

Received August 6, 2020, accepted September 12, 2020, date of current version September 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3024718

Automatic Image Thresholding Based on Shannon Entropy Difference and Dynamic Synergic Entropy

YAOBIN ZOU^{ID}, JINYU ZHANG^{ID}, MANISH UPADHYAY^{ID}, SHUIFA SUN^{ID}, AND TINGYAO JIANG^{ID}

College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China

Corresponding author: Tingyao Jiang (jiangtyctgu@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61871258 and Grant 61502274 and in part by the National Key Research and Development Program of China under Grant 2016YFB0800403.

ABSTRACT An automatic thresholding method based on Shannon entropy difference and dynamic synergic entropy is proposed to select a reasonable threshold from the gray level image with a unimodal, bimodal, multimodal, or peakless gray level histogram. Firstly, a new concept called Shannon entropy difference is proposed, and the stopping condition of a multi-scale multiplication transformation is automatically controlled by maximizing Shannon entropy difference to produce edge images. Secondly, the gray level image is thresholded by the gray levels in order from smallest to largest to generate a series of binary images, followed by extracting contour images from the binary images. Then, a series of gray level histograms that can dynamically reflect gray level distributions and pixel positions are constructed using the edge images and the contour images synergically. Finally, dynamic synergic Shannon entropy is calculated from this series of gray level histograms, and the threshold corresponding to maximum dynamic synergic entropy is taken as the final segmentation threshold. The experimental results on 40 synthetic images and 50 real-world images show that, although the proposed method is not superior to 8 automatic segmentation methods in computational efficiency, it has more flexible adaptivity of selecting threshold and better segmentation accuracy.

INDEX TERMS Automatic thresholding, principle of maximum entropy, Shannon entropy difference, dynamic synergic entropy.

I. INTRODUCTION

Image segmentation is one of the most fundamental, useful, and studied topics in image processing and analysis. The goal is a partition of the image into coherent regions, which is an important initial step in the analysis of image content. Numerous image segmentation algorithms have been developed in the last several decades, from the earliest methods, such as image thresholding [1], region growing and merging [2], [3], clustering [4], [5], watershed segmentation [6], [7], to more complex algorithms, such as active contours [8], graph cuts [9], [10], and deep learning-based methods [11], [12].

Among these methods, image thresholding is a simple, yet effective, way of separating targets from the background, when the gray levels of the pixels belonging to targets are substantially different from the gray levels of the pixels

belonging to the background [1], [13]–[15]. Image thresholding is also one of the most commonly used low-level image processing methods in various image analysis systems [1], [16]–[19]. Image thresholding compares the gray level of each pixel in a gray level image with a selected threshold to determine whether the pixel belongs to the targets or the background. Selecting an appropriate threshold thus becomes a key step for accurate thresholding segmentation, and the core goal of image thresholding is to make the selected threshold as close as possible to the expected optimal threshold [20].

In terms of automatically selecting threshold, one representative idea is to apply the principle of maximum entropy [21] in information theory to select the segmentation threshold [22], [23]. The types of entropy involved in the thresholding methods derived from this idea are mainly Shannon entropy, Rényi entropy, Tsallis entropy, Arimoto entropy, and Masi entropy [24]. The basis for calculating the above various entropies is first to obtain the gray level distribution, or the gray level histogram, from a gray level image. According to

The associate editor coordinating the review of this manuscript and approving it for publication was Jiachen Yang^{ID}.

the dimensions of gray level histogram and the spatial position relationship between pixels involved in constructing the gray level histogram, the gray level histogram can be divided into two categories: global one-dimensional histogram and local two-dimensional histogram. The change from global one-dimensional histogram to local two-dimensional histogram reflects a main idea in the development of thresholding methods based on the principle of maximum entropy: considering both gray level distributions and pixel positions.

The global one-dimensional gray level histogram can be established by counting the occurrence frequency of each gray level in a gray level image. A thresholding method based on the principle of maximum entropy with the global one-dimensional gray level histogram is usually called a global entropy method [25]–[30]. The global entropy method can be traced back to maximum Shannon entropy (MSE) method [25] proposed by Kapur *et al.* The core idea of MSE method is that when the sum of the background Shannon entropy and the target Shannon entropy takes a maximum value, the corresponding gray level in that situation is selected as a segmentation threshold. The theoretical premise of applying the MSE method is that Shannon entropy satisfies the additivity principle for statistics independent subsystems. When the gray level distributions of the background and the target are independent of each other, for example, the gray level distributions of the background and the target are both a uniform distribution and their distributions do not overlap, the MSE method can obtain a theoretical optimal segmentation threshold [31]. However, many real-world images, affected by factors such as noise and low pass filter, usually contain non-extensive information that is long-range correlation or long-term memory [30]. Shannon entropy cannot effectively express non-extensive information. Thus, it is difficult for the MSE method to select a reasonable segmentation threshold from those real-world images.

Many subsequent methods, such as maximum Rényi entropy method [26], maximum Tsallis entropy method [27], maximum Arimoto entropy method [28], and maximum Masi entropy method [30], continue this line of thinking similar to the MSE method, but they adopt Rényi entropy, Tsallis entropy, Arimoto entropy, or Masi entropy in terms of entropy model. For statistics independent subsystems, Rényi entropy, Tsallis entropy, and Arimoto entropy have the ability to describe non-extensive information, while Masi entropy can describe both extensive and non-extensive information [30]. Theoretically, the maximum Rényi, Tsallis, Arimoto, or Masi entropy method has the potential to improve the MSE method. However, Rényi entropy, Tsallis entropy, Arimoto entropy, and Masi entropy are all the entropy with parameters, and they just utilize additional parameters to characterize the extensibility and/or non-extensibility of random system. In practice, segmentation results are often sensitive to the used parameters when these entropies with parameters are applied to image thresholding [30]. This means that before a reasonable threshold can be automatically selected, an appropriate entropy parameter should be first automatically

evaluated [29]. However, this is still currently an open issue. If an empirical fixed parameter or manual adjustment parameter is adopted, it will inevitably limit the automatic adaptability of these methods based on the principle of maximum entropy.

A thresholding method based on the principle of maximum entropy with the local two-dimensional gray level histogram is usually called a local entropy method [22], [32]–[36]. Local entropy methods realize a common problem of global entropy methods: the spatial correlation between pixels is not considered when the one-dimensional gray level histogram is constructed. For a specific global entropy method, as long as the one-dimensional gray level histograms of the gray level images are the same, the method will obtain a same segmentation threshold even if the contents of the gray level images are different. Local entropy methods attempt to alleviate the deficiencies of global entropy methods in the description of image content by constructing two-dimensional gray level histograms. Local entropy methods mainly adopt two ways to construct two-dimensional gray level histograms. The first way is to use an original gray level image and its local mean image [23], [32], [33], [35]. The second way is to construct a gray level co-occurrence matrix from an original gray level image [22], [34], [36]. Although it is possible to construct various gray level co-occurrence matrices with different directions and pixel distances, it mainly uses simple four-neighbor relationship of pixel pairs to build the gray level co-occurrence matrix because of the difficulty in automatically determining the most reasonable direction and pixel distance for a given gray level image.

While extending global entropy methods, local entropy methods implicitly or explicitly inherit three shortcomings of global entropy methods. First, although local entropy methods consider the spatial correlation between pixels, it limits the correlation to a small range of local neighborhoods. Second, if local entropy methods adopt Shannon entropy, the non-extensive information in a two-dimensional gray level histogram still acts as an obstacle to evaluating segmentation thresholds; if Rényi entropy, Tsallis entropy, Arimoto entropy or Masi entropy is adopted, local entropy methods will also face the problem of automatically evaluating the corresponding entropy parameters. Third, local entropy methods extend the dimension of the gray level histogram, thus there are more choices when applying the principle of maximum entropy. However, local entropy methods no longer consider the original gray level image after establishing two-dimensional gray level histograms, and this cuts off the association between the original image and the segmentation threshold.

In addition to the inherent shortcomings of global entropy methods and local entropy methods, the complex and variable targets and backgrounds also objectively increase the difficulty of automatically selecting a reasonable segmentation threshold. Affected by factors such as random noise, low pass filter, as well as the size of target and background, the gray level histogram of an image may be unimodal, bimodal,

multimodal, or even peakless. When the basic distribution constituting a gray level histogram is a non-Gaussian distribution, such as a gamma distribution, an extreme value distribution, a Rayleigh distribution, a uniform distribution or a beta distribution, it remains challenging to automatically select the segmentation threshold that is as reasonable as possible.

To overcome the shortcomings of global entropy methods and local entropy methods, and also to deal with automatic threshold selection under the above different gray level distributions in a unified framework, this article proposes an automatic thresholding method guided by maximizing dynamic synergic entropy (MDSE). The MDSE method constructs dynamically a series of one-dimensional gray level histograms using synergically an invariant guiding edge image and a series of changing contour images. The guiding edge image is produced by performing a multi-scale multiplication transformation on an original gray level image, where the stopping condition of the multi-scale multiplication transformation is automatically controlled by maximizing Shannon entropy difference. The changing contour images are generated by continuously thresholding the original gray level image with every possible gray level, and then extracting internal and external contours from these binary images. The one-dimensional gray level histograms are synergically constructed from the guiding edge image and the changing contour images, which considers both gray level distributions and pixel positions. Thus, the calculated Shannon entropy (dynamic synergic entropy, DSE) based on this series of one-dimensional gray level histograms lays a foundation for applying the principle of maximum entropy to automatically select a reasonable threshold.

Some main contributions of this study can be summarized as follows: ① In terms of extending the theory of Shannon entropy, a new concept called Shannon entropy difference is proposed, and it is used for automatically controlling the stopping condition of a multi-scale multiplication transformation. ② Based on the analysis of the monotonicity of DSE, a new objective function is proposed to automatically select appropriate segmentation thresholds from the gray level images with different gray level distributions. ③ A novel image thresholding method called MDSE is proposed, which has better segmentation adaptability and robustness than other image thresholding methods. ④ 40 synthetic images and 50 real-world images with different gray level distributions are generated or collected, and their corresponding segmentation reference images are also provided. All these images are shared online.

The rest of this article is organized as follows: Section II focuses on a new concept, i.e., Shannon entropy difference, and its application. In particular, Section II.A defines the concept of Shannon entropy difference, and discusses the relationship between Shannon entropy difference and gray level histogram; Section II.B analyzes a technique called multi-scale multiplication transformation and proposes a criterion maximizing Shannon entropy difference to automatically stop computing multi-scale multiplication.

Section III proposes a new criterion of selecting threshold based on DSE, and analyze the rationality of calculating the final threshold according to the new criterion. Section IV describes the corresponding algorithm steps. Section V analyzes and discusses the experimental results of the proposed MDSE method and 9 compared methods on 40 synthetic images and 50 real-world images. Finally, Section VI draws several conclusions and describes future works.

II. SHANNON ENTROPY DIFFERENCE AND ITS APPLICATION

A. DEFINITION AND ANALYSIS OF SHANNON ENTROPY DIFFERENCE

Given a gray level histogram of a gray level image, let a gray level l divide this histogram into two parts (see Fig. 1). Suppose that there are m gray levels on the left part, the discrete probability distribution of these gray levels is $q_i (1 \leq i \leq m)$ and $\sum_{i=1}^m q_i = 1$, and the total frequency on the left part is L . Suppose that there are n gray levels on the right part, the discrete probability distribution of these gray levels is $p_i (1 \leq i \leq n)$ and $\sum_{i=1}^n p_i = 1$, and the total frequency on the right part is R .

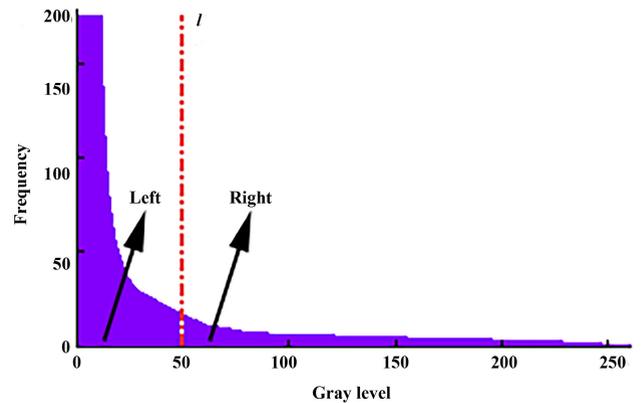


FIGURE 1. Schematic diagram of the left division and the right division of gray level histogram.

Let H^{Left} and H^{Right} denote the Shannon entropy corresponding to the left and the right parts, respectively. According to the definition of Shannon entropy, the following equations can be given:

$$H^{Left} = - \sum_{i=1}^m q_i \log_2 q_i \tag{1}$$

$$H^{Right} = - \sum_{i=1}^n p_i \log_2 p_i \tag{2}$$

Let $H^{Left \cup Right}$ indicate the Shannon entropy corresponding to the whole histogram, and its expression is

$$H^{Left \cup Right} = - \sum_{i=1}^{m+n} o_i \log_2 o_i \tag{3}$$

where

$$o_i = \begin{cases} \frac{L}{L+R}q_i & (1 \leq i \leq m) \\ \frac{R}{L+R}p_i & (m+1 \leq i \leq m+n) \end{cases} \quad (4)$$

Further, we define $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ as Shannon entropy difference. The next 3 propositions will be used for proving and analyzing: to increase the Shannon entropy difference, what kind of distribution characteristics the gray level histogram in Fig. 1 should have.

Proposition 1: $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}} = s(H^{\text{Right}} - H^{\text{Left}}) + s\log_2 s + (1-s)\log_2(1-s)$, where $s = \frac{L}{L+R}$.

