_l^2(t_{l-1})$ , and  $\sigma_q^2(t_{q-1})$  are the mean values and variances of  $l$ th and  $q$ th regions, respectively.

Considering the homogeneity information from all regions, we construct the thresholding criterion as

$$\begin{aligned}
 J_{OB}(t_1^*, t_2^*, \dots, t_{K-1}^*) &= J_{Ol}(t_1^*, t_2^*, \dots, t_{K-1}^*) \\
 &+ \sum_{\substack{q=1 \\ q \neq l}}^K J_{Oq}(t_1^*, t_2^*, \dots, t_{K-1}^*). \tag{20}
 \end{aligned}$$

Then, the optimal threshold values are

$$(t_1^*, t_2^*, \dots, t_{K-1}^*) = \text{Arg} \min_{0 < t_1^*, t_2^*, \dots, t_{K-1}^* < L-1} \{J_{OB}(t_1^*, t_2^*, \dots, t_{K-1}^*)\}, \tag{21}$$

where  $l, q, r \in [1, K]$ . The parameter  $\alpha$  can be selected as with the binary thresholding method.

## 6 Experimental results and analysis

### 6.1 Thresholding results

To verify the performance of our proposed method, we designed an experiment to compare results from the following eight methods, including our own method: the 1-dimensional Otsu’s method (denoted as Otsu) [7], the modified valley-emphasis method (denoted as MOtsu) [15], the parameter-weight method on the object variance (denoted as WOtsu) [16], the iterative three-class method for sub-regions (denoted as IOtsu) [20], the focusing-on-objects method (denoted as FO) [26], our relative homogeneity method without optimization (denoted as ROB,  $\alpha = 0.5$ ) [27], our relative homogeneity criterion with  $\alpha$  parameter optimization (denoted as ROBP), and our relative homogeneity method with optimized valley points and  $\alpha$  (denoted as ROBVP). Here, we set parameter  $\lambda < 1$ .

We validated all methods through simulations in MATLAB. We used a computer with an Intel Core i7-4720HQ CPU at 2.60 GHz and 8 GB of memory running Microsoft Windows 8.

Figure 6a–j shows the 10 representative images from several fields: color, material structure, SAR, red blood cells, rice grains, numbers, lymph, cells under an optical microscope, an MRI slice, and a barcode. Their sizes are  $158 \times 159$ ,  $340 \times 304$ ,  $256 \times 256$ ,  $272 \times 265$ ,  $256 \times 256$ ,  $730 \times 203$ ,  $130 \times 130$ ,  $213 \times 166$ ,  $256 \times 256$  and  $256 \times 256$  pixels, respectively. Figure 7a–j shows their corresponding 1d\_histograms. Among them, the color, material structure, and SAR images have large between-class variance. The number, lymph, cells, MRI slice, and barcode images have skewed or heavy-tailed histograms. The red blood cells and rice grain images have small between-class variance. Figures 8, 9, 10, 11, 12, 13, 14, 15, 16, and 17 show the segmentation results of the 10 images with (a)–(j) showing the results of Otsu, MOtsu, WOtsu, IOtsu, FO, ROB, ROBP, and ROBVP, respectively, for each source image.

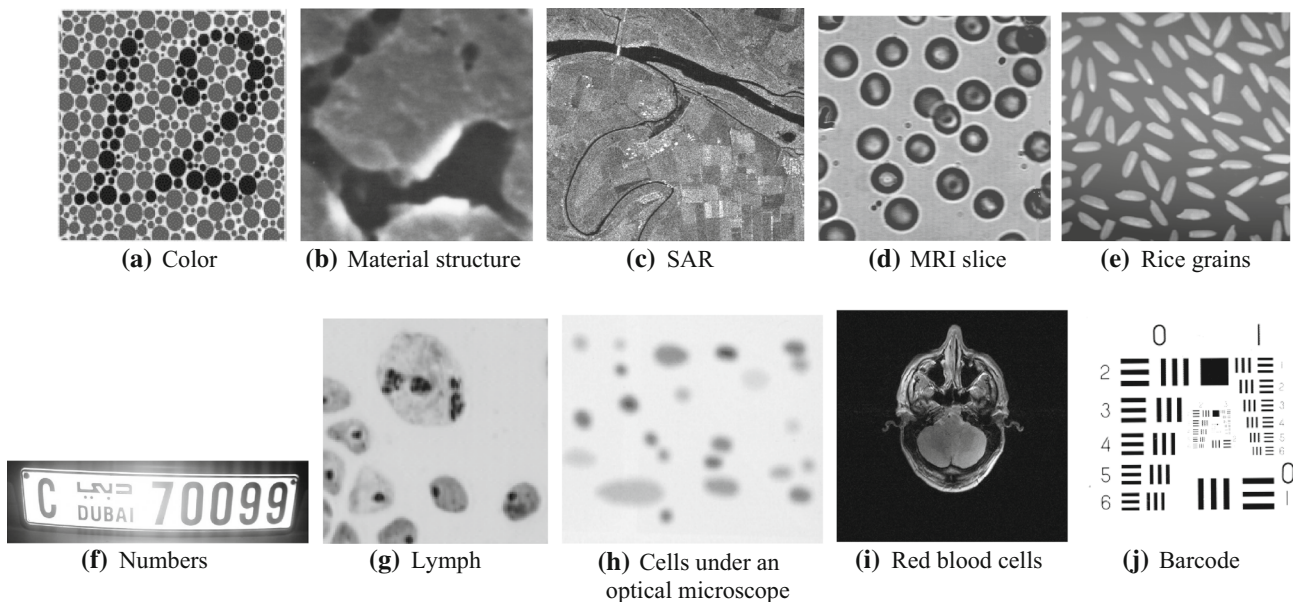


Fig. 6 Original images

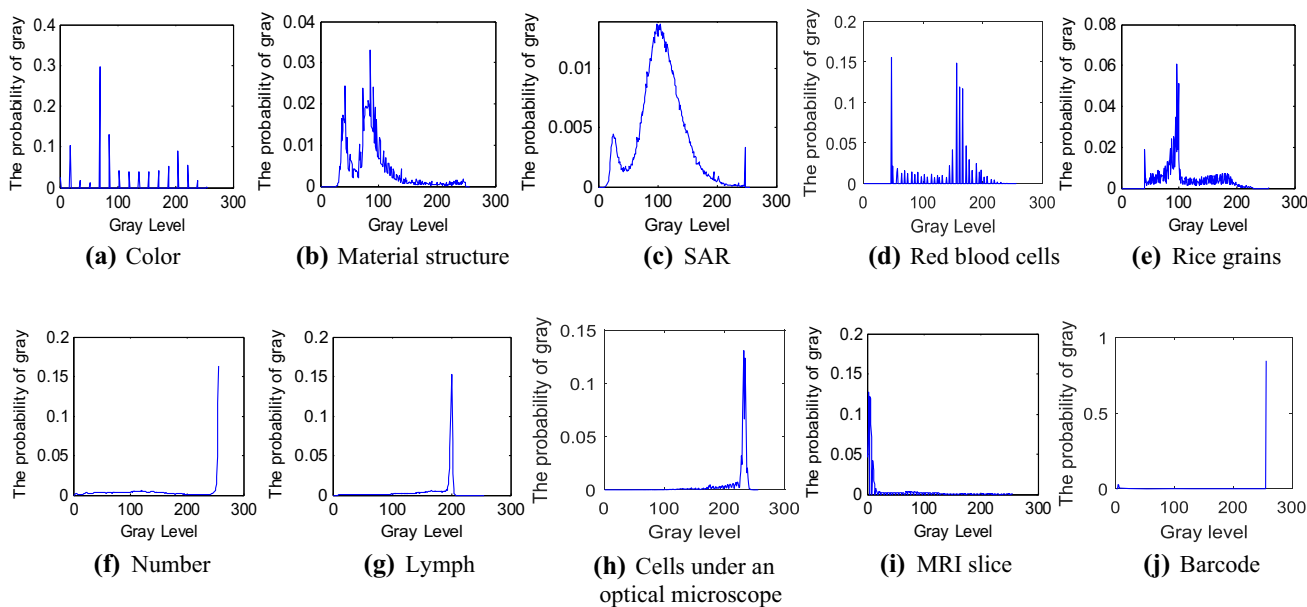


Fig. 7 1d\_histogram

The results from Figs. 8a, 9a, and 10a show that the simple Otsu’s method did not segment the object properly in the color, material structure, and SAR images. From Figs. 11, 12, 13, 14, 15, 16, and 17a, the integrity of segmented object was insufficient. The MOTsu method yielded an excellent result for the SAR image in Fig. 10b, with middling results for the red blood cells, rice grains, numbers, and MRI slice. MOTsu did not correctly extract the objects in the color, material structure, lymph, cells, and barcode images in Figs. 8, 9b, 14, 15b, and 17b. The Wotsu obtained perfect segmentation results in Figs. 11c,

14c, and 17c. Similarly, the IOtsu method had fine results in Figs. 11d and 15 and 16d, but both IOtsu and Wotsu showed limited adaptability to the overall mix of images.

