Image segmentation based on equivalent three-dimensional entropy method and artificial fish swarm optimization algorithm

Xiangxiao Lei
Honglin Ouyang
Lijuan Xu

Image segmentation based on equivalent three-dimensional entropy method and artificial fish swarm optimization algorithm

Xiangxiao Lei,‡ Honglin Ouyang,§ and Lijuan Xu
Hunan University, College of Electrical and Information Engineering, Changsha, China
Changsha Social Work College, School of Electronic Information Engineering, Changsha, China

Abstract. To improve the timeliness of the three-dimensional (3-D) maximum entropy method, an image segmentation method based on equivalent 3-D entropy and artificial fish swarm optimization algorithm is proposed. An equivalent 3-D entropy method without logarithmic operation is developed, and its equivalence is proved theoretically. The optimal threshold is determined based on the artificial fish swarm optimization algorithm so as to avoid exhaustive search and improve algorithm efficiency. The experimental results demonstrate that the proposed method is more time-efficient than the traditional 3-D entropy method and the equivalent 3-D entropy method without affecting segmentation. Compared with the one-dimensional entropy method and the two-dimensional entropy method, it is obviously superior in noise immunity and detail preservation.

Keywords: image segmentation; maximum entropy; three-dimensional histogram; optimization algorithm; artificial fish-swarm algorithm.

1 Introduction
Image segmentation refers to the technique of dividing an image into nonoverlapping homogeneous regions and extracting the objects of interest. It is the basis of image analysis and computer vision. Recently, considerable research on image segmentation has been conducted, and various segmentation algorithms have been proposed. In general, these algorithms can be divided into: edge detection segmentation algorithms, region segmentation algorithms, threshold segmentation algorithms, and fuzzy segmentation algorithms. Among them, threshold segmentation algorithms have been widely studied and applied owing to their simplicity and effectiveness.

The purpose of threshold segmentation is to separate the object from the background as much as possible and retain maximum image information entropy. Therefore, the essence of an entropy-based threshold segmentation algorithm is to use different entropy functions to establish the objective function, with the segmentation threshold as the independent variable, and then to determine the threshold where the objective function takes the extreme value according to a certain criterion. The traditional one-dimensional (1-D) maximum entropy method only considers the gray-scale information of pixels and neglects their spatial correlation. Although the processing speed is high, the noise immunity is poor. In this regard, by using the information contained in the two-dimensional (2-D) histogram of the gray-scale distribution of image pixels and the average gray-scale distribution of their neighborhood, Abutaleb proposed a 2-D maximum entropy method that takes into account the neighborhood mean value information. Although the 2-D maximum entropy method enhances the algorithm’s ability to suppress noise and improves segmentation accuracy, the introduction of neighborhood gray scale mean value information increases running time and resource consumption.

To improve efficiency, Zhang et al. proposed a 2-D maximum entropy threshold segmentation algorithm by combining the artificial fish swarm optimization algorithm with Abutaleb entropy, which has satisfactory noise immunity and convergence performance. To obtain a better threshold, a combination of optimization algorithms is usually employed, such as the particle swarm optimization algorithm, the artificial swarm optimization algorithm, and the artificial fish swarm optimization algorithm. To further improve the performance of the algorithm, Lei et al. introduced the neighborhood gray-scale median based on the 2-D histogram; moreover, they proposed a three-dimensional (3-D) cross-entropy method as well as a fast recursive method to improve efficiency. It achieved a better segmentation effect.

Compared with the 2-D maximum entropy method, the 3-D maximum entropy method has better segmentation quality; however, algorithm complexity greatly increases. To some extent, the efficiency of the algorithm has been improved by the aforementioned recurrence method and an intelligent optimization algorithm, but the entropy calculation still involves a logarithmic operation. Thus, the algorithm efficiency should be improved. Accordingly, in this study, an equivalent 3-D entropy method is proposed by establishing an equivalent objective function. The method can effectively simplify the calculation and improve efficiency. In view of the searching performance of the artificial fish swarm optimization algorithm, a new threshold calculation method is developed based on the proposed objective function so as to further enhance the timeliness of the algorithm. The experimental result shows that this method not only yields segmentation results that are equivalent to
those obtained by the logarithmic method but also improves the efficiency of the algorithm.