Proof:

$$\begin{aligned} H^{\text{Left} \cup \text{Right}} &= -\sum_{i=1}^{m+n} o_i \log_2 o_i \\ &= -\sum_{i=1}^m o_i \log_2 o_i - \sum_{i=m+1}^{m+n} o_i \log_2 o_i \\ &= -\sum_{i=1}^m \left(\frac{L}{L+R}q_i\right) \log_2 \left(\frac{L}{L+R}q_i\right) - \sum_{i=1}^n \left(\frac{R}{L+R}p_i\right) \\ &\quad \times \log_2 \left(\frac{R}{L+R}p_i\right) \\ &= -\sum_{i=1}^m \left(\frac{L}{L+R}q_i\right) \log_2 \left(\frac{L}{L+R}\right) - \sum_{i=1}^m \left(\frac{L}{L+R}q_i\right) \log_2 q_i \\ &\quad - \sum_{i=1}^n \left(\frac{R}{L+R}p_i\right) \log_2 \left(\frac{R}{L+R}\right) - \sum_{i=1}^n \left(\frac{R}{L+R}p_i\right) \log_2 p_i \\ &= -\left(\frac{L}{L+R}\right) \log_2 \left(\frac{L}{L+R}\right) - \left(\frac{L}{L+R}\right) \sum_{i=1}^m q_i \log_2 q_i \\ &\quad - \left(\frac{R}{L+R}\right) \log_2 \left(\frac{R}{L+R}\right) - \left(\frac{R}{L+R}\right) \sum_{i=1}^n p_i \log_2 p_i \\ &= -s\log_2 s + sH^{\text{Left}} - (1-s)\log_2(1-s) + (1-s)H^{\text{Right}} \end{aligned}$$

Then we have

$$H^{\text{Right}} - H^{\text{Left} \cup \text{Right}} = s(H^{\text{Right}} - H^{\text{Left}}) + s\log_2 s + (1-s)\log_2(1-s).$$

Proposition 2: When $s > \frac{1}{1+2(H^{\text{Right}} - H^{\text{Left}})}$, $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ is a monotonically increasing function about s .

Proof: The variable s is discontinuous, so we cannot directly find the derivative of $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ with respect to s . For this problem, we can extend the domain of s from the original discontinuous point set to the continuous interval $(0, 1)$, then we can obtain the derivative of Shannon entropy difference $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ with respect to s :

$$\begin{aligned} H^{\text{Right}} - H^{\text{Left}} + \log_2 s + s \times \frac{1}{s} \times \log_2 e - \log_2(1-s) \\ + (1-s) \times \frac{1}{1-s} \times (-1) \times \log_2 e. \end{aligned}$$

After simplifying it, we have

$$H^{\text{Right}} - H^{\text{Left}} + \log_2 s - \log_2(1-s)$$

Further, let the above expression be greater than 0, then we get the solution $s > \frac{1}{1+2(H^{\text{Right}} - H^{\text{Left}})}$.

According to the relationship between the derivative and the monotonicity of a function, we can infer that when $s > \frac{1}{1+2(H^{\text{Right}} - H^{\text{Left}})}$, $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ is a monotonically increasing function about $s \in (0, 1)$.

Proposition 3: When $H^{\text{Right}} > H^{\text{Left}}$, $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ is a monotonically increasing function about $H^{\text{Right}} - H^{\text{Left}}$.

Proof: According to **Proposition 1**, the derivative of $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ with respect to $H^{\text{Right}} - H^{\text{Left}}$ is s , and it is obvious that s is greater than 0. According to the relationship between the derivative and the monotonicity of a function, we can infer that $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ is a monotonically increasing function about $H^{\text{Right}} - H^{\text{Left}}$.

Propositions 1, 2 and 3 show that when $H^{\text{Right}} > H^{\text{Left}}$ and $s > \frac{1}{1+2(H^{\text{Right}} - H^{\text{Left}})}$, the greater s and $H^{\text{Right}} - H^{\text{Left}}$, the greater the Shannon entropy difference $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$. Further, it is easy to infer that a relatively greater L or a relatively smaller R , and a relatively greater H^{Right} or a relatively smaller H^{Left} will make the Shannon entropy difference $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ take a relatively greater value. In other words, the Shannon entropy H^{Right} corresponding to the right histogram composed of relatively few pixels should be as great as possible, and the Shannon entropy H^{Left} corresponding to the left histogram composed of relatively many pixels should be as small as possible, which will make the Shannon entropy difference $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ take a relatively greater value. Further, according to the property that more uniform the probability distribution tends to, the greater the Shannon entropy becomes [37], a kind of gray level histogram that makes the Shannon entropy difference $H^{\text{Right}} - H^{\text{Left} \cup \text{Right}}$ tend to take a relatively greater value is as follows: the right histogram composed of relatively few pixels should be distributed as uniformly as possible over a wider range of gray levels, and the left histogram composed of relatively many pixels should be distributed as narrowly as possible over a narrower range of gray levels. The next Section II.B will elaborate on how to obtain the gray level histogram with the above distribution characteristics.

B. GUIDING EDGE IMAGE BASED ON MAXIMIZING SHANNON ENTROPY DIFFERENCE

Let the symbol f denote a gray level image, and let the symbol ∇g_x denote the partial derivative of two-dimensional Gaussian function $g(x, y; \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$ with respect to x . Let the symbol $\kappa_x(\sigma)$ denote the absolute value of the output image resulting from the convolution of the image f with the kernel ∇g_x :

$$\kappa_x(\sigma) = |f * \nabla g_x| \quad (5)$$

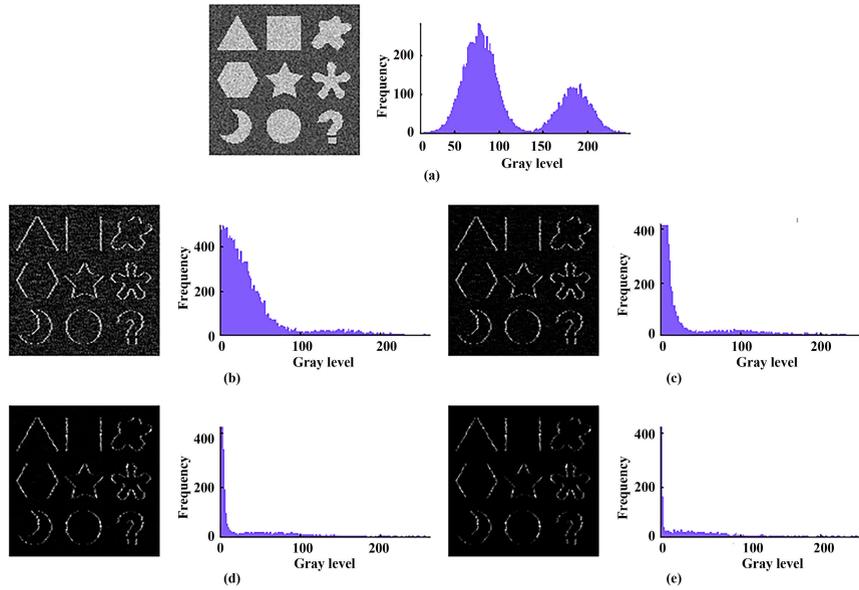


FIGURE 2. The changing trend of the gray level histogram of the image K_{u_x} with gradually increasing value u_x . (a) a gray level image f and its gray level histogram; (b)-(e) show the corresponding image K_{u_x} and its gray level histogram when u_x is 1, 2, 3, and 4, respectively. To clearly show the frequency of gray level in the interval $[0,255]$, the gray level histograms in (c)-(e) are truncated at the frequency around 400.

The image $\kappa_x(\sigma)$ is closely related to convolution scale σ , and different σ will produce different image $\kappa_x(\sigma)$.

Define the multi-scale multiplication transformation of the image f in the x -axis direction as the product of u_x different images $\kappa_x(\sigma_i)$:

$$K_{u_x} = \prod_{i=1}^{u_x} \kappa_x(\sigma_i) \quad (6)$$

According to the sampling theory based on the Gaussian kernel function, convolving an image with a $(8\sigma_i + 1) \times (8\sigma_i + 1)$ Gaussian kernel can ensure that the convolution result sufficiently approximates the result obtained by convolving the image with the full Gaussian distribution [38]. In addition, for the convolution operation on digital images, the neighborhood defined by a convolution kernel is often centered on a pixel and the neighborhood size is usually odd [39], such as $3 \times 3, 5 \times 5, 7 \times 7$, etc. Thus, each convolution scale can be calculated in turn as $\sigma_i = 0.25 \times i (i \geq 1)$.

In Eq. (5), the existence of Gaussian function makes edge signals and noises (or random details) have different response characteristics to ∇g_x . When the convolution scale σ gradually increases, the responses of noises (or random details) rapidly decrease, while the responses of edge signals show the following characteristics: ① the response at the center of the edge remains relatively good; ② the response gradually decreases as the pixel position gets farther and farther from the center of the edge.

Under the premise that the gray level of the image K_{u_x} is normalized to $[0, 255]$, as the number u_x of images participating in multi-scale multiplication transformation gradually increases, the product of the response values of noises

(or random details) will gradually approach the minimum value 0, and the product of the response values of the edge signals will be spread between 0 and 255 (see the left images in Figs. 2(b)-(e)). As u_x gradually increases, the gray level histogram of the image K_{u_x} will show the following characteristics or tendencies (see the right images in Figs. 2(b)-(e)): ① the gray level histogram shows a heavy right tail distribution. ② the mode of the gray level histogram gradually shifts to the left, and the mode will be equal to 0 when u_x is large enough. ③ if the value u_x is continuously increased so that it exceeds a certain critical value, the multiplicative effect will cause more response values of edge signals to gather near 0. As a result, the gray level distribution between 0 and 255 becomes increasingly sparse, and the number of gray levels corresponding to the frequency 0 is gradually increasing.

The relationship between the gray level histogram of the image K_{u_x} and the value u_x indicates that a relatively appropriate value u_x needs to be found so that the gray level histogram of the image K_{u_x} is more consistent with the gray level distribution characteristics expected in Section II.A. To this end, a natural choice is to maximize the Shannon entropy difference of the gray level histogram of the image K_{u_x} , which can be formalized as:

$$(u_x^*, l_x^*) = \arg \max_{u_x \in \mathbb{Z}^+} \max_{l_x \in [0, 255]} (H_x^{\text{Right}} - H_x^{\text{Left} \cup \text{Right}}) \quad (7)$$

Once the number u_x^* of images participating in the multi-scale multiplication transformation is determined, the corresponding image $K_{u_x^*}$ can be calculated by combining the Eq. (5) and (6). Here the image $K_{u_x^*}$ will be called the guiding edge image in the direction of x -axis.

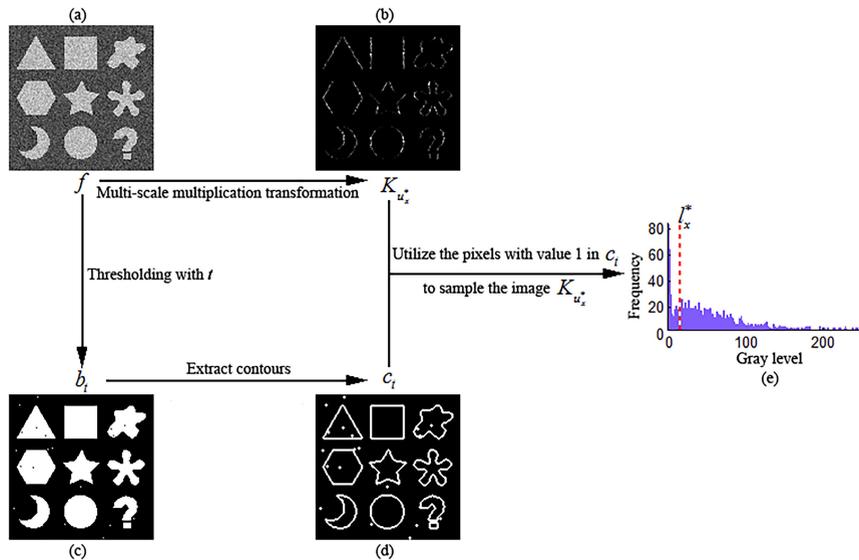


FIGURE 3. Schematic diagram of core concepts and key steps in the MDSE method. (a) a gray level image f ; (b) a guiding edge image $K_{u_x^*}$ obtained by performing multi-scale multiplication transformation on the gray level image f ; (c) a binary image b_t obtained by thresholding the gray level image f with a gray level t ; (d) a contour image c_t obtained by extracting the inner and outer contours from the binary image b_t ; (e) a gray level histogram constructed by sampling the image $K_{u_x^*}$ with the pixels taking value 1 in the image c_t , to clearly show the frequency of gray level in the interval $[0,255]$, the gray level histogram is truncated at the frequency near 80.

The above analysis and reasoning to the image f in the direction of x -axis are also applicable to the case in the direction of y -axis. Thus, for the image K_{u_y} in the direction of y -axis, the relatively appropriate number u_y^* of images participating in multi-scale multiplication transformation can be calculated according to the following Eq. (8):

$$(u_y^*, l_y^*) = \arg \max_{u_y \in Z^+} \max_{l_y \in [0,255]} \left(H_y^{\text{Right}} - H_y^{\text{Left} \cup \text{Right}} \right) \quad (8)$$

Similarly, the image $K_{u_y^*}$ will be called the guiding edge in the direction of y -axis.

III. CRITERION OF SELECTING THRESHOLD BASED ON DSE

A binary image b_t is produced by thresholding a gray level image f with a gray level t (see Fig. 3(c)), and a contour image c_t is extracted from the binary image b_t (see Fig. 3(d)). The pixels with the value 1 in the image c_t are further utilized to sample the guiding edge images $K_{u_x^*}$ and $K_{u_y^*}$ calculated in Section II.B, and then two new gray level histograms Υ_x and Υ_y are reconstructed from the sampled pixels. Figs. 3(b), (d), and (e) show a visual example of constructing a gray level histogram Υ_x .

The MDSE method proposes the following objective function to select the final threshold t^* :

$$t^* = \arg \max_{t \in [t_{\min}, t_{\max}]} \left(\left(H_x^{\text{Left} \cup \text{Right}_t} + H_y^{\text{Left} \cup \text{Right}_t} \right) / 2 \right) \quad (9)$$

where t_{\min} and t_{\max} represent the minimum and maximum gray level in the gray level image f , and $H_x^{\text{Left} \cup \text{Right}_t}$ and $H_y^{\text{Left} \cup \text{Right}_t}$ are Shannon entropy calculated from the gray level histogram Υ_x and Υ_y , respectively.

When the gray level t gradually varies from t_{\min} to t_{\max} , the contour image c_t changes continuously, and consequently the corresponding gray level histogram Υ_x and Υ_y also change continuously, so that the Shannon entropy $H_x^{\text{Left} \cup \text{Right}_t}$ and $H_y^{\text{Left} \cup \text{Right}_t}$ also change dynamically. Moreover, the gray level histograms Υ_x and Υ_y are synergically constructed from the guiding edge image and the contour image. Consequently, Υ_x and Υ_y consider both gray level distributions and pixel positions and reflect the gray level statistical characteristics of different local regions distributed in the original image. Since Shannon entropy is computed from Υ_x and Υ_y , this allows $H_x^{\text{Left} \cup \text{Right}_t}$ and $H_y^{\text{Left} \cup \text{Right}_t}$ to dynamically and synergically reflect the spatial relationship between the pixels that constitute targets and the background. The Shannon entropy calculated in this way is hereinafter called DSE (dynamic synergic entropy).