The FO, ROB, ROBP, and ROBPV methods each had a different construction principle in the corresponding discriminant function. The FO method focused its information statistic on the object itself, segmenting the color image perfectly, but failing to do so for the other nine images with scattered objects or backgrounds. The ROB method, applying the relative homogeneity measurement, achieved preferable results in the red blood cells, lymph, and MRI



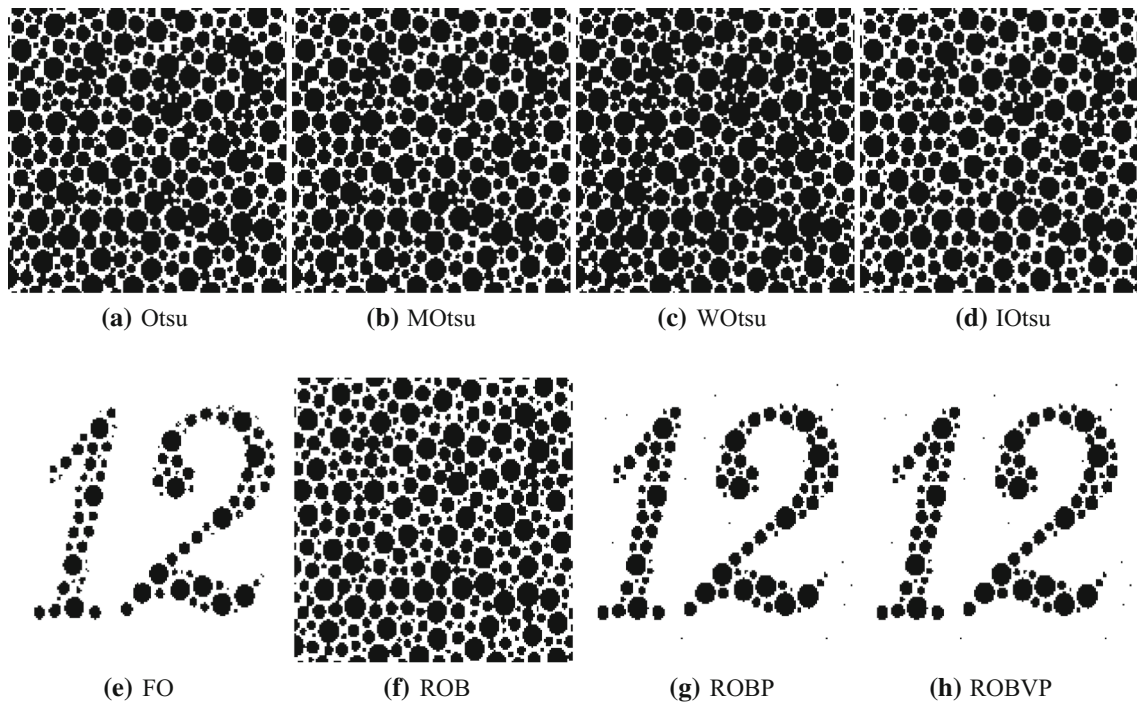


Fig. 8 Results from the color image

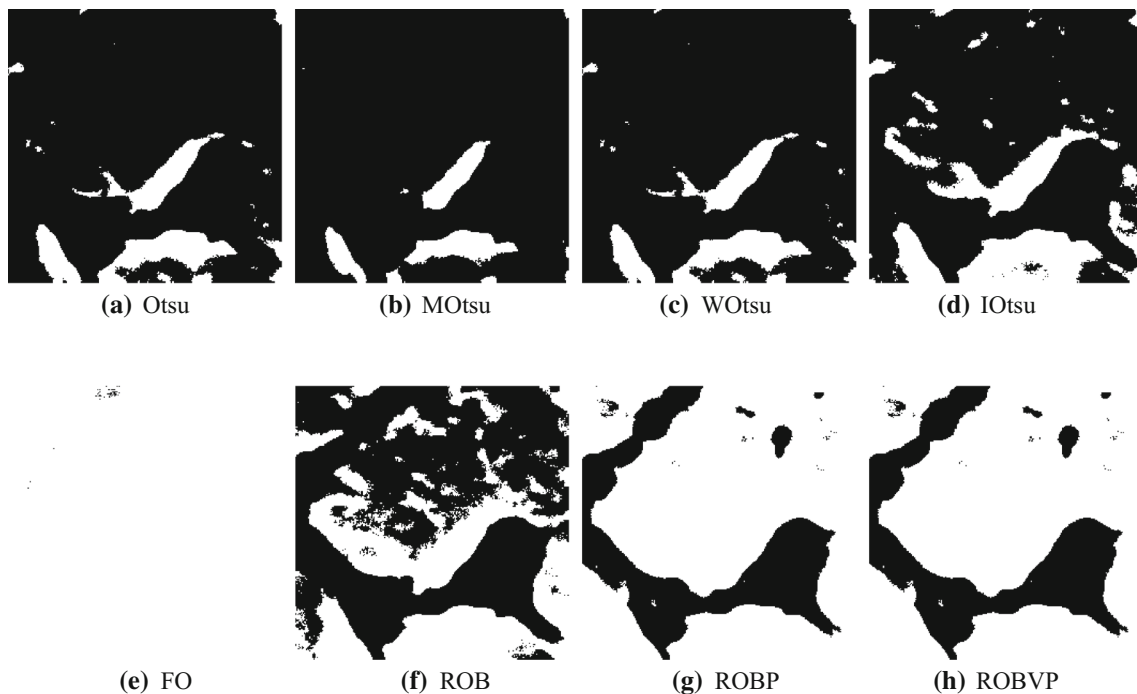


Fig. 9 Results from the material structure image

slice images shown in Figs. 11f, 14f, and 16f. However, the invariant  $\alpha$  value limited ROB’s adaptability to the other images. Our proposed ROBP and ROBVP methods achieved optimal results by extracting the most complete objects. This was especially true of the ROBVP method as

shown in (h) of Figs. 8, 9, 10, 11, 12, 13, 14, 15, 16, and 17.

Table 1 shows the optimal threshold value  $t^*$  and optimized parameter  $\alpha$  of the eight methods for the 10 images. The FO method’s threshold value tended to the lower gray