2 Three-Dimensional Maximum Entropy Method

Let $I$ be an image of size $M \times N$, and let $f(x, y)$, $f(x, y) \in \{0, 1, \ldots, L-1\}$, be the gray-scale value of the pixel $(x, y)$. The neighborhood gray-scale mean value $g(x, y)$ and gray-scale median $h(x, y)$ of a $K \times K$ neighborhood of the pixel $(x, y)$ can be defined as follows:

$$g(x, y) = \frac{\sum_{i=-\frac{K}{2}}^{\frac{K}{2}} \sum_{j=-\frac{K}{2}}^{\frac{K}{2}} f(x+i, y+j)}{K^2}.$$  \hspace{1cm} (1)

$$h(x, y) = \text{med}\left\{ \sum_{i=-\frac{K}{2}}^{\frac{K}{2}} \sum_{j=-\frac{K}{2}}^{\frac{K}{2}} f(x+i, y+j) \right\}.$$ \hspace{1cm} (2)

$f(x, y)$, $g(x, y)$, and $h(x, y)$ can define a triple $(i, j, k)$. For example, if $n_{ijk}$ represents the frequency at which the pixel with gray-scale value $i$, neighborhood gray-scale mean value $j$, and neighborhood gray value $k$ appears in the image $I$, then its frequency $p_{ijk}$ is

$$p_{ijk} = \frac{n_{ijk}}{M \times N}.$$ \hspace{1cm} (3)

Based on the definition of Shannon entropy, the 3-D discrete entropy can be defined as

$$H = -\sum_{i} \sum_{j} \sum_{k} p_{ijk} \ln p_{ijk}.$$ \hspace{1cm} (4)

Based on the 3-D histogram shown in Fig. 1, the 3-D maximum entropy method divides the image into the objective area 0 and the background area 1 by using the threshold $(s, t, q)$. Areas 2 to 7 are the edges and noise regions. Clearly, it can be seen that the 3-D maximum entropy method uses the maximum amount of object and background information in the image, thus preserving details as much as possible.

According to the definition of entropy, the total entropy of the image $I$ is

$$H(s, t, q) = -\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \sum_{k=0}^{L-1} p_{ijk}^0 \ln p_{ijk}^0 - \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} \sum_{k=q+1}^{L-1} p_{ijk}^1 \ln p_{ijk}^1.$$ \hspace{1cm} (5)

where $L$ is gray levels, $s$ is gray value of image pixels, $t$ is neighborhood mean, $q$ is neighborhood median, $p_{ijk}^0 = \frac{\sum_{i=j-k}^{j+k-1} \sum_{j=s}^{s-1} \sum_{k=q}^{q-1} p_{ijk}}{(L-s-1)(L-1)(L-q-1)}$, $p_{ijk}^1 = \frac{\sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} \sum_{k=q+1}^{L-1} p_{ijk}}{(L-s-1)(L-t-1)(L-q-1)}$. According to the principle of maximum entropy, the optimal threshold value is

$$(s^*, t^*, q^*) = \text{Arg}\left\{ \max_{0 \leq (s, t, q) \leq L-1} [H(s, t, q)] \right\}.$$ \hspace{1cm} (6)

3 Equivalent Three-Dimensional Entropy

3.1 Equivalent Three-dimensional Entropy Method and Equivalence Proof

According to Eq. (5), there are a large number of logarithmic operations in the process of calculating the 3-D entropy. However, the logarithmic operation efficiency is low, which seriously restricts the application of the algorithm. Maximum entropy means the kind of situation in which the probability distribution is the most uniform. And in the ideal situation, $p_{ijk}^0$ is infinitely close to $\frac{1}{s^*q^*}$, and $p_{ijk}^1$ is close to $\frac{1}{(L-t-1)(L-q-1)}$. In view of this, a new 3-D maximum entropy method is proposed in this study. Its efficiency is improved by replacing the logarithmic operations with more efficient subtraction operations. The equivalence between this algorithm and the traditional 3-D maximum entropy method is now proved theoretically.