Here we further analyze the rationality of calculating the final threshold t^* according to Eq. (9). Because the analysis of $H_x^{\text{Left} \cup \text{Right}_t}$ and $H_y^{\text{Left} \cup \text{Right}_t}$ is similar, here we will only take $H_x^{\text{Left} \cup \text{Right}_t}$ for an example. There are two ways to calculate the DSE $H_x^{\text{Left} \cup \text{Right}_t}$ corresponding to the gray level histogram Υ_x . One is to do the calculation based on the entire Υ_x , the other is to divide the gray level histogram Υ_x into left and right parts with the gray level l_x^* calculated by Eq. (7), and compute Shannon entropy in the same way as $H^{\text{Left} \cup \text{Right}}$ in Section II.A. The first method is more direct and more efficient, and it will be used in the specific implementation of the proposed algorithm. The second method can construct a multivariate function consisting of 3 independent variables, which facilitates theoretical analysis of the rationale for the

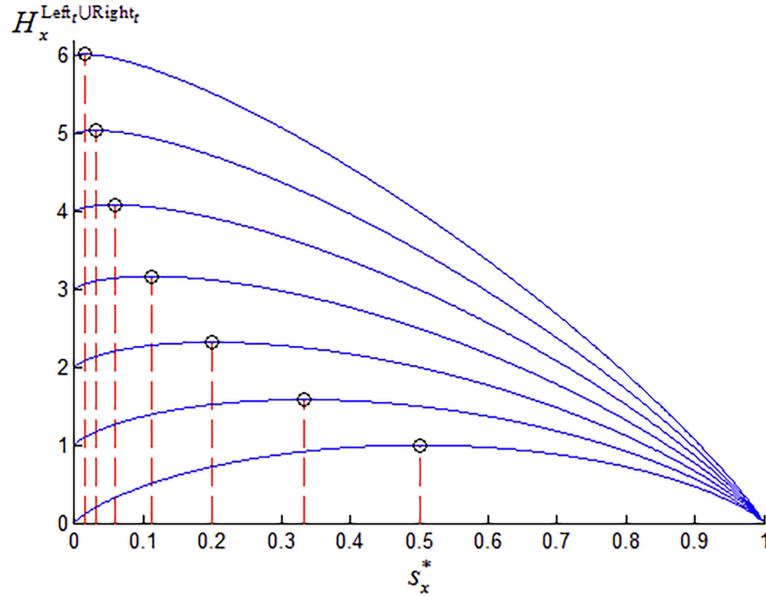


FIGURE 4. The family of curves of the DSE $H_x^{Left_t \cup Right_t}$, where the circle on each curve represents the maximum value of that curve, and the intersections of the red dashed lines and the horizontal axis indicate the corresponding s_x^* . For each curve drawn from top to bottom, the corresponding value $H_x^{Right_t}$ is 6, 5, 4, 3, 2, 1, and 0, while $H_x^{Left_t}$ remains 0.

threshold t taking a reasonable value. According to the second method, we can calculate $H_x^{Left_t \cup Right_t}$ as follows:

$$H_x^{Left_t \cup Right_t} = (1 - s_x^*)H_x^{Right_t} + s_x^*H_x^{Left_t} - s_x^* \log_2 s_x^* - (1 - s_x^*) \log_2 (1 - s_x^*) \quad (10)$$

where $H_x^{Right_t}$ denotes the Shannon entropy corresponding to the right part in the gray level histogram Υ_x greater than the gray level l_x^* , $H_x^{Left_t}$ denotes the Shannon entropy corresponding to the left part in the gray level histogram Υ_x less than the gray level l_x^* , and $s_x^* = \frac{L_x^*}{L_x^* + R_x^*}$, where L_x^* and R_x^* are the number of pixels corresponding to the left and right part in the gray level histogram Υ_x , respectively.

Proposition 4: Let $s_x^\perp = \frac{1}{(H_x^{Right_t} - H_x^{Left_t})}$, then when $s_x^* \in (s_x^\perp, 1)$, the DSE $H_x^{Left_t \cup Right_t}$ decreases monotonically with respect to s_x^* ; when $s_x^* \in (0, s_x^\perp)$, the DSE $H_x^{Left_t \cup Right_t}$ increase monotonically with respect to s_x^* .

Proof: The variable s_x^* is discontinuous, so we cannot directly find the derivative of $H_x^{Left_t \cup Right_t}$ with respect to s_x^* . For this problem, we can extend the domain of s_x^* from the original discontinuous point set to the continuous interval $(0, 1)$, then we can obtain the derivative of the DSE $H_x^{Left_t \cup Right_t}$ with respect to s_x^* :

$$\begin{aligned} & -H_x^{Right_t} + H_x^{Left_t} - \log_2 s_x^* - s_x^* \times \frac{1}{s_x^*} \times \log_2 e \\ & + \log_2 (1 - s_x^*) + (1 - s_x^*) \times \frac{1}{(1 - s_x^*)} \times \log_2 e \\ & = -H_x^{Right_t} + H_x^{Left_t} - \log_2 s_x^* + \log_2 (1 - s_x^*) \end{aligned}$$

when $s_x^* \in (s_x^\perp, 1)$, $-H_x^{Right_t} + H_x^{Left_t} - \log_2 s_x^* + \log_2 (1 - s_x^*) < 0$, according to the relationship between the derivative and the monotonicity of a function, we can infer that $H_x^{Left_t \cup Right_t}$ decreases monotonically with respect to s_x^* in the interval $(s_x^\perp, 1)$; when $s_x^* \in (0, s_x^\perp)$, $-H_x^{Right_t} + H_x^{Left_t} - \log_2 s_x^* + \log_2 (1 - s_x^*) > 0$, according to the relationship between the derivative and the monotonicity of a function, we can infer that $H_x^{Left_t \cup Right_t}$ increases monotonically with respect to s_x^* in the interval $(0, s_x^\perp)$.

Proposition 4 points out two important properties of the DSE $H_x^{Left_t \cup Right_t}$: ① when $s_x^* \in (s_x^\perp, 1)$, the less s_x^* , the greater the DSE $H_x^{Left_t \cup Right_t}$; ② when $s_x^* \in (0, s_x^\perp)$, the greater s_x^* , the greater the DSE $H_x^{Left_t \cup Right_t}$. The gray level histogram Υ_x is constructed by sampling the guiding edge image $K_{u_x^*}$ with the pixels taking the value 1 in the contour image c_t , and therefore the gray level histogram Υ_x is composed of the edge and non-edge pixels in the guiding edge image $K_{u_x^*}$. In the gray level histogram Υ_x , the left part less than the gray level l_x^* is mostly composed of non-edge pixels in the image $K_{u_x^*}$, and the right part greater than the gray level l_x^* is mostly composed of edge pixels in the image $K_{u_x^*}$. Thus the value s_x^* reflects the ratio of the non-edge regions to the edge regions in the contour image c_t located in the guiding edge image $K_{u_x^*}$. For the contour image c_t , the less s_x^* , the less the proportion of non-edge regions, which is an important feature expected for good segmentation. Thus, we can find a greater DSE $H_x^{Left_t \cup Right_t}$ in the monotonically decreasing interval $(s_x^\perp, 1)$ according to Property ①, or we can find a less DSE $H_x^{Left_t \cup Right_t}$ in the monotonically increasing interval

(0, s_x^\perp) according to Property ②. Both strategies are expected to achieve a more reasonable segmentation threshold.

There are two key techniques to increase the chance that s_x^* is in the monotonically decreasing interval ($s_x^\perp, 1$), thus the strategy of “finding a greater DSE $H_x^{\text{Left}_t \cup \text{Right}_t}$ in the monotonically decreasing interval ($s_x^\perp, 1$) according to Property ①” is more desirable in practice. One technique is to maximize the Shannon entropy difference in Section II.A and II.B, and the other technique is to separately process in the directions of x -axis and y -axis. First, since $s_x^\perp = \frac{1}{1+2\left(\frac{H_x^{\text{Right}_t} - H_x^{\text{Left}_t}}{H_x^{\text{Right}_t} - H_x^{\text{Left}_t}}\right)}$,

then the greater $H_x^{\text{Right}_t} - H_x^{\text{Left}_t}$ is, the less s_x^\perp is, and the wider the interval ($s_x^\perp, 1$) (see Fig. 4). Ideally, if s_x^\perp approaches or even equals 0, then the DSE $H_x^{\text{Left}_t \cup \text{Right}_t}$ will decrease monotonically over the entire interval (0, 1). Maximizing Shannon entropy difference in Section II.A and II.B tends to make $H_x^{\text{Right}_t} - H_x^{\text{Left}_t}$ as great as possible, while $H_x^{\text{Right}_t}$ and $H_x^{\text{Left}_t}$ are samples of H_x^{Right} and H_x^{Left} , respectively. A relatively greater $H_x^{\text{Right}_t} - H_x^{\text{Left}_t}$ also facilitates to generate a relatively greater $H_x^{\text{Right}_t} - H_x^{\text{Left}_t}$. Second, the y -direction edge in the image f is also suppressed to near 0 in the guiding edge image $K_{u_x^*}$, whereas the y -direction contour still exists in the contour image c_t (see Figs. 3(b) and (d)), which facilitates s_x^* easier to be greater than s_x^\perp . This is another reason why to separately process in the directions of x -axis and y -axis.

IV. ALGORITHM DESCRIPTION

Algorithm 1. Algorithm Name: MDSE

Input: a gray level image f to be segmented

Output: a selected threshold t^* and a thresholding result image

Step 1: Calculate the guiding edge images $K_{u_x^*}$ and $K_{u_y^*}$ according to Eqs. (5), (6), (7), and (8).

Step 2: Use H_{x_t} and H_{y_t} to record $H_x^{\text{Left}_t \cup \text{Right}_t}$ and $H_y^{\text{Left}_t \cup \text{Right}_t}$ during the subsequent loop, respectively, and use H_t to record the arithmetic mean of H_{x_t} and H_{y_t} . Use H_{\max} to record the maximum H_t during the subsequent loop, and use t^* to record the gray level corresponding to H_{\max} . The initial values of H_{x_t} , H_{y_t} , H_t , H_{\max} , and t^* are all 0. For each possible gray level $t \in [t_{\min}, t_{\max}]$ in the image f , repeat the following Steps 3 to 6 in ascending order.

Step 3: Utilize the gray level t to threshold the image f to obtain the corresponding binary image b_t .

Step 4: Extract the contour image c_t from the binary image b_t , which can be specifically divided into 3 small steps. Initially, let all pixels in the image c_t take value 1; then extract the inner contour: if the value of a pixel in b_t is 1 and its 4-neighborhood pixels all take value 1, set the pixel value of its corresponding position in c_t to 0; finally extract the outer contour: after generating the complement image \tilde{b}_t by reversing 0 and 1 in the image b_t , the pixel and its 4-neighborhood pixels in the image \tilde{b}_t are analyzed and judged in the same way as in the image b_t , and the image c_t is also set accordingly.

Step 5: Utilize the pixels with the value 1 in the image c_t to sample the guiding edge image $K_{u_x^*}$, and reconstruct a normalized gray level histogram from the sampled pixels, then calculate the DSE H_{x_t} from this normalized gray level histogram. Use the same way to process the guiding edge image $K_{u_y^*}$ and calculate the corresponding DSE H_{y_t} .

Step 6: Calculate the mean H_t of H_{x_t} and H_{y_t} , and then judge the relationship between H_t and H_{\max} . If $H_t > H_{\max}$, let H_{\max} equal H_t and t^* equal t .

Step 7: Generate the binary image b_{t^*} by thresholding the image f with the final threshold t^* , output the final threshold t^* and the binary image b_{t^*} .

V. EXPERIMENTAL RESULTS AND DISCUSSION

A. TEST ENVIRONMENT, QUANTITATIVE EVALUATION INDICATOR, AND METHODS PARTICIPATING COMPARISONS

The main software and hardware used for the test experiments are as follows: Intel Core i3-2350M 2.3GHz CPU, 4GB DDR2 memory, Windows 7 32-bit operating system, and Matlab 2009a 32-bit development platform. The test image set contains 40 synthetic images and 50 real-world images, and these test images have different gray level histograms. The test image set and the image set of segmentation reference are available online [40].

Misclassification error (ME) [20], [30] is a commonly used quantitative evaluation indicator when the segmentation accuracy of a thresholding method is evaluated. ME indicates the percentage of background pixels that are misclassified as target pixels and target pixels that are misclassified as background pixels in a segmentation result image, and its calculation formula is given as follows:

$$ME = \left(1 - \frac{|\phi \cap \phi_t| + |\varphi \cap \varphi_t|}{|\phi| + |\varphi|}\right) \times 100\% \quad (11)$$

where ϕ and φ respectively denote the set of the target and background pixels in the reference image, ϕ_t and φ_t respectively denote the set of the target and background pixels in the result image obtained by thresholding the test image with a threshold t , the symbol \cap represents the intersection operation, and the symbol $|\cdot|$ means to calculate the number of elements in a set. ME will be 0% when the thresholding result image is exactly the same as the reference image; ME will be 100% when the thresholding result image is the complement of the reference image.