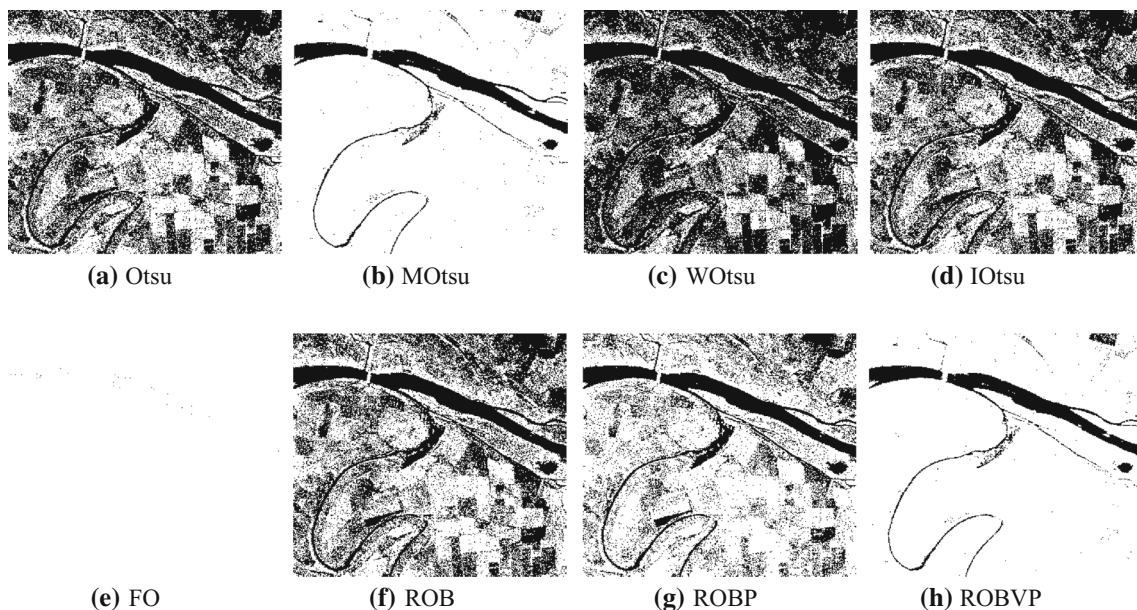


Fig. 10 Results from the SAR image

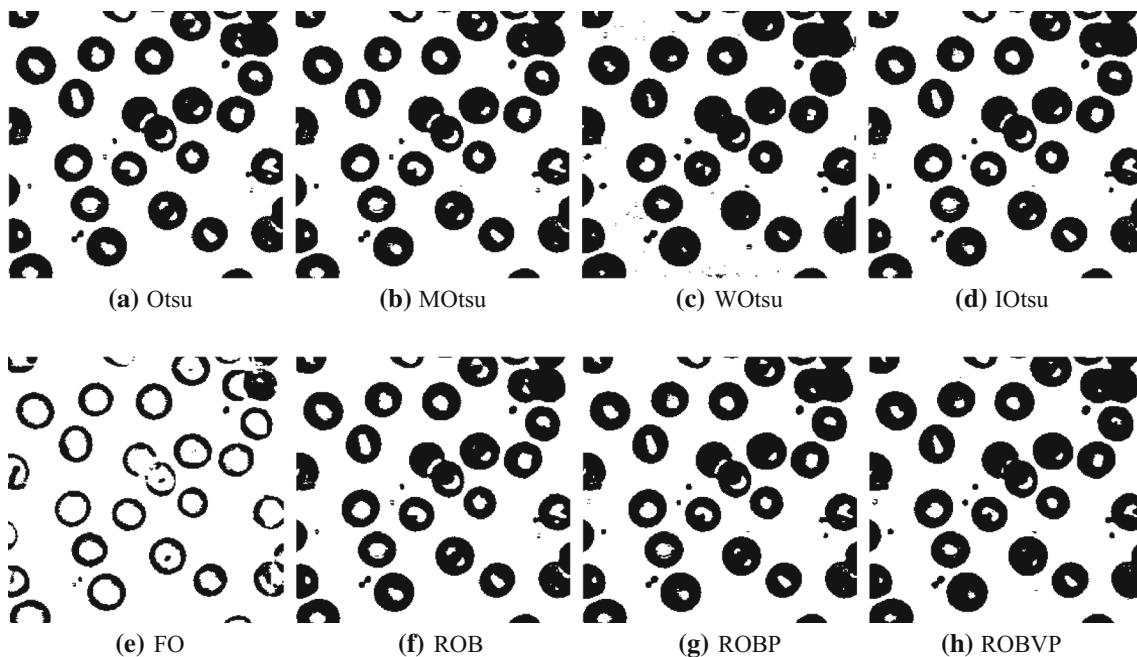


Fig. 11 Results from the red blood cells image

level, which was smaller than that of the other methods as a result of the monotonic criterion  $J_O(t)$ . The ROB and ROBVP methods had same threshold values in the material structure and number images, but ROBVP’s threshold value achieved more complete extraction results on the SAR, cells, and barcode images.

Moreover, the ROBVP method was faster as a result of its valley-points optimization strategy, as shown in

Table 2. For the 10 images, our proposed ROBVP had the best times, taking between 45.9 and 75% of the time required by the ROB method. It is remarkable that ROBVP’s running time was even shorter than that of the Otsu’s method on the color, rice, lymph, and MRI slice images.

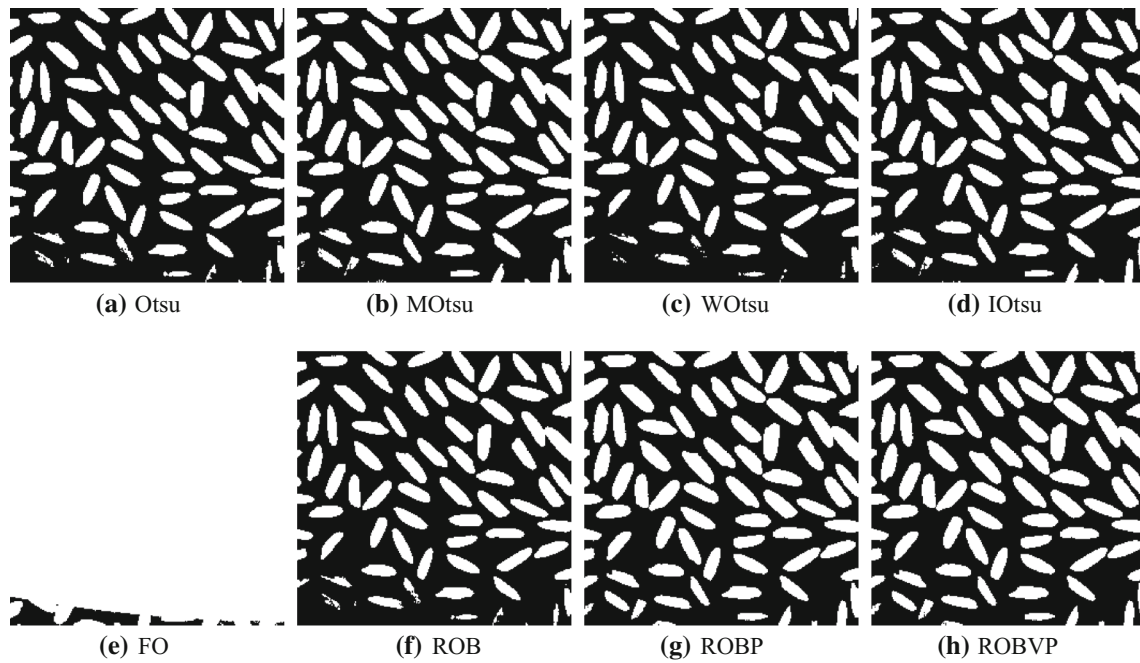


Fig. 12 Results from the rice grains image



Fig. 13 Results from the numbers image

### 6.2 Segmentation results of noisy image

The discriminant criteria in Eqs. (2)–(8) include neighborhood average information. If  $\lambda < 1$ , some denoising performance results. In this experiment, we compared the ROBVP method with  $\lambda = 1$  and  $\lambda = 0.5$ . Figures 18a, 19, and 20a are 3 representative images: color, SAR, and red blood cells with Gaussian noise of 0 mean value and 0.003 variance. Figures 18, 19, and 20b show the 1d\_histograms of the 3 noisy images. Figures 18, 19, and 20i, j are the results of ROBVP with  $\lambda$  set to 1 and 0.5, respectively. Figures 18, 19, and 20c–h show the results from the Otsu, MOtsu, WOtsu, IOtsu, ROB, and 2d\_Otsu methods for comparison. We do not include FO because of its non-ideal results.

When  $\lambda = 0.5$ , the denoising performance was better than with  $\lambda = 1$ . For the color image, only ROBVP ( $\lambda = 0.5$ ) obtained the correct threshold value. For the SAR and red blood cells images, the denoising results of the ROBVP

methods were superior. Table 3 shows the threshold values of the eight methods. When  $\lambda = 0.5$ , the threshold values from ROBVP were more accurate than ROBVP with  $\lambda = 1$  as well as the 2d\_Otsu method.

### 6.3 Performance evaluation

To analyze the segmentation results quantitatively, we assessed the performance of the methods according to the misclassification error (ME). ME can reflect the percentage of background pixels wrongly assigned to foreground and, conversely, foreground pixels wrongly assigned to background. For the two-class segmentation, ME can be expressed as [28]:

$$ME = 1 - \frac{|B_O \cap B_T| + |F_O \cap F_T|}{|B_O| + |F_O|}, \tag{22}$$

where  $B_O$  and  $B_T$  denote the object of the ground truth and test images,  $F_O$  and  $F_T$  are the background of the two