**Definition 1.** The equivalent 3-D entropy is defined as follows:

$$H_1(s, t, q) = -\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \sum_{k=0}^{L-1} \left( p_{ijk}^0 - \frac{1}{s^*q^*} \right)^2 + \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} \sum_{k=q+1}^{L-1} \left( \frac{1}{(L-s-1)(L-t-1)(L-q-1)} - p_{ijk}^1 \right)^2.$$ \hspace{1cm} (7)

By Eq. (6), the optimal threshold value $(s^*, t^*, q^*)$ is

---

**Fig. 1** 3-D histogram and its division. (a) 3-D histogram area division, (b) division of areas 0, 2, 3, and 4, and (c) division of areas 1, 5, 6, and 7.
\[ (s^*, t^*, q^*) = \text{Arg}\left\{ \max_{0 \leq (s, t, q) \leq L} [H_1(s, t, q)] \right\}. \]  

(8)

**Theorem 1.** The optimal threshold value calculated by Eq. (6) is equal to that calculated by Eq. (8). Namely, the proposed 3-D entropy method is equivalent to the traditional 3-D entropy method.

**Proof.**

\[ \sum_{i=0}^{s} \sum_{j=0}^{t} \sum_{k=0}^{q} p_{ijk}^0 = 1. \]

(9)

According to Lagrange Multiplier Method,\(^{11}\) constants are introduced, then it can be obtained from Eq. (5):

\[
H(s, t, q) = -\sum_{i=0}^{s} \sum_{j=0}^{t} \sum_{k=0}^{q} p_{ijk}^0 \ln p_{ijk}^0
- \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{t} \sum_{k=q+1}^{L-1} p_{ijk}^1 \ln p_{ijk}^1
+ \alpha \left( \sum_{i=0}^{s} \sum_{j=0}^{t} \sum_{k=0}^{q} p_{ijk}^0 \ln p_{ijk}^0 - 1 \right)
+ \beta \left( \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{t} \sum_{k=q+1}^{L-1} p_{ijk}^1 - 1 \right).
\]

(10)

Take the partial derivatives of \( p_{ijk}^0 \) and \( p_{ijk}^1 \), respectively, and let the derivatives be 0, which can get

\[
\frac{\partial H}{\partial p_{ijk}^0} = \sum_{i=0}^{s} \sum_{j=0}^{t} \sum_{k=0}^{q} (\alpha - \ln p_{ijk}^0) = 0
\]

\[
\frac{\partial H}{\partial p_{ijk}^1} = \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{t} \sum_{k=q+1}^{L-1} (\beta - \ln p_{ijk}^1) = 0.
\]

(11)

Solve the equation set shown in the Eq. (11), which can get

\[
p_{ijk}^0 = e^{\alpha - 1}
p_{ijk}^1 = e^{\beta - 1}.
\]

(12)

It can be obtained from Eqs. (9) and (12)

\[
e^{\alpha - 1} = \frac{1}{\lambda^{s} \prod_{i=s+1}^{L-1} \prod_{j=s+1}^{t} \prod_{k=q+1}^{L-1} p_{ijk}^0}
\]

\[
e^{\beta - 1} = \frac{1}{\lambda^{s} \prod_{i=s+1}^{L-1} \prod_{j=s+1}^{t} \prod_{k=q+1}^{L-1} p_{ijk}^1}.
\]

(13)

Therefore,

\[
p_{ijk}^0 = \frac{1}{Q(s, t, q)}
p_{ijk}^1 = \frac{1}{Q(s, t, q)}.
\]

(14)

It is known from the above solving process, \( \frac{1}{Q(s, t, q)} \) is the maximum value of Eq. (5). And it is also the maximum point of \( H_1(s, t, q) \).

Therefore, the optimal threshold value calculated according to Eq. (6) is equal to that calculated according to Eq. (8). Namely, the equivalent 3-D entropy method is equivalent to the traditional 3-D entropy method.

### 3.2 Artificial Fish-Swarm Algorithm

According to Eq. (7), for any threshold vector \((s, t, q)\), the entropy is calculated through \( stq + (L - s - 1) (L - t - 1) (L - q - 1) \) summations. The complexity of this calculation is \( o(L^3) \). In addition, \( 0 \leq (s, t, q) \leq L - 1 \). Hence, for any image, the complexity of the calculation is \( o(L^3) \). It can be seen that if the optimal threshold of Eq. (7) is obtained by the exhaustive search method, the algorithm is relatively inefficient. Intelligent optimization algorithms can effectively avoid exhaustive calculation and improve efficiency. They are widely used to solve complex problems.\(^{15-17}\) In view of this, Dai et al.\(^{17}\) put forward an optimization strategy from bottom to top, which is called artificial fish swarm optimization algorithm (AFSA), by analyzing the behavioral characteristics of shoals and adopting the model of autonomous animation. AFSA exhibits high convergence rate, high robustness, and strong global searching ability.\(^{17}\)