Intersection over Union (IoU), also known as the Jaccard index, is also a popular evaluation metric for tasks such as segmentation, object detection and tracking [41]. IoU emphasizes similarity between finite sample sets, and is formally defined as the size of the intersection divided by the size of the union of the sample sets. The mathematical representation of IoU is written as:

$$IoU = \frac{|\phi \cap \phi_t|}{|\phi| + |\phi_t| - |\phi \cap \phi_t|} \quad (12)$$

IoU will be 0 when the target pixels of the reference image and the thresholding result image do not overlap at all;

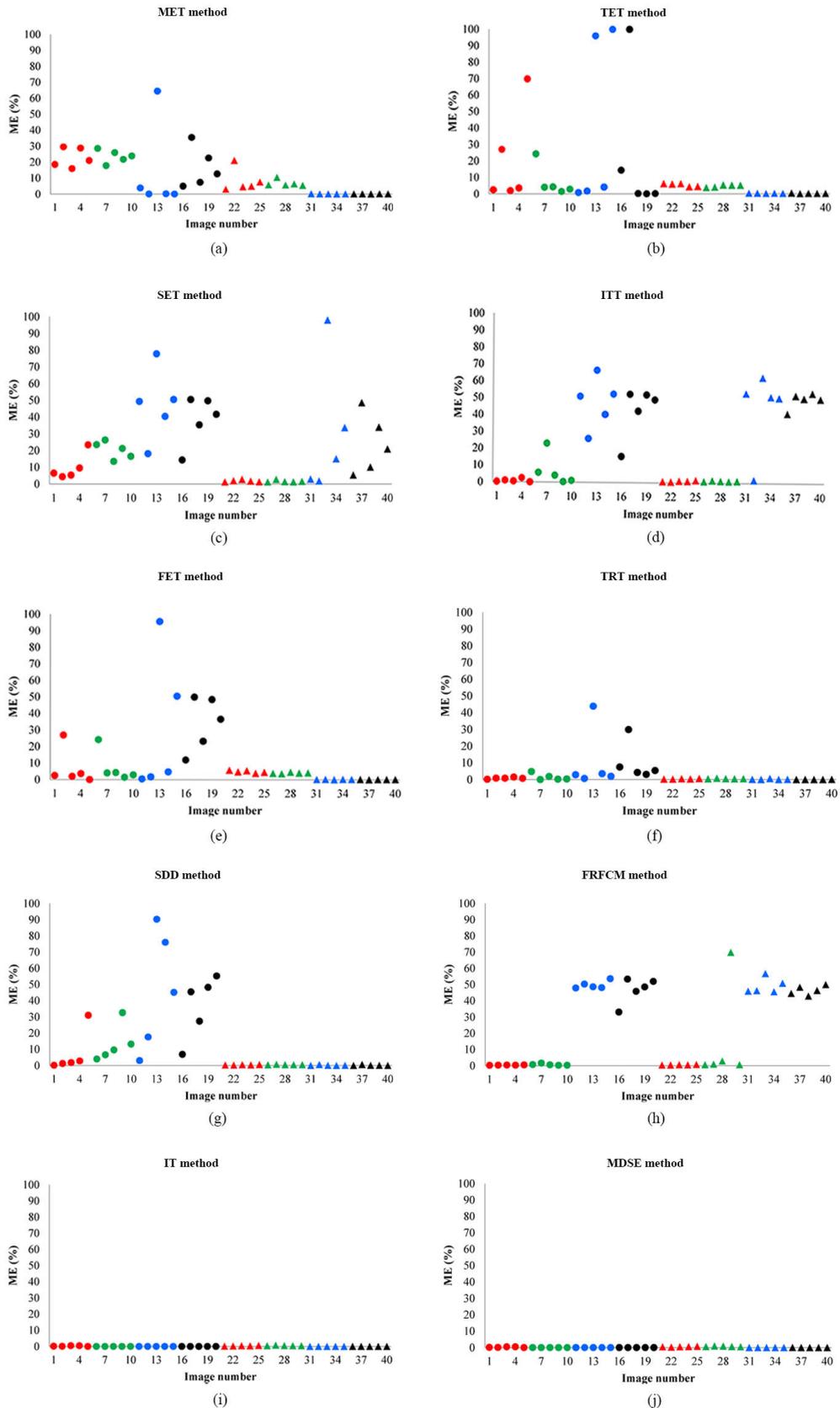


FIGURE 5. ME values of 10 segmentation methods on 40 synthetic images with different gray level histograms. In each subfigure, the red dots, the green dots, the blue dots, the black dots, the red triangles, the blue triangles, and the black triangles indicate the ME values of the corresponding method on the first to eighth group of test images, respectively.

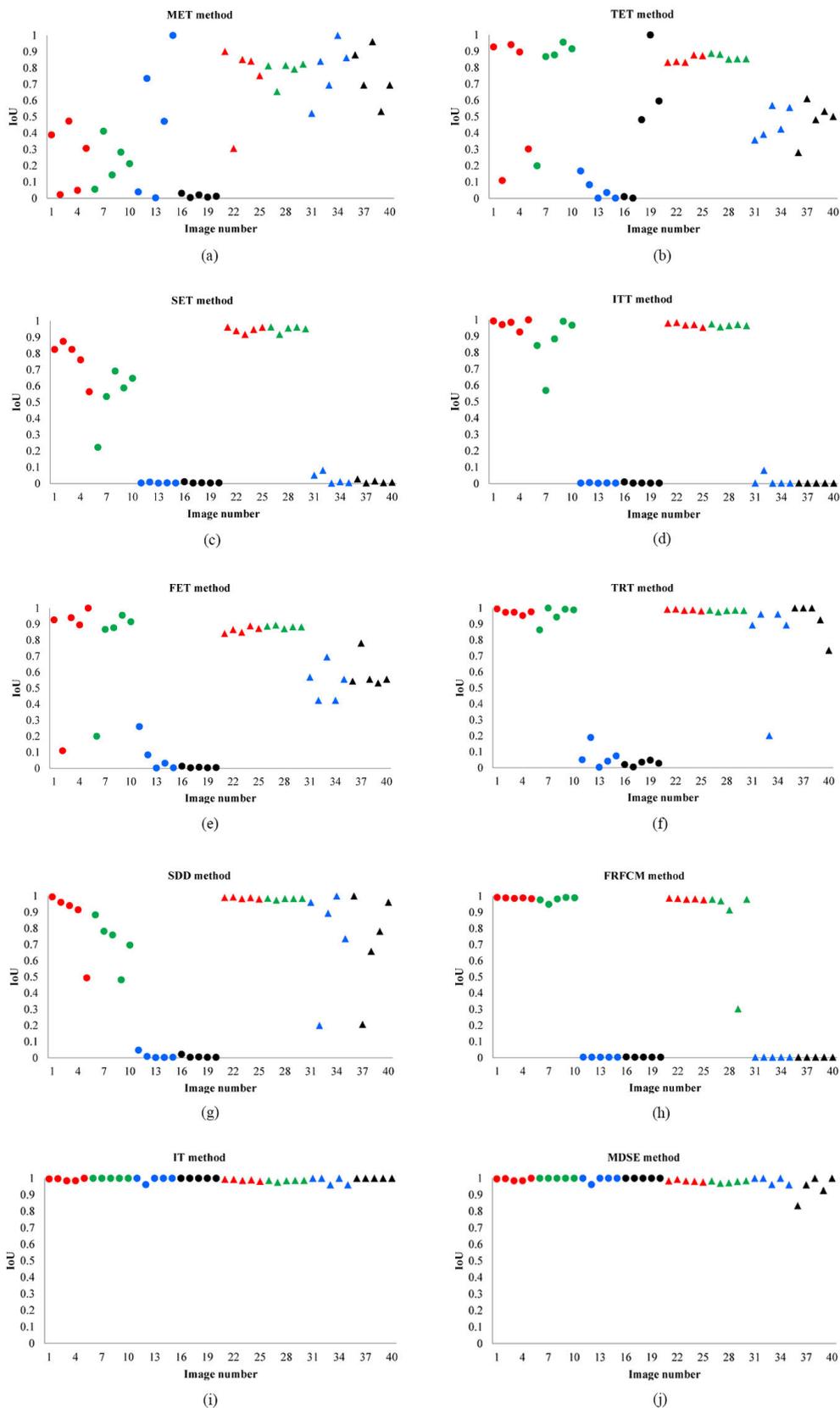


FIGURE 6. IoU values of 10 segmentation methods on 40 synthetic images with different gray level histograms. In each subfigure, the red dots, the green dots, the blue dots, the black dots, the red triangles, the green triangles, the blue triangles, and the black triangles indicate the IoU values of the corresponding method on the first to eighth group of test images, respectively.

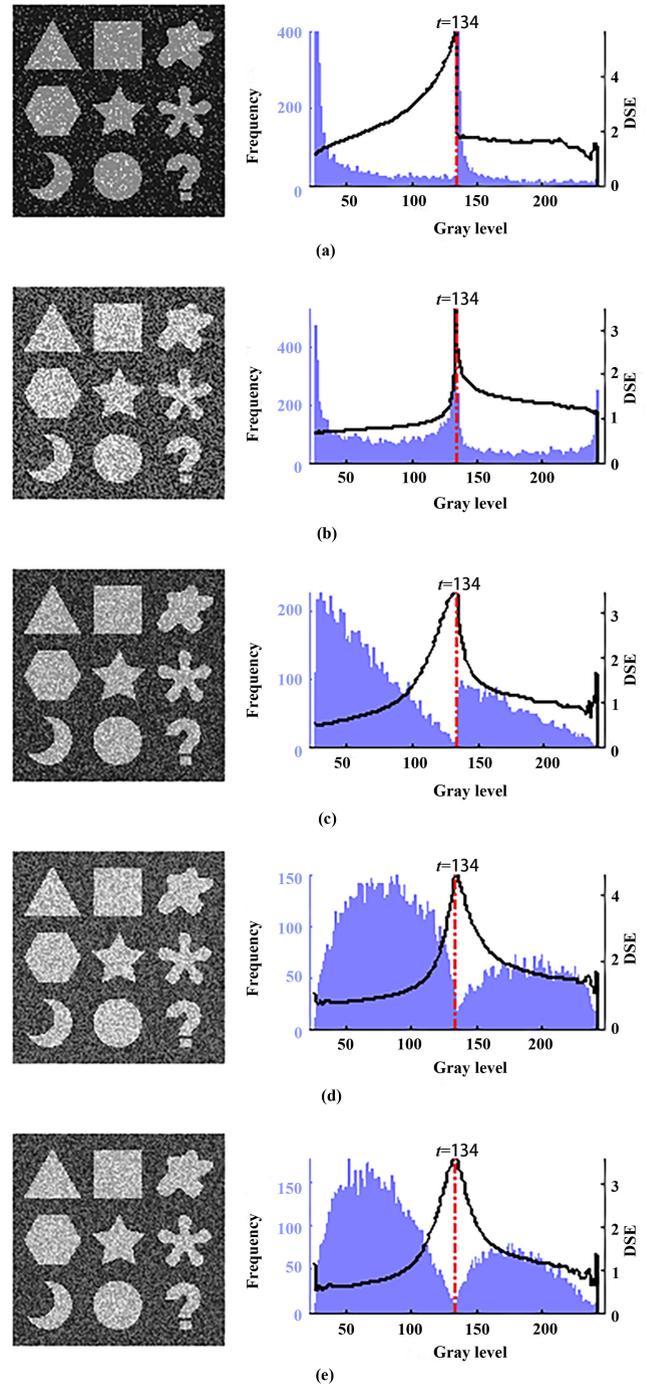
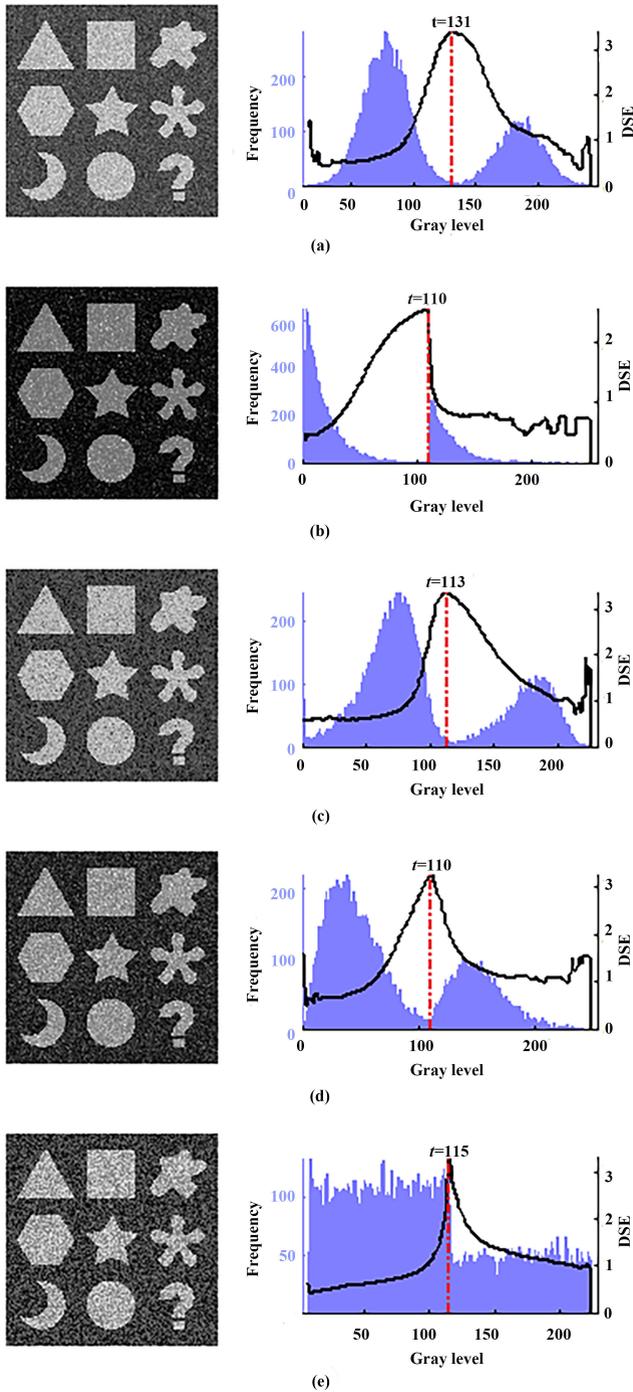


FIGURE 7. The first group test images and the thresholds selected by the MDSE method. In each subfigure, the test image is shown on the left, its gray level histogram is displayed in the blue area on the right. In addition, the black curve on the right shows the objective function curve of the MDSE method for calculating thresholds, and the red dotted line and the number next to it indicate the threshold selected by the MDSE method (the same below). (a)-(e) sequentially show that the gray level distributions of the target and the background are a Gaussian distribution, a gamma distribution, an extreme value distribution, a Rayleigh distribution, and a uniform distribution.

FIGURE 8. The second group test images and the thresholds selected by the MDSE method. (a)-(e) sequentially show that the gray level distributions of the target and the background are a beta distribution with different parameters: (a) $\alpha = 0.1, \beta = 0.8$, (b) $\alpha = 0.5, \beta = 0.5$, (c) $\alpha = 1.0, \beta = 2.0$, (d) $\alpha = 1.5, \beta = 1.5$, (e) $\alpha = 1.5, \beta = 2.0$.

IoU will be 1 when the target pixels of the reference image and the thresholding result image are a perfect match.