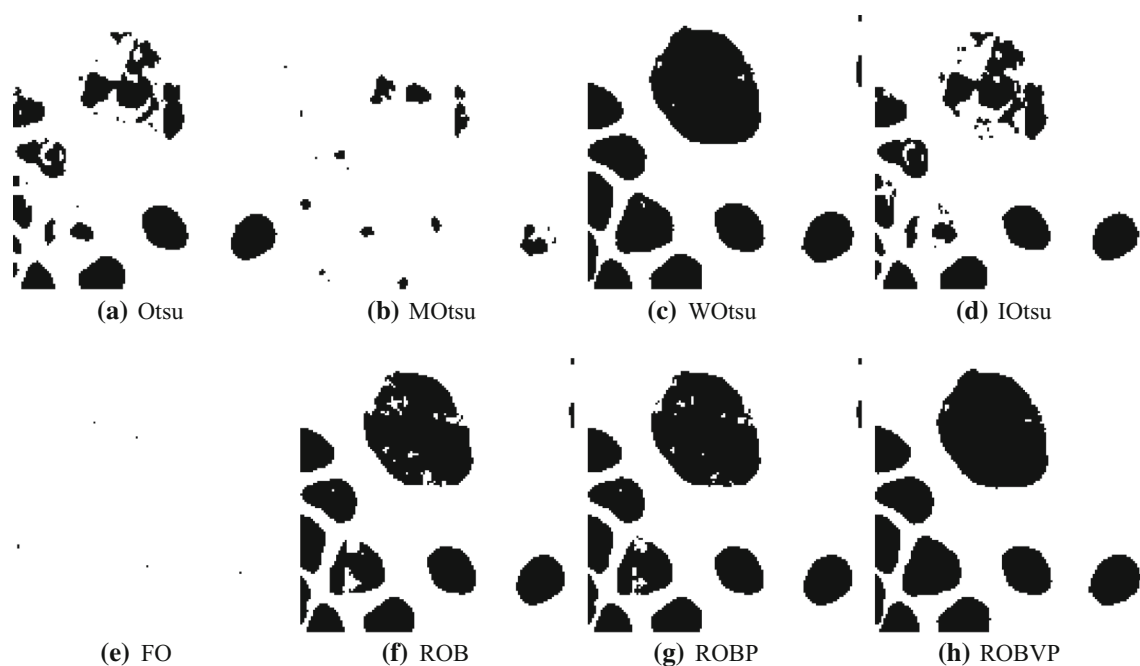


Fig. 14 Results from the lymph image

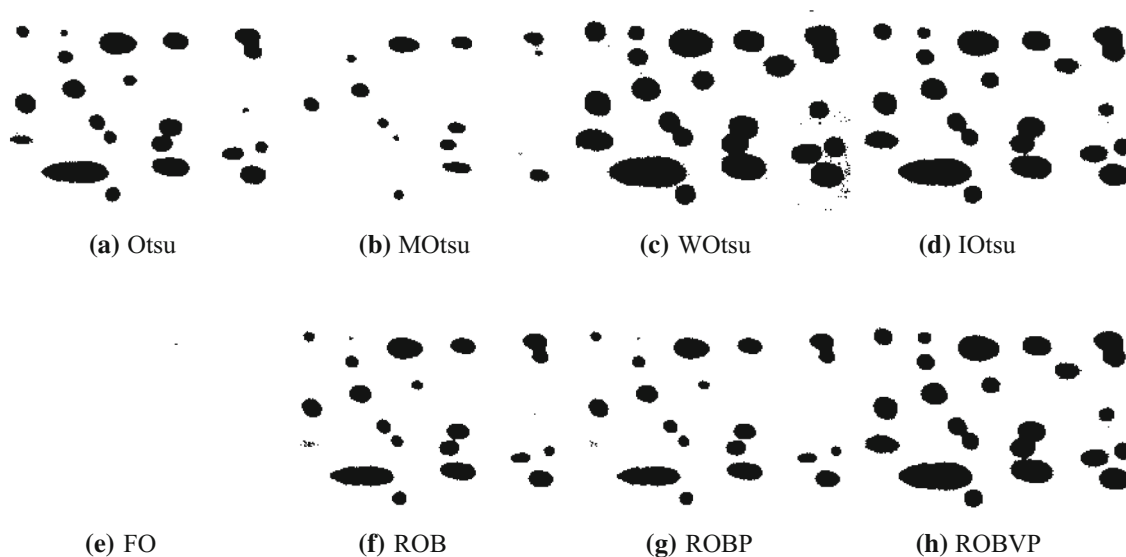


Fig. 15 Results from the cells under an optical microscope image

images, and  $|\bullet|$  is the operator to obtain the cardinality of a set. The range of ME is from 0 to 1, 0 denotes totally correct segmentation, and 1 means totally erroneous case; their scores vary from 0 for a totally correct segmentation to 1 for a totally erroneous case.

We selected four test images from a database of non-destructive testing (NDT) images for quantitative evaluation of the thresholding methods [28]. The ground truth images for the NDT images were chosen interactively by experts [29], and we downloaded them from the Web site

of Sezgin [30]. The size of the Test1 image is  $232 \times 131$  pixels, and the sizes of Test2, Test3, and Test4 are all  $256 \times 256$  pixels. Figures 21, 22, 23, and 24 show the results. Figures 21, 22, 23, and 24a are the original images, while Figs. 21, 22, 23, and 24b are the ground truth images. Figures 21, 22, 23, and 24c–j show the results of the Otsu, MOtsu, WOtsu, IOtsu, FO, ROB, ROBP, and ROBVP methods, respectively. Table 4 shows the segmentation threshold values and MEs of the 8 methods. The ROBVP method with  $\lambda$  set to 0.5 had an ME lower than the other 7

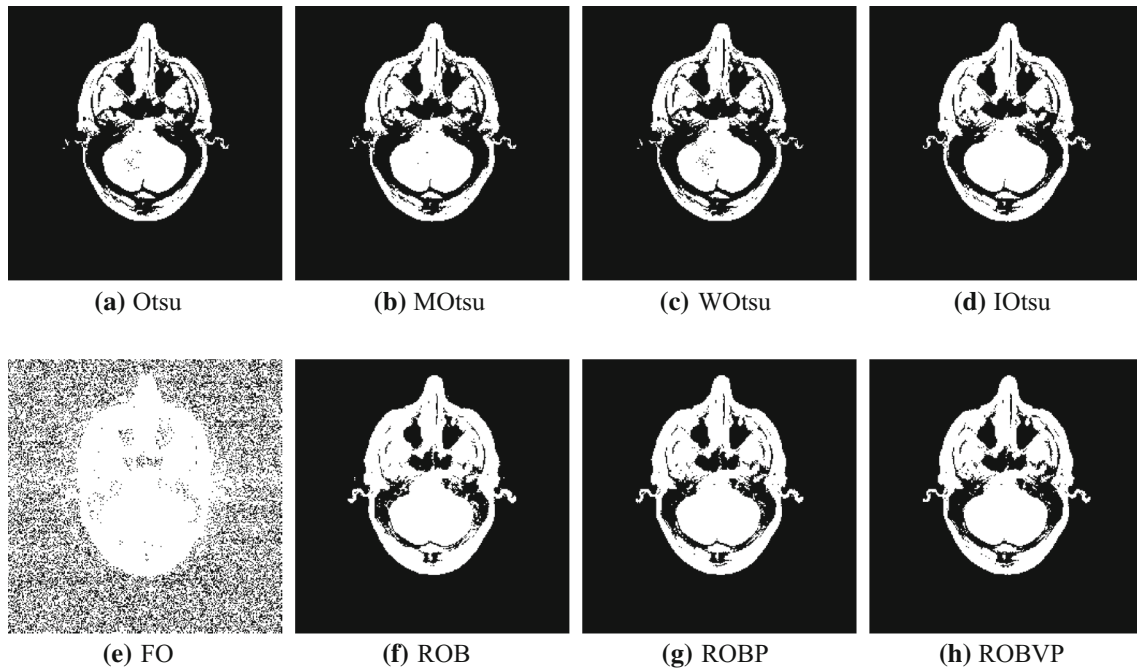


Fig. 16 Results from the MRI slice image

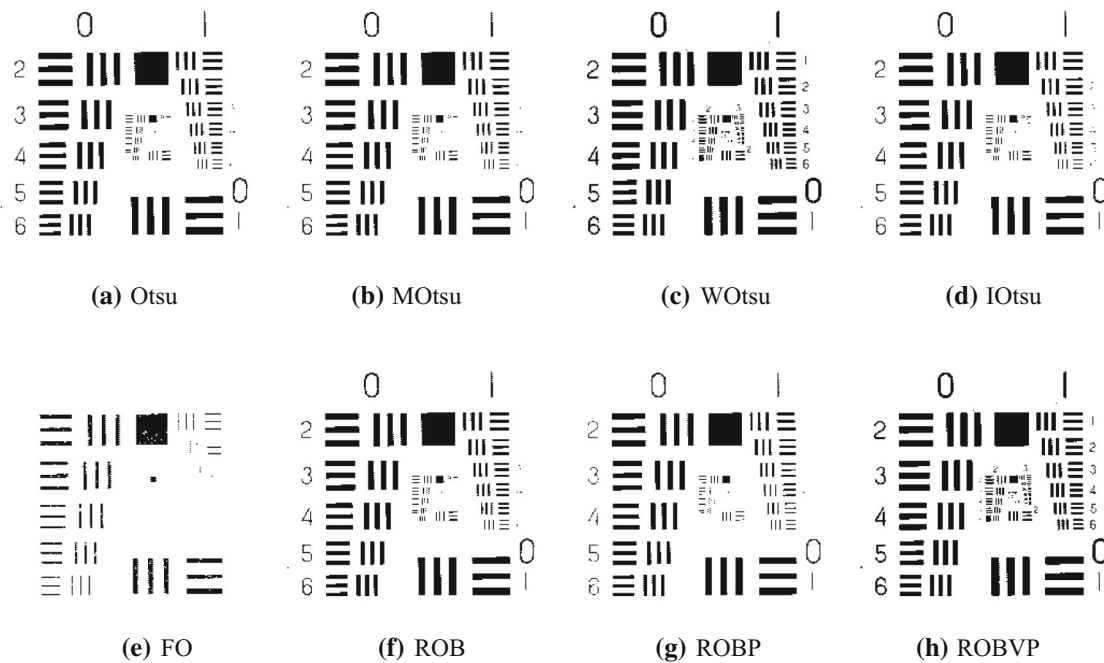


Fig. 17 Results from the barcode image

methods. The ME of the Test4 image was only 0.0054. We have marked the lowest ME values in bold in Table 4.