In view of the high convergence rate and strong global searching capability of AFSA, in this study, AFSA is used to determine the optimal threshold value in Eq. (7). In the 3-D gray-scale space \( Q = \{(s, t, q) | 0 \leq s, t, q \leq L - 1 \} \), an artificial fish individual model for threshold value calculation is constructed. Each artificial fish represents a potential threshold solution \((s, t, q)\). The behavior types of artificial fish are also defined, namely, foraging behavior, swarm behavior, following behavior, and random behavior.\(^{17,18}\)

1. **Foraging behavior.** Let the current state of the artificial fish be \( X_i \). A state \( X_j \) is randomly selected within its range of perception. If in the problem of determining the maximum it holds that \( Y_i < Y_j \), a step is taken in that direction; otherwise, a state \( X_j \) is selected again, and it is determined whether the forward condition is met. If the number of repeated attempts has reached Try-max, the maximum number of predetermined attempts, then a step is taken randomly:

\[
X_{i_{\text{next}}} = \begin{cases} X_i + \text{rand}() \cdot \frac{X_j - X_i}{\|X_j - X_i\|}, & Y_i < Y_j, \\ X_i + \text{rand}(), & Y_i > Y_j, \end{cases}
\]

(15)

where \( \text{rand}() \) is a random number in \([0,1]\).

2. **Swarm behavior.** It is assumed that the current state of the artificial fish is \( X_i \), and the concentration of food is \( Y_i \). Then, the number \( n_i \) and the central position \( X_c \) of the current neighborhood are considered \((d_{ij} < \text{visible})\). If \( Y_i / n_i > \hat{Y}_{c} \), then there is sufficient food, and the shoal is not crowded. In this case, a step is taken toward the center; otherwise, foraging behavior is performed:
\[
\begin{align*}
X_{i, \text{next}} &= X_i + \text{rand}(\delta \cdot (\frac{X_{\text{max}} - X_i}{\|X_{\text{max}} - X_i\|}), Y_{\text{c}}) , \\
& \quad \text{ForagingBehavior}, Y_{i, n} < \delta Y_i \\
& \quad \text{otherwise, foraging behavior is performed:} \\
\{ X_{i, \text{next}} &= X_i + \text{rand}(\delta \cdot (\frac{X_{\text{max}} - X_i}{\|X_{\text{max}} - X_i\|}), Y_{\text{c}}) , \\
& \quad \text{ForagingBehavior}, Y_{\text{c}} > \delta Y_i \\
\} & \quad \text{Step 5: If the algorithm termination conditions are met, then the optimal threshold value } \chi^* = (s^*, r^*, q^*) \text{ is the output, and image segmentation is performed according to the threshold; otherwise, execution returns to step 3 and continues.}
\end{align*}
\]

The more the population size SN is, the richer the information representing the original image will be, the more accurate the segmentation threshold will get, the higher the optimization accuracy of the algorithm will become. But the increase of population will increase the complexity of the algorithm if other conditions remain unchanged. Reference 19 shows that the efficiency and accuracy will reach the peak value at the same time and the image segmentation effect will be the best when the population size is between 15 and 35. The maximum number of iterations is one of the conditions for the end of the optimization algorithm. It is set so that the algorithm can find the optimal solution under the prescribed conditions. With a large number of experiments, it can be proved that the algorithm can find the optimal solution when the maximum number of iterations is 30.

In this study, the population size SN is 20, and the maximum number of iterations is 30, the step length \( \eta \) is 10, and the crowd factor \( \delta \) is 0.8. Thus, the total number of summation operations is \( 20 \times 30 \times \{ \frac{1}{2} s_{i} + (L - s - 1)(L - i - 1)(L - q - 1) \} \), and the complexity of the algorithm is \( O(L^3) \), whereas the complexity of the exhaustion method is \( O(L^5) \). Thus, the computation time of this method is significantly reduced, which improves the efficiency of the algorithm.