The proposed MDSE method is compared with global Masi entropy thresholding (MET) method [30], global

adaptive Tsallis entropy thresholding (TET) method [29], local Shannon entropy thresholding (SET) method [22], iterative triclass thresholding (ITT) method [42], fuzzy entropy thresholding (FET) method [31], transition region thresholding (TRT) method [20], slope difference distribution (SDD) clustering method [4], fast and robust fuzzy c-means (FRFCM) clustering method [5], and interactive

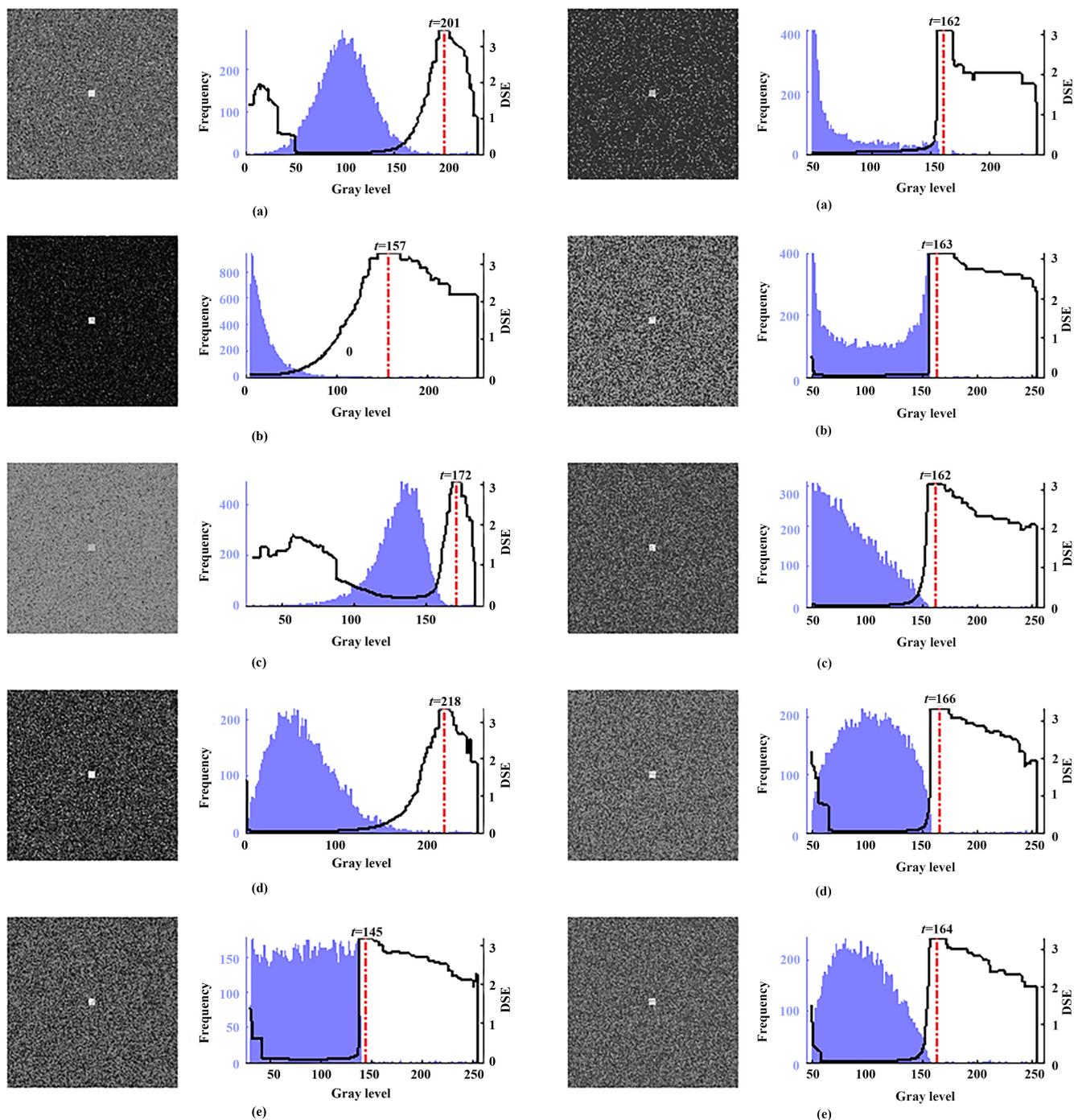


FIGURE 9. The third group test images and the thresholds selected by the MDSE method. (a)-(e) sequentially show that the gray level distributions of the target and the background are a Gaussian distribution, a gamma distribution, an extreme value distribution, a Rayleigh distribution, and a uniform distribution.

FIGURE 10. The fourth group test images and the thresholds selected by the MDSE method. (a)-(e) sequentially show that the gray level distributions of the target and the background are a beta distribution with different parameters: (a) $\alpha = 0.1, \beta = 0.8$, (b) $\alpha = 0.5, \beta = 0.5$, (c) $\alpha = 1.0, \beta = 2.0$, (d) $\alpha = 1.5, \beta = 1.5$, (e) $\alpha = 1.5, \beta = 2.0$.

thresholding (IT) method [43]. The IT method interactively selects a segmentation threshold, and the binary image corresponding to this threshold has the smallest ME value. Therefore, the IT method can be used as a reference method for other methods participating comparisons in terms of segmentation accuracy.

B. EXPERIMENTS ON SYNTHETIC IMAGES

Affected by factors such as random noise, low pass filter, as well as the size of target and background, the gray level histograms of different gray level images may show different histogram patterns. In addition to the common bimodal pattern, there are unimodal, peakless, and even multimodal

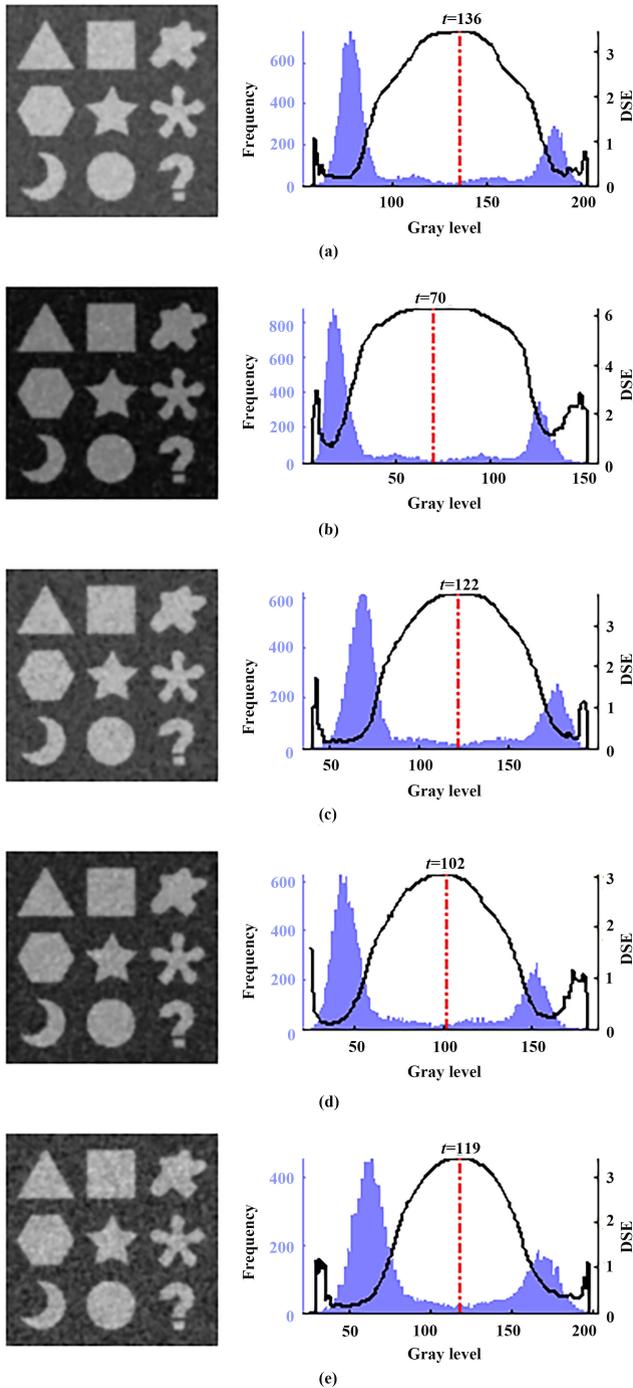


FIGURE 11. The fifth group test images and the thresholds selected by the MDSE method. The test images sequentially displayed in (a)-(e) are obtained by performing a Gaussian point diffusion on the test images in Figs. 7(a)-(e).

patterns. Moreover, there are many possibilities for the basic distribution consisting of the histogram pattern. In addition to the common Gaussian distribution, there are also gamma distribution, extreme value distribution, Rayleigh distribution, uniform distribution, and beta distribution.

In comparison experiments on synthetic images, to comprehensively test the segmentation adaptability of 10 segmentation methods to different histogram patterns, 40 different

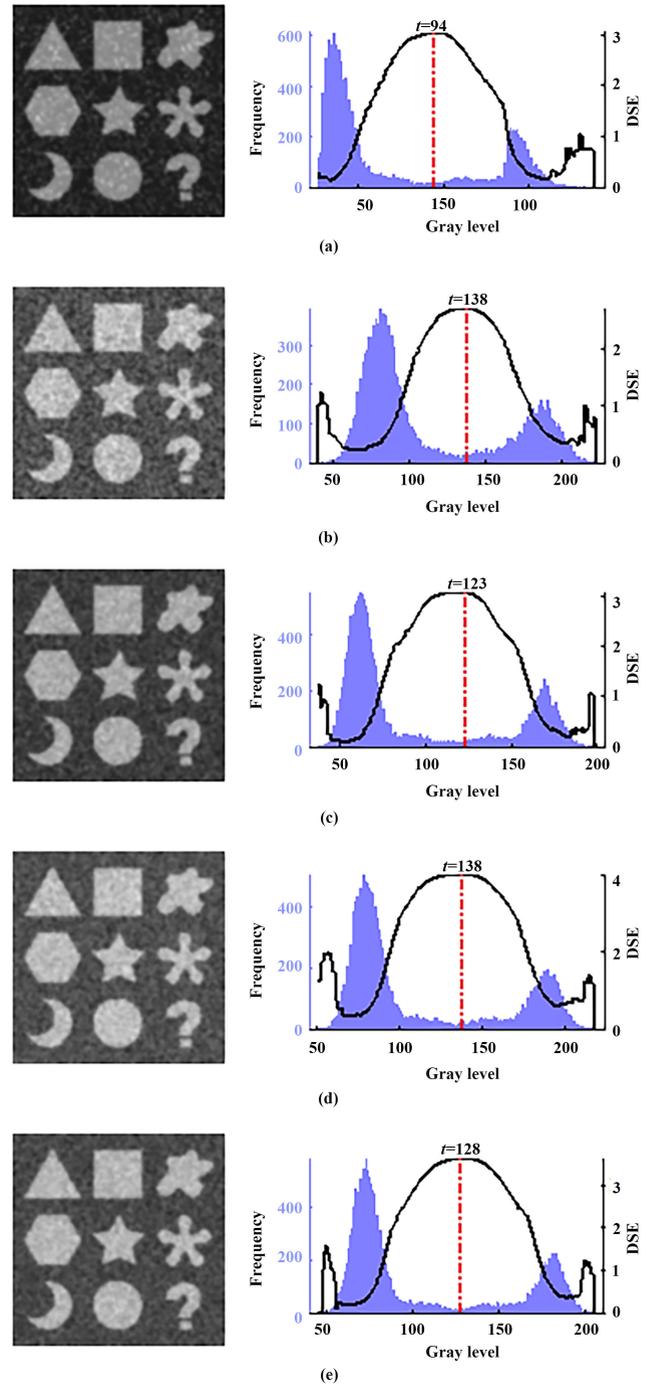


FIGURE 12. The sixth group test image and the thresholds selected by the MDSE method. The test images sequentially displayed in (a)-(e) are obtained by performing a Gaussian point diffusion on the test images in Figs. 8(a)-(e).

synthetic images are divided into 8 groups and tested separately. Figs. 5-6 shows the quantitative results about the segmentation accuracy of these 10 segmentation methods on these 8 groups of test images, and Figs. 7-14 show these 8 groups of test images in sequence.

In the first group test images, the size ratio of the target to the background is approximately equal to 3:7. To simulate the basic distributions such as Gaussian distribution, gamma

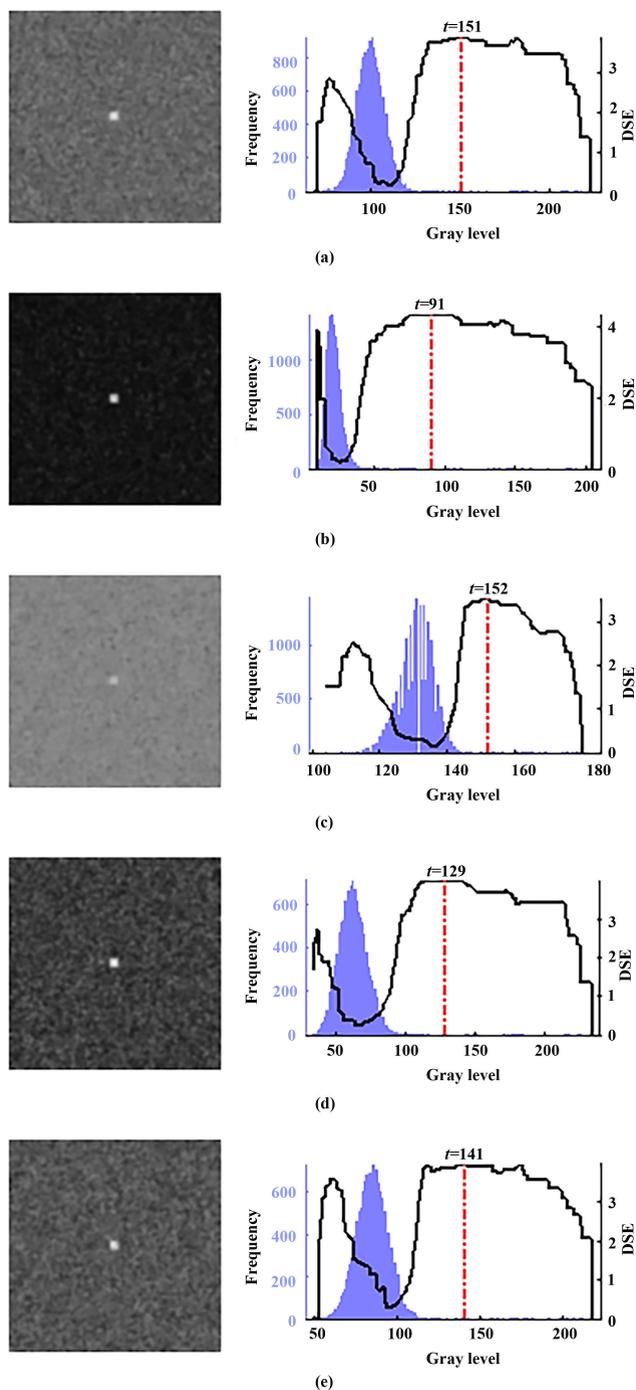


FIGURE 13. The seventh group test image and the thresholds selected by the MDSE method. The test images sequentially displayed in (a)-(e) are obtained by performing a Gaussian point diffusion on the test images in Figs. 9(a)-(e).