For evaluating the denoising performance of the methods, we used the common peak signal-to-noise ratio (PSNR) [31] measure. PSNR gives the similarity of a thresholded image against a reference image based on the root mean square error (RMSE) [31] of each pixel. The

larger the PSNR, the better the similarity. It is given by Eq. (23):

$$PSNR(x, y) = 20 \log_{10} \left( \frac{255}{\sqrt{RMSE(x, y)}} \right) \tag{23}$$

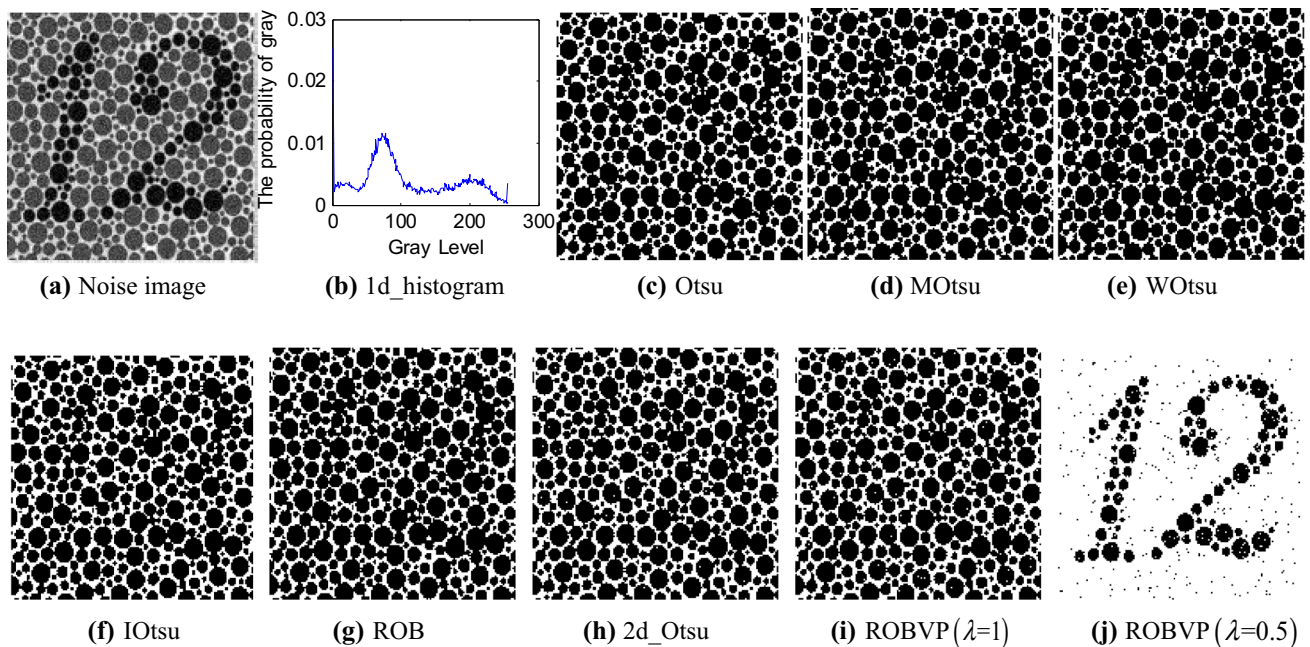
The segmentation results of two test images with Gaussian noise of 0 mean value and 0.003 variance are

**Table 1** Optimal threshold value  $t^*$  and optimized parameter  $\alpha$

Image	Otsu		MOtsu		WOtsu		IOtsu		FO		1d_OB		ROBP		ROBVP	
	$t^*$	$\alpha$	$t^*$	$\alpha$	$t^*$	$\alpha$	$t^*$	$\alpha$	$t^*$	$\alpha$	$t^*$	$\alpha$	$t^*$	$\alpha$	$t^*$	$\alpha$
Color	120	–	131	–	151	–	103	–	18	–	86	0.5	52	0.01	57	0.01
Material structure	126	–	150	–	127	–	106	–	29	–	89	0.5	63	0.01	63	0.01
SAR	107	–	49	–	121	–	107	–	14	–	102	0.5	88	0.08	47	0.01
Red blood cells	107	–	116	–	142	–	119	–	49	–	49	0.5	119	0.47	128	0.72
Rice grains	126	–	115	–	132	–	116	–	42	–	128	0.5	108	0.79	109	0.03
Number	171	–	205	–	195	–	158	–	4	–	173	0.5	222	0.03	222	0.03
Lymph	151	–	84	–	185	–	157	–	10	–	177	0.5	180	0.28	186	0.01
Cells under an optical microscope	197	–	156	–	224	–	213	–	65	–	190	0.5	189	0.03	214	0.74
MRI slice	55	–	51	–	56	–	44	–	2	–	23	0.5	29	0.87	33	0.01
Barcode	139	–	130	–	233	–	147	–	4	–	136	0.5	89	0.01	235	0.24

**Table 2** Processing time for the 10 images

Image	Otsu (s)	ROBP (s)	ROBVP (s)	Time ratio of ROBVP–ROBP (%)
Color	1.5	1.9	1.2	63.2
Material structure	1.5	2.7	1.6	59.3
SAR	1.6	3.1	1.6	51.6
Red blood cells	1.6	2.8	1.7	60.7
Rice grains	1.7	2.8	1.5	53.6
Number	1.7	3.7	1.7	45.9
Lymph	1.4	1.6	1.2	75.0
Cells under an optical microscope	1.5	2.2	1.6	72.7
MRI slice	1.7	2.5	1.4	56.0
Barcode	1.8	2.6	1.8	69.2