### 4 Experimental Analysis

To test the segmentation performance of the proposed method, the following two sets of experiments were conducted using the four images shown in Fig. 2: (i) Quantitative analysis experiments regarding quantitative indicators of the segmentation effect. Two indicators, namely, threshold value and running time, were selected. A quantitative comparison was made among the traditional 3-D maximum entropy method, the equivalent 3-D maximum entropy method, and the AFSA-based equivalent 3-D maximum entropy method to analyze the effectiveness of the proposed method in improving efficiency. (ii) Qualitative analysis experiments regarding visual effects. The segmentation results of the traditional 1-D, 2-D, and 3-D maximum entropy methods were compared, and the effectiveness of the proposed method was qualitatively analyzed from a visual perspective. The experimental platform was a PC with 2-GB RAM and a Q8300 CPU, which was programmed using Matlab2010.

#### 4.1 Effectiveness Validation of Improvement Measures

To verify the equivalence of the proposed 3-D maximum entropy and the rationality of using AFSA to determine the optimal threshold value, the equivalent 3-D maximum entropy method based on AFSA was compared with the traditional 3-D maximum entropy method and the equivalent 3-D maximum entropy method. As the purpose of the experiment is to verify the equivalence between the solution by the artificial fish with the largest adaptive value is set as the current optimal \( \chi^* \).
improved method and the solution by the traditional 3-D maximum entropy method, it is necessary to compare the threshold values calculated by each method. To test the efficiency of the improved method, the running time of the algorithms should be compared. From the comparison results in Table 1, it can be seen that the threshold value obtained by the equivalent 3-D maximum entropy method is consistent with that obtained by the traditional 3-D maximum entropy method, which experimentally verifies Theorem 1. As the equivalent 3-D maximum entropy method can avoid the time-consuming logarithmic operation, the efficiency of the algorithm is greatly improved. As AFSA is used to determine the optimal threshold value in the AFSA-based equivalent 3-D maximum entropy method, which avoids exhaustive search, the efficiency of the algorithm is further improved. It can be seen that the equivalent 3-D maximum entropy method based on AFSA can give better consideration to the time factor and achieves the same segmentation effect.

### 4.2 Comparison of Segmentation Effect

To test the segmentation performance of the proposed method and verify its advantages, comparisons with the 1-D maximum entropy method, the 2-D maximum entropy method, and the 3-D maximum entropy were conducted. As shown in Figs. 3–6, the 1-D maximum entropy method only considers the pixel gray-scale information and fails to make the right decision. The 2-D and 3-D maximum entropy methods can consider the spatial information and make a more reasonable judgment. The equivalent 3-D maximum entropy method based on AFSA can give a more accurate segmentation effect.
consider the spatial correlation of pixels, which leads to poor noise immunity. By introducing the neighborhood mean value information, the 2-D maximum entropy method improves the noise immunity of the algorithm. However, as the probability of the main diagonal region is assumed to be \( \sim 1 \) during the threshold calculation process, the image information is lost, which leads to inaccurate image details. The traditional 3-D maximum entropy method considers median information and more neighborhood information on the basis of 2-D maximum entropy method, and the details of the image are relatively richer. In the proposed method, the optimal threshold is determined based on the 3-D histogram. The noise immunity of the algorithm is relatively strong. As there is no approximate calculation, the same segmentation effect as the maximum entropy method is achieved. The comparison results show that the proposed method exhibits strong antinoise performance and is able to better retain image details. Thereby, the advantages of the method are verified from the visual perspective.
5 Conclusions

In view of the low efficiency of the traditional 3-D maximum entropy method, an image segmentation method based on the equivalent 3-D entropy and AFSA was proposed. In this method, the optimal threshold is determined by the equivalent 3-D entropy equation shown in Eq. (8), which avoids the time-consuming logarithmic operation. The optimal threshold is calculated by AFSA, which avoids exhaustive search. The results of the quantitative comparison with the traditional 3-D maximum entropy method showed that the segmentation effect of the proposed method is the same as that of the traditional 3-D maximum entropy method, but the efficiency of the algorithm is greatly improved. The qualitative analysis of the 1-D maximum entropy method and the 2-D maximum entropy method showed that the method has strong antinoise performance and better retains image details. The qualitative and quantitative comparison and analysis demonstrated that the proposed method has high accuracy, strong noise immunity, and high speed and may be successfully applied in the segmentation of SAR and noisy images.

Acknowledgments

This work was supported by the National Natural Science Foundation of China, Grant No. 51677063. The authors have no relevant financial interests in the manuscript and no other potential conflicts of interest to disclose.

References