distribution, extreme value distribution, Rayleigh distribution, and uniform distribution, the test images are generated by adding Gaussian noise, gamma noise, extreme value noise, Rayleigh noise, and uniform noise to the noise-free synthetic image, respectively (see Fig. 7). Because the size ratio of the target to the background is relatively balanced, when the basic distribution is a Gaussian distribution, a gamma distribution,

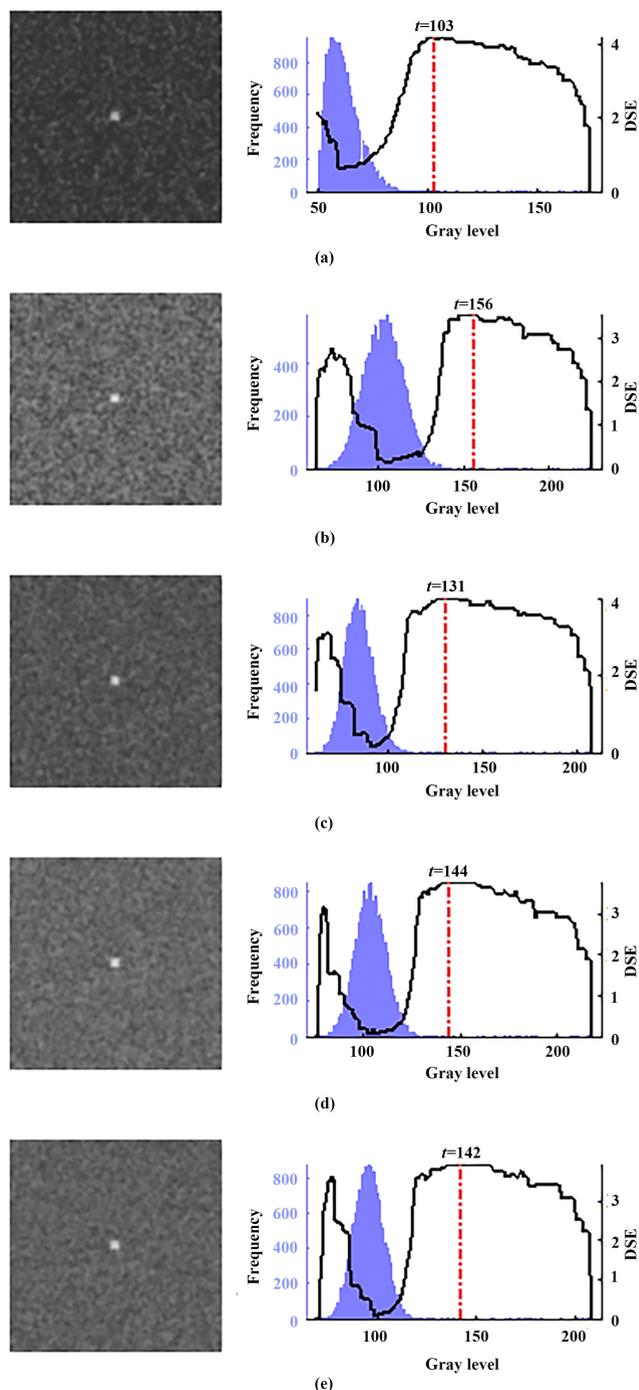


FIGURE 14. The eighth group test image and the thresholds selected by the MDSE method. The test images sequentially displayed in (a)-(e) are obtained by performing a Gaussian point diffusion on the test images in Figs. 10(a)-(e).

an extreme value distribution, and a Rayleigh distribution, the gray level histograms of these test images show typical bimodal patterns (see Figs. 7(a)-(d)); when the basic distribution is a uniform distribution, the gray level histogram of the test image show peakless pattern (see Fig. 7(e)). As shown by the red dots in Figs. 5-6 (the corresponding image numbers are 1 to 5): ① the overall segmentation result of the MET

method is the worst. ② the segmentation results of the TET method are slightly better than the MET method in general, but its mis-segmentation is serious when the basic distribution is a gamma distribution and a uniform distribution. ③ the overall segmentation results of the SET, FET, and SDD method are better than the MET and TET methods, but their mis-segmentation are also serious when the basic distribution is a uniform distribution or a gamma distribution. ④ the segmentation results of the ITT, TRT, FRFCM, and MDSE methods are obviously better than the other 5 compared methods; in particular, the segmentation results of the MDSE and IT methods are completely consistent, and they achieve the theoretically optimal segmentation in terms of minimizing ME and maximizing IoU.

In the second group test images, the size ratio of the target to the background is approximately equal to 3:7. The test images are generated by adding beta noise with different parameters to the noise-free synthetic image, respectively. As a result, the gray level histograms of the test images are bimodal (see Figs. 8(a), (c)-(e)) or multimodal (see Fig. 8(b)). As shown by the green dots in Figs. 5-6 (the corresponding image numbers are 6 to 10): ① both the MET and SET methods have serious mis-segmentations. ② the segmentation results of the TET, ITT, FET, and SDD methods are unstable, and the ME value is sometimes large and sometimes small. ③ compared with the other 6 methods, the TRT, FRFCM, MDSE, and IT methods obtain significantly better segmentation accuracy; in particular, the proposed MDSE method and the IT method achieve completely consistent segmentation results, and they once again get the theoretically optimal segmentation.

In the third and fourth group test images, the size ratio of the target to the background is approximately equal to 3:1997. The test images in the third group are generated by adding Gaussian noise, gamma noise, extreme value noise, Rayleigh noise, and uniform noise to the noise-free synthetic image, respectively (see Fig. 9), while the test images in the fourth group are generated by adding beta noise with different parameters to the noise-free synthetic image (see Fig. 10). Because the size ratio of the target to the background is seriously unbalanced, the gray level histograms of the test images show a unimodal pattern (see Figs. 9(a)-(d) and Figs. 10(a), (c)-(e)) or a bimodal pattern (see Fig. 10(b)) or a peakless pattern (see Fig. 9(e)). We can observe that from the blue dots and the black ones in Figs. 5-6 (the corresponding image numbers are 11 to 15, and 16 to 20, respectively): ① the MET and TET methods are more suitable for segmenting test images with right-biased unimodal histogram, but not for test images with left-biased unimodal histogram (see Fig. 9(c)); ② the SET, ITT, FET, TRT, SDD, and FRFCM methods have serious mis-segmentations on these two groups of test images, which shows that these methods are not suitable for segmenting images with a serious imbalance between the size of target and background. ③ On these two groups of test images, the thresholds calculated by the MDSE method are always consistent with the thresholds selected by the IT

method, so their ME values and IoU values are also always consistent.

The test images of the fifth and sixth group are generated by performing Gaussian point diffusion on the test images of the first and second group, respectively (see Figs. 11-12). The targets and the backgrounds in these two groups of test images are both blurred. As a result, the gray level histograms of these test images are bimodal and with flat valleys, but they are different in the flat valleys. We can observe that from the red triangles and green ones in Figs. 5-6 (the corresponding image numbers are 21 to 25, and 26 to 30, respectively): ① the ME values of MET, SET, FET, and SDD methods significantly decrease in general, and the IoU values of MET and SET methods significantly increase in general, which indicates that, contrasted with the test images with bimodal deep-valley histograms, the MET and SET methods are more suitable for segmenting the images with bimodal flat-valley histograms; ② some ME or IoU values of the TET method have risen and some have fallen, but the TET method has relatively better a adaptability to the test image with bimodal flat-valley histograms; ③ the average ME values of the ITT method, the MDSE method, the TRT method, and the IT method on the test images of Group 5 and Group 6 are 0.94%, 0.59%, 0.44%, and 0.42%, respectively, which indicates the ITT method, the TRT method, and the MDSE method can segment test images with bimodal flat-valley histogram very well.

The test images of the seventh and eighth group are generated by performing Gaussian point diffusion on the test images of the third and fourth group, respectively (see Figs. 13-14). The two groups of test images all show a small blurred target on a blurred background. The gray level histograms of these test images are right-biased unimodal. As shown by the blue triangles and black ones in Figs. 5-6 (the corresponding image numbers are 31 to 35, and 36 to 40, respectively): ① the MET, TET, FET, TRT, and SDD methods are more suitable for segmenting test images with right-biased unimodal histogram; ② the SET, ITT, and FRFCM methods have more serious mis-segmentation on these two groups of test images; ③ the MDSE and IT methods have completely consistent segmentation results on the seventh group test images; only on the first, second, and fourth test images in the eighth group, the segmentation results of the MDSE and IT methods are slightly different.

C. EXPERIMENTS ON REAL-WORLD IMAGES

50 real-world images with different gray level histograms are utilized to further test the segmentation adaptability of 10 segmentation methods. The gray level histograms of these real-world images can be approximated by a mixture of several distributions such as Gaussian distribution, gamma distribution, extreme value distribution, Rayleigh distribution, uniform distribution, and beta distribution. The 50 test images are divided into 4 groups. The first group of images are numbered from 1 to 11, the second group of images are numbered from 12 to 28, the third group of images are

numbered from 29 to 46, and the fourth group of images are numbered from 47 to 50. The gray level histograms of the first group of test images appear as unimodal left-biased, unimodal right-biased or unimodal unbiased patterns; the gray level histograms of the second group of test images are shown as bimodal deep-valley or bimodal flat-valley patterns; due to the complex background, the gray level histograms of the third group of test images appear as multimodal patterns with the number of modalities greater than or equal to 3; the gray level histograms of the fourth group of test images appear as peakless patterns.

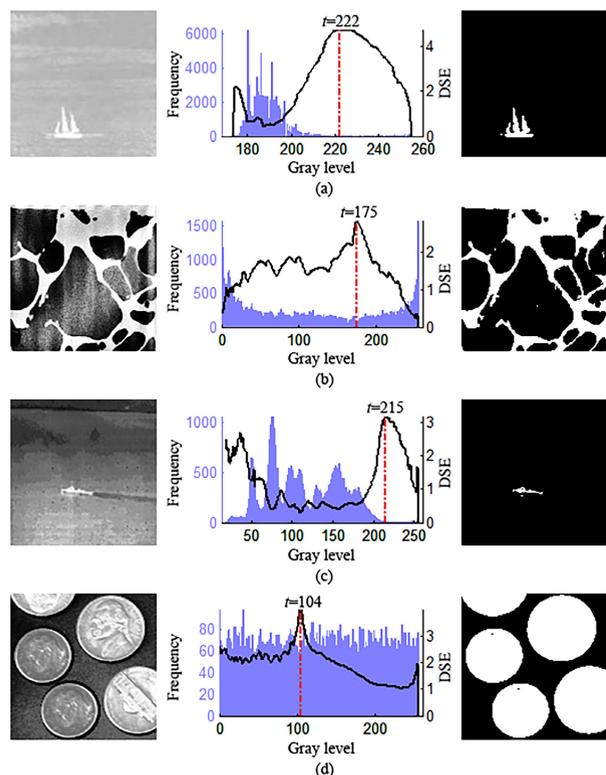


FIGURE 15. 4 real-world images numbered 3, 18, 39, and 47 (left column), their corresponding gray level histograms (middle column), and the segmentation results of the MDSE method (right column). In the middle column, each gray level histogram is displayed in a blue area, each black curve shows the objective function curve of the MDSE method for calculating thresholds, and the red dotted line and the number next to it indicate the threshold selected by the MDSE method.

As 4 representative examples, Fig. 15 shows 4 test images numbered 3, 18, 39, and 47 and their corresponding gray level histograms, as well as the segmentation results obtained by the MDSE method. It can be observed that the MDSE method accurately separates the target from the background. In fact, for the 4 test images, the corresponding differences of the ME values between the MDSE method and the IT method are 0.088%, 0.092%, 0.003%, and 0.317%, respectively; the corresponding differences of the IoU values between the MDSE method and the IT method are -0.053, -0.003, -0.01, and -0.005, respectively. Additionally, the differences between the selected thresholds obtained by the MDSE method and the IT method are -4, -1, 1, and 1, respectively. All this

indicates that the segmentation results of the MDSE method are very close to the optimal segmentation results determined by the IT method.

Figs. 16-17 shows the quantitative comparison results about segmentation accuracy of 10 segmentation methods on 50 real-world images. As shown in Figs. 16-17: ① for the test images with unimodal, bimodal, multimodal, or peakless histograms, the average ME values of the MDSE method and the IT method are much smaller than the average ME values of other methods, and the average IoU values of the MDSE method and the IT method are much larger than the average IoU values of other methods; ② although the ITT and FRFCM methods have good segmentation accuracy for the images with bimodal histogram, they have serious mis-segmentation in the case of unimodal, multimodal, and peakless patterns; ③ the SET method has serious mis-segmentation in all 4 cases, and the respective average ME values all exceed 11%, and the respective average IoU values are less than 0.736; in particular, the ME values on 12 test images are greater than 20%; ④ compared with the SET, ITT, and FRFCM methods, the overall segmentation accuracies of the TET, FET, and MET methods are slightly better, but they have relatively greater ME values for some test images, even greater than 20%.

The IT method interactively selects segmentation thresholds under the criterion of minimizing the ME value, so the thresholds it chooses are optimal ones in the sense of minimizing ME value. By calculating and comparing the differences between the thresholds selected by the IT method and other methods (see Fig. 18), the degree of deviation in the threshold selection of other methods can be quantitatively compared. As shown in Fig. 18: ① the threshold differences between the MDSE, MET, TET, SET, ITT, FET, TRT methods and the IT method are distributed in intervals [-7, 9], [-97, 119], [-169, 24], [-120, 89], [-153,84], [-134,84], and [-108,71]; ② although the thresholds selected by the MET, TET, SET, ITT, FET, and TRT methods on partial test images are relatively close to the optimal thresholds, their selected thresholds are far from the optimal thresholds as a whole: the number of cases where the absolute value of the threshold difference is greater than or equal to 10 are 45, 41, 37, 35, 44, and 27, respectively; the number of cases where the absolute value of the threshold difference is greater than or equal to 20 are still 33, 31, 31, 28, 33, and 15, respectively; ③ overall, the thresholds selected by the MDSE method on different gray level histograms are closer to optimal thresholds, and the absolute value of the threshold difference are all within 9 gray levels.