**Fig. 18** Denoising results of the color image

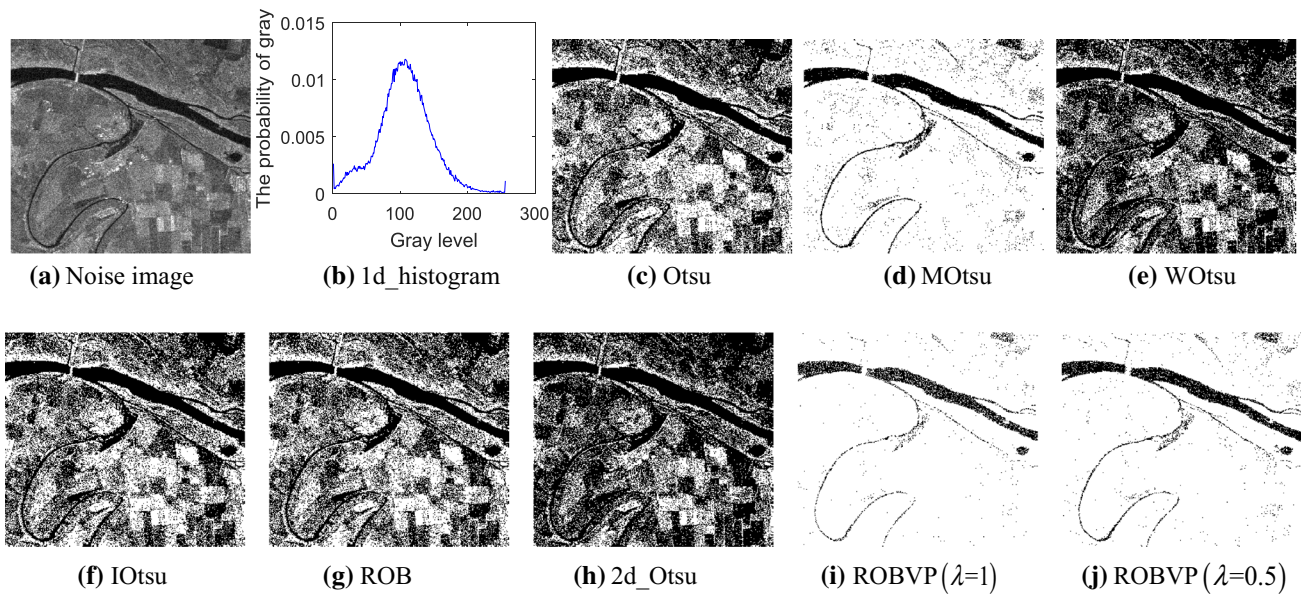


Fig. 19 Denoising results of the SAR image

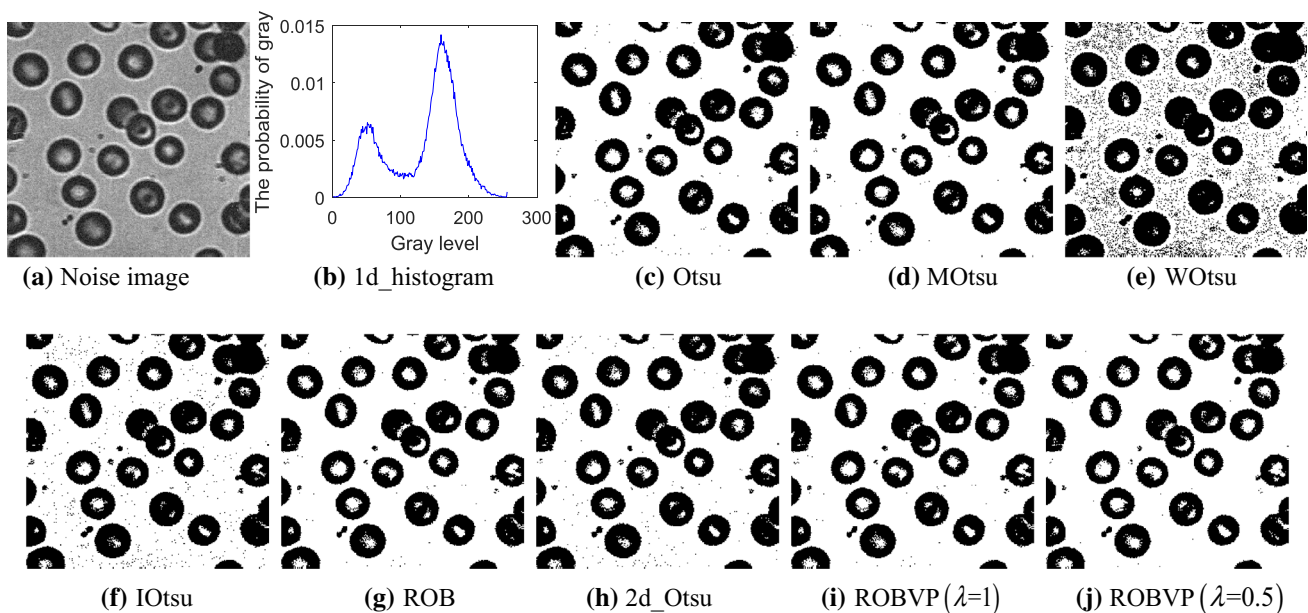


Fig. 20 Denoising results of the red blood cells image

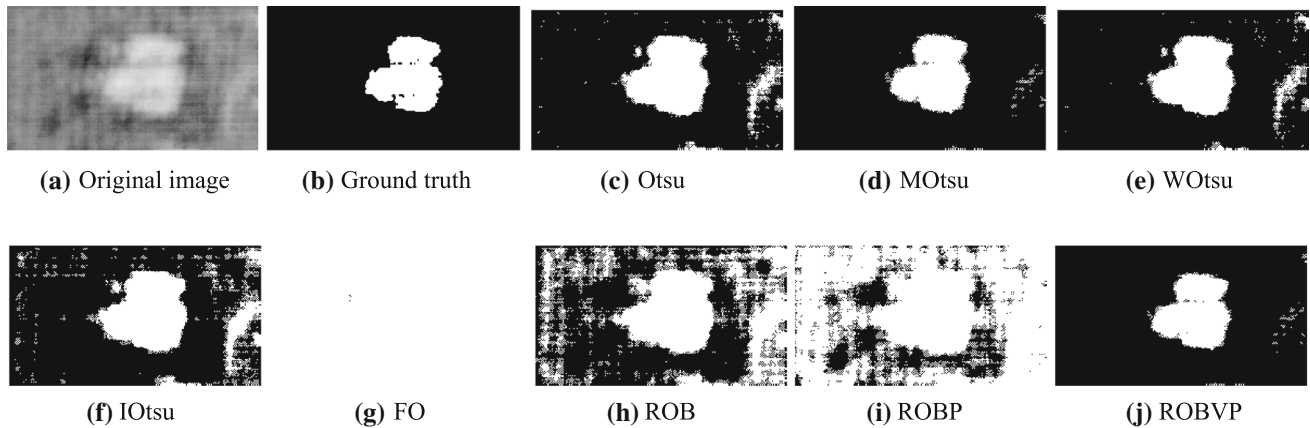
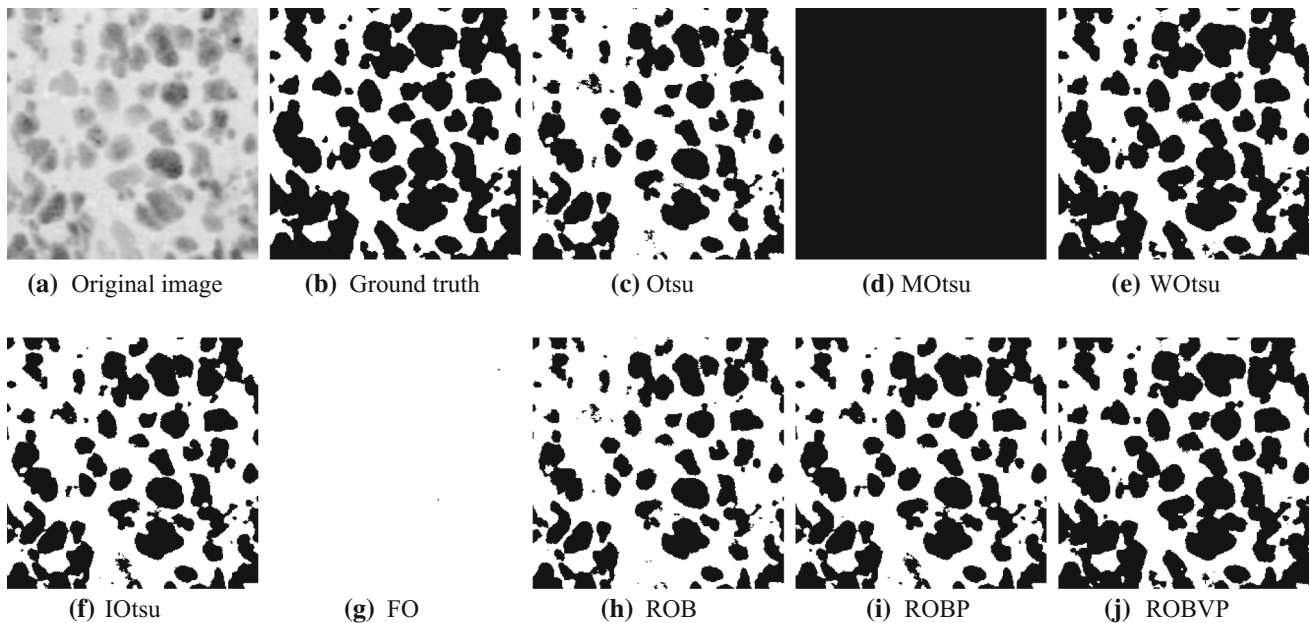
shown in Figs. 25 and 26, and the PSNR of eight methods is displayed in Table 5. We can see that the ROBVP ( $\lambda = 0.5$ ) method obtained maximum PSNR values in the eight methods, which implies that our proposed method has prominent performance for noisy images segmentation. The sizes of Test5 and Test6 image are  $105 \times 104$  and  $166 \times 60$  pixels, respectively.

### 7 Conclusion

In this paper, we have responded to the limitations of the Otsu’s method by focusing on other objects criteria and exploring image thresholding methods. We first developed the relative homogeneity measurement for images with larger distribution between-class variance. We then applied optimization strategies based on the gray distribution and region homogeneity characteristic information to propose a

**Table 3** Denoising threshold results of eight methods

Image	Otsu	MOtsu	WOtsu	IOtsu	ROB	2d_ Otsu	ROBVP ( $\lambda = 1$ )	ROBVP ( $\lambda = 0.5$ )
Color	125	135	138	116	129	119	112	41
SAR	106	54	120	107	107	122	36	43
Red blood cells	112	109	141	125	111	118	111	105

**Fig. 21** Results of the Test1 image**Fig. 22** Results of the Test2 image

new thresholding segmentation method, relative homogeneity method with optimized valley points and parameter  $\alpha$  (ROBVP). Furthermore, we have explored a multilevel segmentation discriminant using a single thresholding criterion.