D. COMPARISON EXPERIMENTS ON COMPUTATIONAL EFFICIENCY

Comparing CPU time of different methods under same software and hardware conditions can intuitively reflect the difference in the computational efficiency of different methods. Note that CPU time of same program will fluctuate slightly when running at different times. To reduce this kind

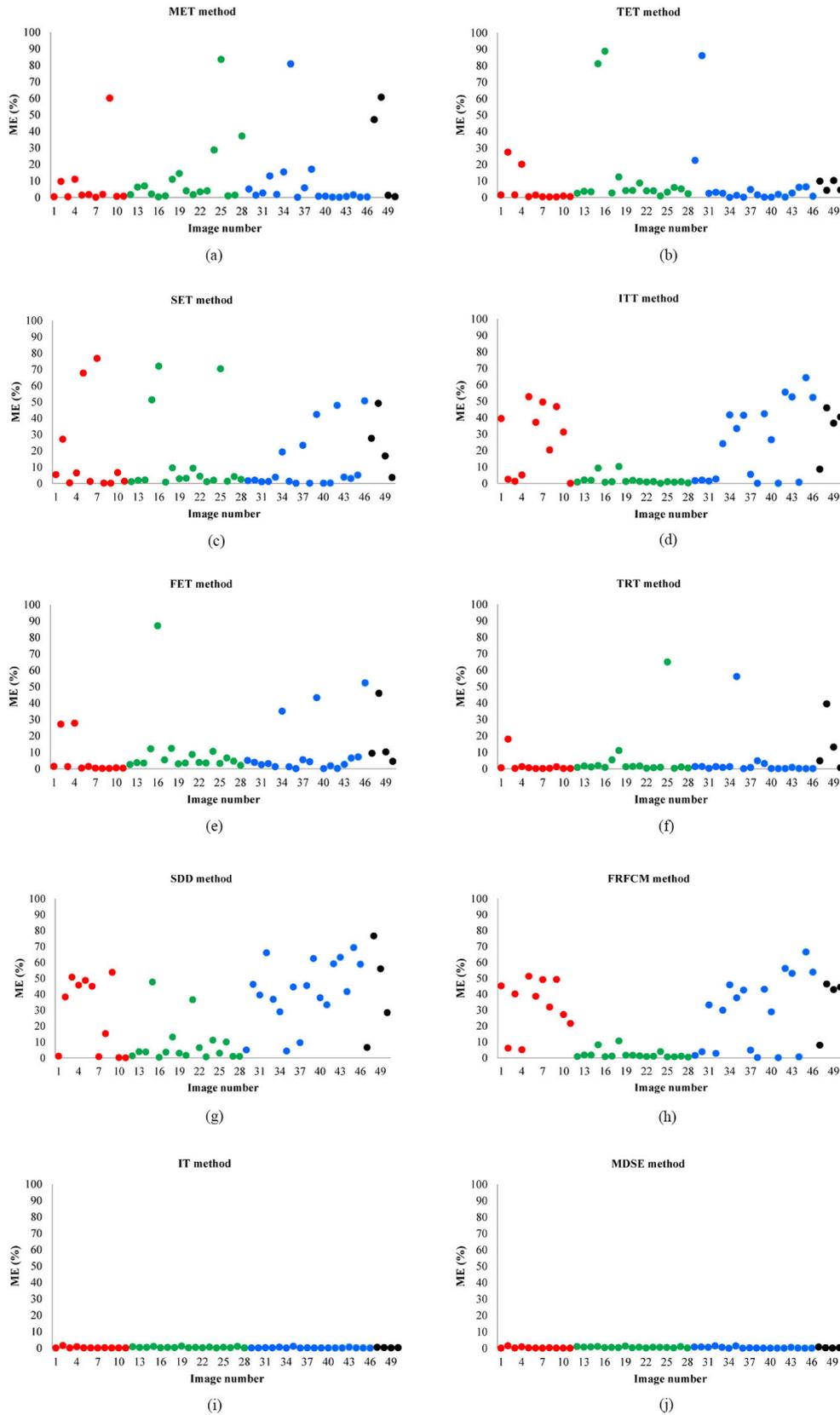


FIGURE 16. ME values of 10 segmentation methods on 50 real-world images with different gray level histograms. In each subfigure, the red dots, the green dots, the blue dots, and the black dots indicate the ME values of the corresponding method on the first to fourth group of test images, respectively.

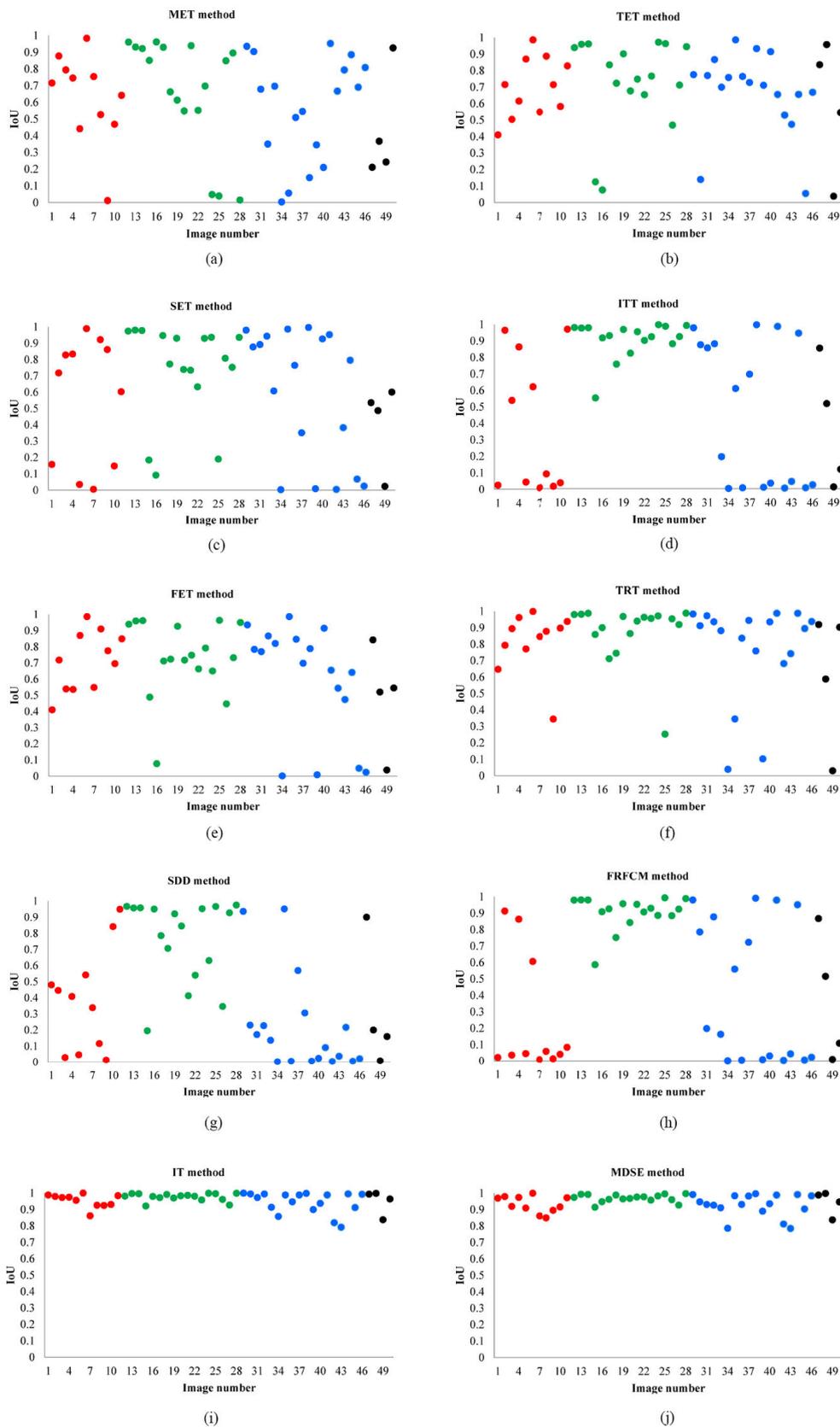


FIGURE 17. IoU values of 10 segmentation methods on 50 real-world images with different gray level histograms. In each subfigure, the red dots, the green dots, the blue dots, and the black dots indicate the IoU values of the corresponding method on the first to fourth group of test images, respectively.

TABLE 2. The average CPU time, average ME, and average IoU of the MDSE method at different jumping step size.

Step size γ	1	2	3	4	5	6	7	8	9	10
Synthetic images										
Average CPU time (s)	0.609	0.354	0.268	0.222	0.198	0.188	0.179	0.172	0.168	0.163
Average ME (%)	0.18	0.18	0.18	0.18	0.18	0.35	5.28	0.37	1.42	3.56
Average IoU	0.98	0.98	0.98	0.98	0.97	0.95	0.83	0.95	0.93	0.85
Real-world images										
Average CPU time (s)	0.939	0.545	0.401	0.332	0.297	0.277	0.263	0.252	0.249	0.244
Average ME (%)	0.43	0.43	0.43	0.43	0.44	2.38	3.53	5.73	2.37	4.87
Average IoU	0.94	0.94	0.94	0.94	0.93	0.88	0.86	0.81	0.88	0.84

MET, FRFCM, and SDD methods are relatively higher, while the computational efficiencies of the SET, FET, and MDSE method is relatively lower.

The ITT method first constructs the gray level distribution of original gray level image, and then performs a relatively simple calculation about mean and variance, therefore it takes the least average CPU time. The MET method also needs to calculate the gray level distribution of original image in advance, and then computes Masi entropy. The computation of Masi entropy involves logarithm operation and is relatively more complicated than the computation of mean and variance, so the computational cost of the MET method is slightly higher than that of the ITT method. The FRFCM method is an improved FCM clustering algorithm. Its main computations include morphological reconstruction of the image to be segmented, constructing gray level histogram of the reconstructed image, and fast filtering of membership. The computational cost of the FRFCM method is higher than that of the ITT method, so its average CPU time is more than that of the ITT method. The SDD method is also a clustering segmentation method. It first constructs a normalized gray level histogram and smooths the histogram in the frequency domain, and then calculates a slope difference of the smoothed gray level histogram. It finally selects clustering centers based on the peaks of the computed slope difference distribution. The last two calculations of the SDD method are more complex than the calculation of Masi entropy of the MET method, so the calculation cost of the former is slightly higher than the latter.

In addition to the gray level distribution of original image, the TET method also needs to automatically calculate the entropy index q of Tsallis entropy by analyzing the redundancy of original image, as well as the power of q of each gray level distribution probability. As a result, its average CPU time is greater than the MET's CPU time. The TRT method is a thresholding method based on transition region. Its key points are to first extract a transition region set between targets and the background, and then follow a basic rule, that is, the elements of stable transition region set are equal or close to each other in the average gray level, to obtain the final segmentation threshold using the maximizing 1-STRS strategy. The main computations of this method occur in the construction of transition region set.

The SET method is a local entropy method. It needs to utilize an original gray level image to first construct a gray level co-occurrence matrix, and then divide the gray level co-occurrence matrix into 4 quadrants, and define the Shannon entropy on each quadrant. Compared with the ITT, MET, and TET methods that all are based on one-dimensional gray level histogram, the SET method is more time-consuming to build a gray level co-occurrence matrix, so its average CPU time is greater than ITT, MET, and TET methods. The FET method is a thresholding method based on fuzzy entropy. It essentially utilizes a fuzzy membership function to modify the gray level histogram of an original gray level image. Its main work involves two parts: one is to construct a fuzzy membership function based on the division of the gray level histogram, and the other is to define new fuzzy entropy based on the constructed fuzzy membership function. The computational cost of the MDSE method mainly arises from the relatively complex calculations involving the guiding edge images, the changing contour images, and the DSE. Consequently, its average CPU time is the relatively largest.

To reduce the CPU time of the MDSE method, the goal of determining maximum DSE can be divided into 2 procedures: jumping processing and stepwise processing. In the jumping processing, processing the gray level one by one in the interval $[t_{\min}, t_{\max}]$ in Step 2 of Algorithm 1 in Section IV is changed to a jumping processing with a certain step size γ , while the other steps for calculating the DSE are unchanged. A gray level t_{temp}^* is output after completing the jumping processing, and it corresponds to the maximum DSE generated during the jumping processing. Then, in the stepwise processing, the gray level in the interval $[t_{temp}^* - \gamma + 1, t_{temp}^* + \gamma - 1]$ is processed one by one, that is, the interval $[t_{\min}, t_{\max}]$ in Step 2 in Algorithm 1 is replaced by the interval $[t_{temp}^* - \gamma + 1, t_{temp}^* + \gamma - 1]$, while the other steps for calculating the DSE are unchanged. The final threshold t^* is output after completing the stepwise processing, and it corresponds to the maximum DSE in the gray level interval $[t_{temp}^* - \gamma + 1, t_{temp}^* + \gamma - 1]$.

As shown in Table 2, for 40 synthetic images and 50 real-world images: ① when the step size γ changes from 1 to 5, the corresponding average CPU time decreases gradually, but the corresponding average ME value and average IoU value remain basically unchanged; ② when the step size γ changes

from 5 to 10, the corresponding average CPU time remains basically stable, but the corresponding average ME value and average IoU value fluctuate. Consequently, when the step size γ is 3, 4, or 5, the MDSE method can approach the TET method or the TRT method in terms of computational efficiency, while maintaining the segmentation accuracy.

VI. CONCLUSION AND FUTURE WORKS

For the gray level images with a unimodal, bimodal, multimodal or peakless histogram, when the gray level distribution of the target or the background can be approximated by a Gaussian distribution, a gamma distribution, an extreme value distribution, a Rayleigh distribution, a uniform distribution, or a beta distribution, compared with the MET, TET, SET, ITT, FET, TRT, SDD, and FRFCM methods, the MDSE method has more flexible adaptivity of selecting threshold and better segmentation accuracy.

The MDSE method possesses the above 2 advantages mainly due to the following 3 aspects: ① the MDSE method generates the guiding edge images in horizontal and vertical directions respectively, which is conducive to highlighting the transitional regions between the target and the background in advance. Moreover, sampling the guiding edge images with the changing contour images favors constructing dynamically the gray level distribution of the transitional and non-transitional regions. ② the MDSE method constructs a series of one-dimensional gray level histograms using the invariant guiding edge image and the dynamically changing contour images, and utilizes the Shannon entropy without parameter as the entropy calculation model. Thus, the MDSE method can calculate the Shannon entropy that indirectly reflects both gray level distributions and pixel positions. ③ for each possible gray level in the original image, the MDSE method utilizes it to threshold the original image to generate a corresponding binary image, and extracts a contour image from the binary image. During this whole procedure, the segmentation threshold maintains a close relationship with the original image by means of the contour image.

The MDSE method needs to calculate the guiding edge image and the changing contour images, which makes it inferior to the MET, TET, SET, ITT, FET, TRT, SDD, and FRFCM methods in terms of computational efficiency. How to improve the computational efficiency of the MDSE method is worth further investigation. In the future, some generalized entropies can be considered to replace the Shannon entropy adopted by the MDSE method, then the automatic evaluation of the parameters in those generalized entropies will be worth further study.

REFERENCES

- [1] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron. Imag.*, vol. 13, no. 1, pp. 146–165, Jan. 2004, doi: [10.1117/1.1631315](https://doi.org/10.1117/1.1631315).
- [2] S.-Y. Wan and W. E. Higgins, "Symmetric region growing," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1007–1015, Sep. 2003, doi: [10.1109/TIP.2003.815258](https://doi.org/10.1109/TIP.2003.815258).
- [3] R. Nock and F. Nielsen, "Statistical region merging," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 11, pp. 1452–1458, Nov. 2004, doi: [10.1109/TPAMI.2004.110](https://doi.org/10.1109/TPAMI.2004.110).
- [4] Z. Wang, "Determining the clustering centers by slope difference distribution," *IEEE Access*, vol. 5, pp. 10995–11002, 2017, doi: [10.1109/ACCESS.2017.2715861](https://doi.org/10.1109/ACCESS.2017.2715861).
- [5] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, and A. K. Nandi, "Significantly fast and robust fuzzy C-means clustering algorithm based on morphological reconstruction and membership filtering," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 5, pp. 3027–3041, Oct. 2018, doi: [10.1109/TFUZZ.2018.2796074](https://doi.org/10.1109/TFUZZ.2018.2796074).
- [6] Y.-C. Lin, Y.-P. Tsai, Y.-P. Hung, and Z.-C. Shih, "Comparison between immersion-based and toboggan-based watershed image segmentation," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 632–640, Mar. 2006, doi: [10.1109/TIP.2005.860996](https://doi.org/10.1109/TIP.2005.860996).
- [7] T. Lei, X. Jia, T. Liu, H. Meng and A. K. Nandi, "Adaptive morphological reconstruction for seeded image segmentation," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5510–5523, Nov. 2019, doi: [10.1109/TIP.2019.2920514](https://doi.org/10.1109/TIP.2019.2920514).
- [8] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape," *Int. J. Comput. Vis.*, vol. 72, no. 2, pp. 195–215, Apr. 2007, doi: [10.1007/s11263-006-8711-1](https://doi.org/10.1007/s11263-006-8711-1).
- [9] B. Peng, L. Zhang, and D. Zhang, "A survey of graph theoretical approaches to image segmentation," *Pattern Recognit.*, vol. 46, no. 3, pp. 1020–1038, Mar. 2013, doi: [10.1016/j.patcog.2012.09.015](https://doi.org/10.1016/j.patcog.2012.09.015).
- [10] X. Chen and L. Pan, "A survey of graph cuts/graph search based medical image segmentation," *IEEE Rev. Biomed. Eng.*, vol. 11, pp. 112–124, 2018, doi: [10.1109/RBME.2018.2798701](https://doi.org/10.1109/RBME.2018.2798701).
- [11] F. Sultana, A. Sufian, and P. Dutta, "Evolution of image segmentation using deep convolutional neural network: A survey," *Knowl.-Based Syst.*, vols. 201–202, pp. 1–38, May 2020, doi: [10.1016/j.knosys.2020.106062](https://doi.org/10.1016/j.knosys.2020.106062).
- [12] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," in *Proc. Comput. Vis. Pattern Recognit.*, 2020, pp. 1–23. [Online]. Available: <https://arxiv.org/pdf/2001.05566.pdf>
- [13] A. A. Ewees, M. A. Elaziz, M. A. A. Al-Qaness, H. A. Khalil, and S. Kim, "Improved artificial bee colony using sine-cosine algorithm for multi-level thresholding image segmentation," *IEEE Access*, vol. 8, pp. 26304–26315, 2020, doi: [10.1109/ACCESS.2020.2971249](https://doi.org/10.1109/ACCESS.2020.2971249).
- [14] Z. Zhang and J. Yin, "Bee foraging algorithm based multi-level thresholding for image segmentation," *IEEE Access*, vol. 8, pp. 16269–16280, 2020, doi: [10.1109/ACCESS.2020.2966665](https://doi.org/10.1109/ACCESS.2020.2966665).
- [15] Z. Yan, J. Zhang, Z. Yang, and J. Tang, "Kapur's entropy for underwater multilevel thresholding image segmentation based on whale optimization algorithm," *IEEE Access*, early access, Jun. 29, 2020, doi: [10.1109/ACCESS.2020.3005452](https://doi.org/10.1109/ACCESS.2020.3005452).
- [16] J. Chen, B. Guan, H. Wang, X. Zhang, Y. Tang, and W. Hu, "Image thresholding segmentation based on two dimensional histogram using gray level and local entropy information," *IEEE Access*, vol. 6, pp. 5269–5275, 2018, doi: [10.1109/ACCESS.2017.2757528](https://doi.org/10.1109/ACCESS.2017.2757528).
- [17] X. Cao, T. Li, H. Li, S. Xia, F. Ren, Y. Sun, and X. Xu, "A robust parameter-free thresholding method for image segmentation," *IEEE Access*, vol. 7, pp. 3448–3458, 2019, doi: [10.1109/ACCESS.2018.2889013](https://doi.org/10.1109/ACCESS.2018.2889013).
- [18] H. Jia, K. Sun, W. Song, X. Peng, C. Lang, and Y. Li, "Multi-strategy emperor penguin optimizer for RGB histogram-based color satellite image segmentation using Masi entropy," *IEEE Access*, vol. 7, pp. 134448–134474, 2019, doi: [10.1109/ACCESS.2019.2942064](https://doi.org/10.1109/ACCESS.2019.2942064).
- [19] C. Kang, C. Wu, and J. Fan, "Lorenz curve-based entropy thresholding on circular histogram," *IEEE Access*, vol. 8, pp. 17025–17038, 2020, doi: [10.1109/ACCESS.2020.2964335](https://doi.org/10.1109/ACCESS.2020.2964335).
- [20] Y. Zou, H. Liu, and Q. Zhang, "Image bilevel thresholding based on stable transition region set," *Digit. Signal Process.*, vol. 23, no. 1, pp. 126–141, Jan. 2013, doi: [10.1016/j.dsp.2012.08.004](https://doi.org/10.1016/j.dsp.2012.08.004).
- [21] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, no. 4, pp. 620–630, May 1957, doi: [10.1103/physrev.106.620](https://doi.org/10.1103/physrev.106.620).
- [22] C. I. Chang, Y. Du, and J. Wang, "Survey and comparative analysis of entropy and relative entropy thresholding techniques," *IEE Proc.-Vis., Image Signal Process.*, vol. 153, no. 6, pp. 837–850, Dec. 2006, doi: [10.1049/ip-vis:20050032](https://doi.org/10.1049/ip-vis:20050032).
- [23] P. K. Sahoo and G. Arora, "A thresholding method based on two-dimensional Rényi's entropy," *Pattern Recognit.*, vol. 37, no. 6, pp. 1149–1161, Jun. 2004, doi: [10.1016/j.patcog.2003.10.008](https://doi.org/10.1016/j.patcog.2003.10.008).
- [24] M. Masi, "A step beyond Tsallis and Rényi entropies," *Phys. Lett. A*, vol. 338, nos. 3–5, pp. 217–224, May 2005, doi: [10.1016/j.physleta.2005.01.094](https://doi.org/10.1016/j.physleta.2005.01.094).

- [25] 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, doi: [10.1016/S0734-189X\(85\)90156-2](https://doi.org/10.1016/S0734-189X(85)90156-2).
- [26] P. K. Sahoo, C. Wilkins, and J. Yeager, "Threshold selection using Rényi entropy," *Pattern Recognit.*, vol. 30, no. 1, pp. 71–84, Jan. 1997, doi: [10.1016/S0031-3203\(96\)00065-9](https://doi.org/10.1016/S0031-3203(96)00065-9).
- [27] M. P. de Albuquerque, I. A. Esquef, A. R. G. Mello, and M. P. de Albuquerque, "Image thresholding using Tsallis entropy," *Pattern Recognit. Lett.*, vol. 25, no. 9, pp. 1059–1065, Jul. 2004, doi: [10.1016/j.patrec.2004.03.003](https://doi.org/10.1016/j.patrec.2004.03.003).
- [28] H. Zhang, "One-dimensional Arimoto entropy threshold segmentation method based on parameter optimization," in *Proc. Int. Conf. Appl. Inform. Commun.*, Wuhan, China, 2010, pp. 573–581.
- [29] R. R. Abdiel, R. H. M. Alejandro, and H. C. Gerardo, "Determining the entropic index q of Tsallis entropy in images through redundancy," *Entropy*, vol. 18, no. 8, pp. 1–14, Aug. 2016, doi: [10.3390/e18080299](https://doi.org/10.3390/e18080299).
- [30] F. Nie, P. Zhang, J. Li, and D. Ding, "A novel generalized entropy and its application in image thresholding," *Signal Process.*, vol. 134, pp. 23–34, May 2017, doi: [10.1016/j.sigpro.2016.11.004](https://doi.org/10.1016/j.sigpro.2016.11.004).
- [31] D. Liu, Z. Jiang, and H. Feng, "A novel fuzzy classification entropy approach to image thresholding," *Pattern Recognit. Lett.*, vol. 27, no. 16, pp. 1968–1975, Dec. 2006, doi: [10.1016/j.patrec.2006.05.006](https://doi.org/10.1016/j.patrec.2006.05.006).
- [32] A. D. Brink, "Thresholding of digital images using two-dimensional entropies," *Pattern Recognit.*, vol. 25, no. 8, pp. 803–808, Aug. 1992, doi: [10.1016/0031-3203\(92\)90034-G](https://doi.org/10.1016/0031-3203(92)90034-G).
- [33] A. S. Abutableb, "Automatic thresholding of gray-level pictures using two-dimensional entropy," *Comput. Vis. Graph. Image Process.*, vol. 47, no. 1, pp. 22–32, Jul. 1989, doi: [10.1016/0734-189x\(89\)90051-0](https://doi.org/10.1016/0734-189x(89)90051-0).
- [34] N. R. Pal and S. K. Pal, "Entropic thresholding," *Signal Process.*, vol. 16, no. 2, pp. 97–108, Feb. 1989, doi: [10.1016/0165-1684\(89\)90090-x](https://doi.org/10.1016/0165-1684(89)90090-x).
- [35] P. K. Sahoo and G. Arora, "Image thresholding using two-dimensional Tsallis–Havrda–Charvát entropy," *Pattern Recognit. Lett.*, vol. 27, no. 6, pp. 520–528, Apr. 2006, doi: [10.1016/j.patrec.2005.09.017](https://doi.org/10.1016/j.patrec.2005.09.017).
- [36] X. Zhang, Z. Xue, and Y. Zheng, "Fast and precise two-dimensional Rényi entropy image thresholding," *Pattern Recogn. Artif. Intell.*, vol. 25, no. 3, pp. 411–418, Jun. 2012, doi: [10.16451/j.cnki.issn1003-6059.2012.03.005](https://doi.org/10.16451/j.cnki.issn1003-6059.2012.03.005).
- [37] M. Kelbert and Y. Suhov, *Information Theory and Coding by Example*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
- [38] T. Lindeberg, "Scale-space for discrete signals," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 3, pp. 234–254, Mar. 1990, doi: [10.1109/34.49051](https://doi.org/10.1109/34.49051).
- [39] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. London, U.K.: Pearson, 2017.
- [40] Y. Zou. (Mar. 2020). *Test Image Set and Segmentation Reference Set*. [Online]. Available: <https://share.weiyun.com/5pdwtWY>
- [41] M. Beneš and B. Zitová, "Performance evaluation of image segmentation algorithms on microscopic image data," *J. Microsc.*, vol. 257, no. 1, pp. 65–85, Jan. 2015, doi: [10.1111/jmi.12186](https://doi.org/10.1111/jmi.12186).
- [42] H. Cai, Z. Yang, X. Cao, W. Xia, and X. Xu, "A new iterative triclass thresholding technique in image segmentation," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1038–1046, Mar. 2014, doi: [10.1109/TIP.2014.2298981](https://doi.org/10.1109/TIP.2014.2298981).
- [43] R. Bemis. (Mar. 2020). *Thresholding Tool*. [Online]. Available: <http://cn.mathworks.com/matlabcentral/fileexchange/6770-thresholding-tool>



YAOBIN ZOU received the Ph.D. degree in computer science and technology from the Huazhong University of Science and Technology, in 2011. He is currently an Associate Professor with the Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering and the College of Computer and Information Technology, China Three Gorges University. His research interests include digital image processing and pattern recognition.



JINYU ZHANG received the B.S. degree from China Three Gorges University, Yichang, China, in 2019. She is currently pursuing the master's degree. Her current research interest includes digital image processing.



MANISH UPADHYAY received the B.S. degree from the Shri Govindram Seksaria Institute of Technology and Science, Indore, India, in 2019. He is currently pursuing the master's degree. His current research interest includes digital image processing.



SHUIFA SUN received the M.S. degree from the Zhejiang University of Technology, China, in 2002, and the Ph.D. degree from Zhejiang University, in 2005. His research interests include intelligent information processing, computer vision, and multimedia information processing.



TINGYAO JIANG received the Ph.D. degree from the School of Computer Science and Technology, Huazhong University of Science and Technology, in 2003. He is currently a Professor with the College of Computer and Information Technology, China Three Gorges University. His research interests include pattern recognition and mobile computing.

...