Our work makes four main contributions. First, our method examines the relative homogeneity between the object and the background to offset the negative effects from the between-class distribution difference and to enable a more detailed and accurate image description.



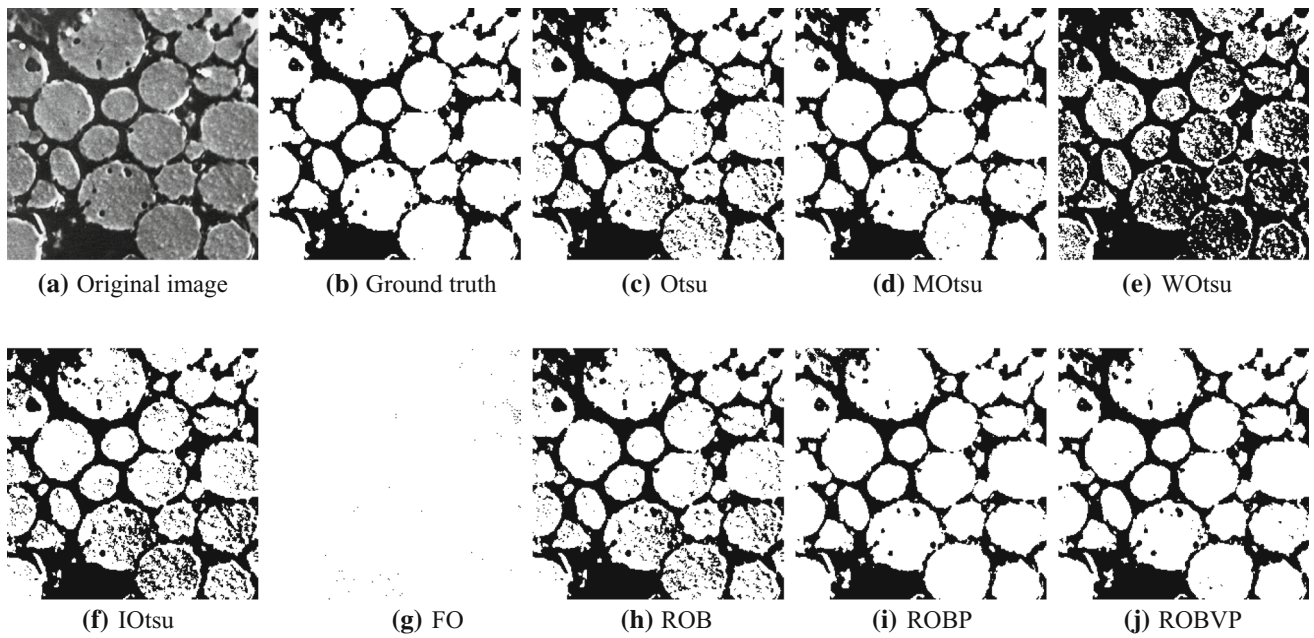


Fig. 23 Results of the Test3 image

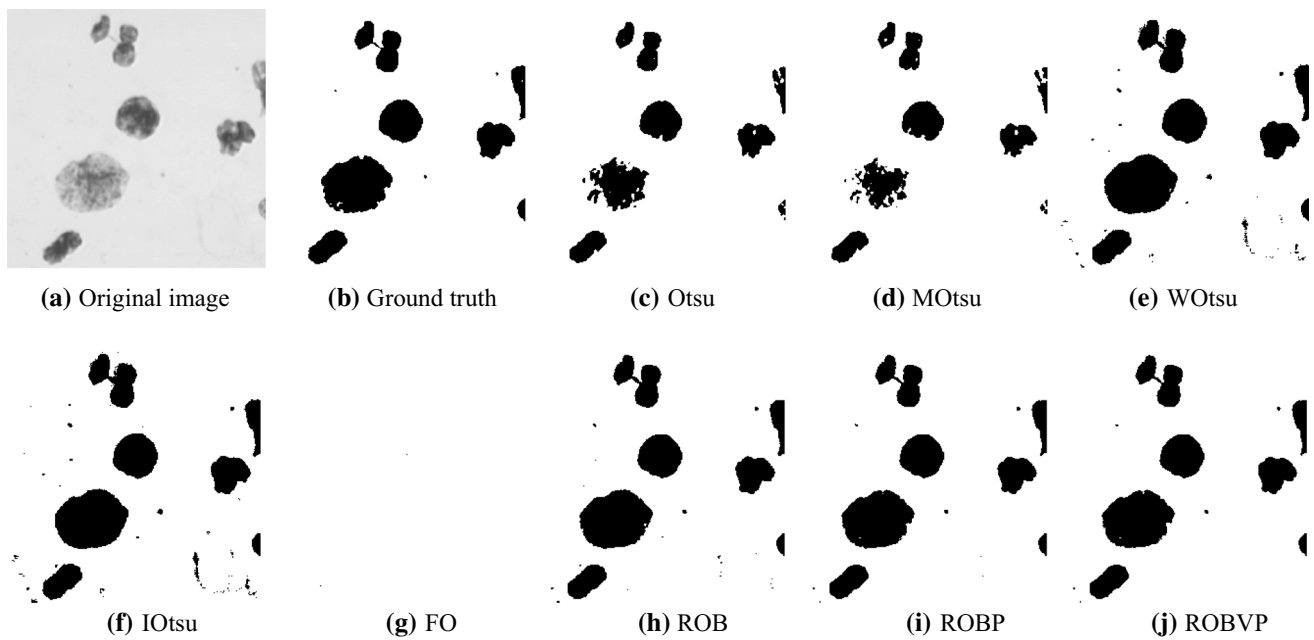


Fig. 24 Results of the Test4 image

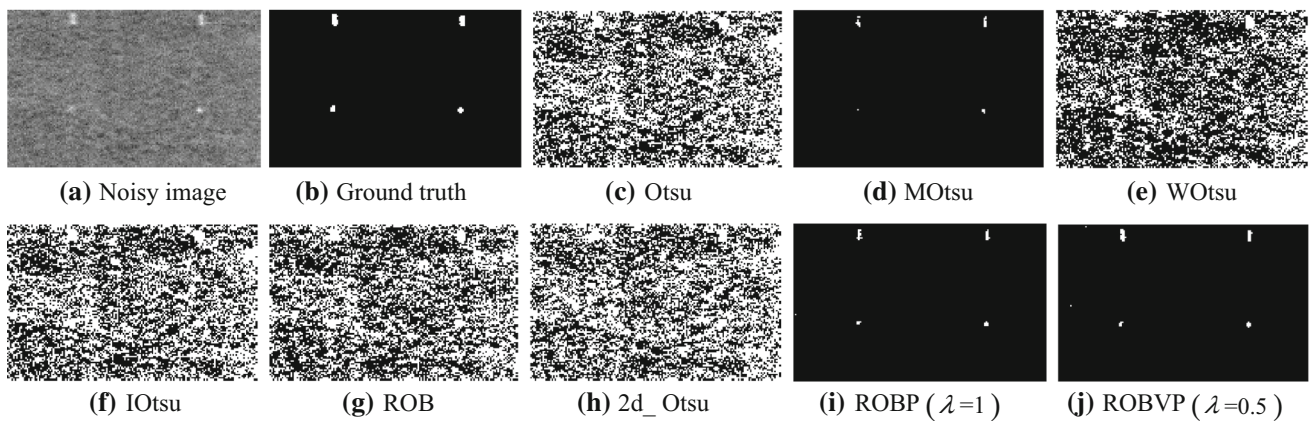
Second, our method uses the uniformity measure to determine the region with better homogeneity and to guide the selection of thresholding criteria. Third, our method improves the valley-points selection strategy by removing false threshold points and improving selection efficiency. Fourth, we have constructed an optimization function to

find the best value for parameter  $\alpha$  as part of obtaining the optimal threshold value.

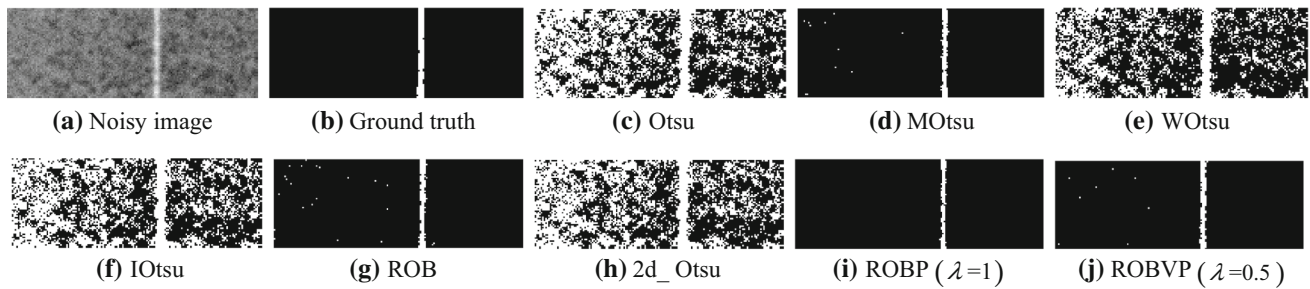
Our experimental results with real images having large between-class variances or skewed or tail-heavy gray histograms show that our proposed ROBVP method obtained the most accurate threshold value and yielded a lower misclassification error than existing methods. In addition,

**Table 4** The misclassification error (ME) of eight methods

Image	Otsu	MOtsu	WOtsu	IOtsu	FO	ROB	ROBP	ROBVP ( $\lambda = 0.5$ )
Test1								
Threshold	142	160	143	132	17	117	96	162
ME	0.0633	0.0183	0.0548	0.1173	0.8918	0.2927	0.6596	<b>0.0161</b>
Test2								
Threshold	167	254	184	171	38	163	171	184
ME	0.1501	0.4861	0.0176	0.1180	0.5155	0.1745	0.1102	<b>0.0176</b>
Test3								
Threshold	80	61	112	69	2	81	45	57
ME	0.0699	0.0183	0.3266	0.1190	0.2868	0.0789	0.0124	<b>0.0109</b>
Test4								
Threshold	182	172	221	222	70	219	216	215
ME	0.0294	0.0400	0.0153	0.0194	0.1259	0.0100	0.0066	<b>0.0054</b>



**Fig. 25** Results of the noisy Test5 image



**Fig. 26** Results of the noisy Test6 image

**Table 5** The PSNR value of eight methods

Image	Otsu	MOtsu	WOtsu	IOtsu	ROB	2d_ Otsu	ROBVP ( $\lambda = 1$ )	ROBVP ( $\lambda = 0.5$ )
Test5	3.1153	25.4432	4.4748	3.1153	3.4803	3.0812	27.0304	27.1600
Test6	3.4650	22.2011	5.1854	3.6854	22.2741	3.6854	20.9517	23.1702

our valley-points filter strategy shortened the time required to find a thresholding value, improving segmentation efficiency. Finally, when neighborhood image information was included, our proposed method offered superior denoising performance and higher PSNR values.

**Acknowledgements** The work is supported by the National Science Foundation of China (Nos. 61571361, 61671377), the Science Plan Foundation of the Education Bureau of Shaanxi Province (No. 15JK1682), and Scientific Research Climbing Project of Xiamen University of Technology, No. XPDKT18016.

## References

- Sima HF, Guo P, Zou YF, Wang ZH, Xu ML (2018) Bottom-up merging segmentation for color images with complex areas. *IEEE Trans Syst Man Cybern Syst* 48(3):354–365
- Oktay O, Ferrante E, Kamnitsas K, Heinrich M, Bai WJ, Caballero J, Cook SA, Marvao A, Dawes T, O'Regan DP, Kainz B, Glocker B, Rueckert D (2018) Anatomically Constrained Neural Networks (ACNNs): application to cardiac image enhancement and segmentation. *IEEE Trans Med Image* 37(2):384–395
- Wang T, Yang J, Ji ZX, Sun QS (2019) Probabilistic diffusion for interactive image segmentation. *IEEE Trans Image Process* 28(1):330–342
- Garcia-Lamont F, Cervantes J, López A, Rodriguez L (2018) Segmentation of images by color features: a survey. *Neurocomputing* 292(5):1–27
- Eltanboly A, Ghazal M, Hajjdiab H, Shalaby A, Switala A, Mahmoud A, Sahoo P, El-Azab M, El-Baz A (2019) Level sets-based image segmentation approach using statistical shape priors. *Appl Math Comput* 40(1):164–179
- Shi CF, Cheng YZ, Wang JK, Wang YD, Mori K, Tamura S (2017) Low-rank and sparse decomposition based shape model and probabilistic atlas for automatic pathological organ segmentation. *Med Image Anal* 38:30–49
- Tong T, Wolz R, Wang ZH, Gao QQ, Misawa K, Fujiwara M, Mori K, Hajnal JV, Rueckert D (2015) Discriminative dictionary learning for abdominal multi-organ segmentation. *Med Image Anal* 23:92–104
- Sahoo PK, Soltani S, Wong AKC (1988) A survey of thresholding techniques. *Comput Vis Graphics Image Process* 41(2):233–260
- Otsu N (1979) A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 9(1):62–66
- Goh TY, Basah SN, Yazid H, Safar MJA, Saad FSA (2018) Performance analysis of image thresholding: Otsu technique. *Measurement* 114:298–307
- Xue JH, Zhang YJ (2012) Ridler and Calvard's, Kittler and Illingworth's and Otsu's methods for image thresholding. *Pattern Recognit Lett* 33:793–797
- Xue JH, Titterton DM (2011) t-Test, F-tests and Otsu's methods for image thresholding. *IEEE Trans Image Process* 20:2392–2396
- Vala HJ, Baxi A (2013) A review on Otsu image segmentation algorithm. *Int J Adv Res Comput Eng Technol* 2(2):387–389
- Zou Y, Dong F, Lei B, Sun S, Jiang T, Chen P (2014) Maximum similarity thresholding. *Digit Signal Proc* 28:120–135
- Fan JL, Lei B (2012) A modified valley-emphasis method for automatic thresholding. *Pattern Recognit Lett* 33(6):703–708
- Yuan XC, Wu LS, Peng QJ (2015) An improved Otsu method using the weighted object variance for defect detection. *Appl Surf Sci* 349(15):472–484
- Suheir M, Harb E, Isa NAM, Salamah SA (2015) Improved image magnification algorithm based on Otsu thresholding. *Comput Electr Eng* 46:338–355
- Farrahi Moghaddam R, Cheriet M (2012) AdOtsu: an adaptive and parameterless generalization of Otsu's method for document image binarization. *Pattern Recognit* 45(6):2419–2431
- Lai YK, Rosin PL (2014) Efficient circular thresholding. *IEEE Trans Image Process* 23(3):992–1001
- Cai H, Yang Z, Cao X, Xia W, Xu X (2014) A new iterative triclass thresholding technique in image segmentation. *IEEE Trans Image Process* 23(3):1038–1046
- Feng YC, Zhao HY, Li XF, Zhang XL, Li HP (2017) A multi-scale 3D Otsu thresholding algorithm for medical image segmentation. *Digit Signal Proc* 60:186–199
- He S, Schomaker L (2019) DeepOtsu: document enhancement and binarization using iterative deep learning. *Pattern Recognit* 91:379–390
- Manikandan S, Ramar K, Willjuice Iruthayarajan M, Srinivasagan KG (2014) Multilevel thresholding for segmentation of medical brain images using real coded genetic algorithm. *Measurement* 47:558–568
- Bhandari AK, Kumar A, Singh GK (2015) Modified artificial bee colony based computationally efficient multilevel thresholding for satellite image segmentation using Kapur's, Otsu and Tsallis functions. *Expert Syst Appl* 42(3):1573–1601
- Zhao F, Liu HQ, Fan JL, Chen CW, Lan R, Li N (2018) Intuitionistic fuzzy set approach to multi-objective evolutionary clustering with multiple spatial information for image segmentation. *Neurocomputing* 312:296–309
- Chen SC, Li DH (2006) Image binarization focusing on objects. *Neurocomputing* 69(16–18):2411–2415
- Zhang H, Hu WY (2018) A modified thresholding method based on relative homogeneity. *J Inf Hiding Multimed Signal Process* 9(2):285–292
- Sezgin M, Sankur B (2004) Survey over image thresholding techniques and quantitative performance evaluation. *J Electron Imaging* 13(1):146–168
- Sezgin M, Sankur B (2003) Image multithresholding based on sample moment function. In: *Proceeding of the 2003 IEEE international conference on image processing*, vol 9. Barcelona, Spain, pp 415–418
- Sezgin M (2017) blt\_image\_references. <http://mehmetsezgin.net>. Accessed 10 Nov 2017
- Akay B (2013) A study on particle swarm optimization and artificial bee colony algorithms for multilevel thresholding. *Appl Soft Comput* 13:3066–3091

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.