Comparative analysis of two leading evolutionary intelligence approaches for multilevel thresholding

Zhengmao Ye* and Hang Yin

College of Science and Engineering,
Southern University,
Baton Rouge, LA 70813, USA
Email: zhengmao_ye@subr.edu
Email: hang_yin@subr.edu
*Corresponding author

Yongmao Ye

Broadcasting Department,
Liaoning Radio and Television Station,
Shenyang, 110003, China
Email: yeyongmao@hotmail.com

Abstract: The rapid advance of artificial intelligence has made complex image processing in real time possible. Multilevel thresholding has become a feasible way for image segmentation, even in the presence of poor contrast and external artefacts. Genetic algorithms (GAs) and particle swarm optimisation (PSO) are broadly recognised by far to be two dominating schemes which outperform classical ones on multilevel thresholding. Qualitative analysis can usually be applied to observe their superiority to all classical approaches. However, no convincing result is reached with respect to differences in performance between GAs and PSO. The existing segmentation practices are either examined by visual appeals exclusively, or evaluated quantitatively assuming perfect statistical distributions. To make thorough comparisons, comparative analysis of two leading multilevel thresholding approaches is conducted for true colour image segmentation. The information theory is also employed to analyse the outcomes of systematic approaches using diverse quantitative metrics from various aspects.

Keywords: artificial intelligence; evolutionary computation; genetic algorithms; particle swarm optimisation; PSO; global optimisation; multilevel thresholding; image segmentation; quantitative analysis; qualitative analysis; information theory.

Reference to this paper should be made as follows: Ye, Z., Yin, H. and Ye, Y. (2018) 'Comparative analysis of two leading evolutionary intelligence approaches for multilevel thresholding', *Int. J. Signal and Imaging Systems Engineering*, Vol. 11, No. 1, pp.20–30.

Biographical notes: Zhengmao Ye received his BS from Tianjin University, China; MS from Tsinghua University, China; 2nd MS and PhD from Wayne State University, USA, respectively. His research interests include control and optimisation with diverse control applications on electrical, mechanical, automotive and biomedical systems, as well as signal and image processing. Currently, he is a Professor in Electrical Engineering at Southern University, Baton Rouge, USA. He is a multi-disciplinary researcher who has the first author publications covering all leading control proceedings in three prestigious engineering societies (IEEE, ASME, SAE), specifically, IEEE (CDC, CCA, SMC, ACC, ISIC, FUZZ, IJCNN, CEC, CASE, ICCA, SOSE, MSC Congress, WCCI Congress), ASME (Congress, ICES, JRCICE), SAE (USA Congress, EAEC Congress, PFL Congress).

Hang Yin obtained his BS and MS from Tsinghua University, China, and 2nd MS and PhD from Louisiana State University, USA. He has also been conducting research at Tokyo University, Japan. He is the holder of US patent. Currently, he is an Assistant Professor in Civil Engineering at Southern University, Baton Rouge, USA. His major research interests include hydrology and hydrological modelling, water resources management, geotechnical engineering, geophysics as well as hydropower engineering.

Yongmao Ye graduated from Tianjin University, China. Currently, he serves as the Supervisor of Broadcasting Department, Liaoning Radio and Television Station, Shenyang, China.

This paper is a revised and expanded version of a paper entitled 'Qualitative and quantitative study of GAs and PSO based evolutionary intelligence for multilevel thresholding' presented at *International Symposium on Advanced Topics in Electrical Engineering*, Bucharest, Romania, 23–25 March , 2017.

1 Introduction

The natural evolution inspired artificial intelligence schemes have been widely applied to digital image segmentation via thresholding or clustering. Different from conventional principal component analysis (PCA) and independent component analysis (ICA) based clustering pattern recognition approaches, the classical Otsu thresholding is a straightforward means of automatic image thresholding which consists of two classes of histogram-based pixels to generate bi-level thresholding. For multilevel thresholding cases, however, the fundamental Otsu approach turns out to be time-consuming whose optimisation can only be reached by exhaustive search processes. Classical approaches are also lack of robustness against noises and consistency in pixels. The role of Otsu approach is sometimes limited to local optimisation. Clustering is another major approach in which the greater similarity occurs within the same cluster while the smaller similarity occurs in diversified clusters. Integration of level set and fuzzy C-Means clustering has been presented where fuzzy C-Means clustering is applied to identify the initial surface, which will be crucial to dynamic level set evolution and segmentation with changing boundaries in space and time. For both thresholding or clustering, evolution computation schemes are further carried out to achieve global optimisation for complex image processing, to enhance the accuracy of segmentation as well as to speed up the convergence rate to reach the optimal performance such as genetic algorithms (GAs), ant colony optimisation (ACO) and particle swarm optimisation (PSO) (Gonzalez and Woods, 2007; Duda et al., 2000; Engelbrecht, 2007; Otsu, 1979; Ye et al., 2007; Ye and Mohamadian, 2013; MacKay, 2005).

There are many GAs applications to solve medical signal and image processing problems (e.g., biomedical sample differentiation to determine optimal baselines). There are also various cases that employ GAs based medical image segmentation to be against poor contrast and external artefacts, which give rise to diffusing organ and tissue boundaries (Ye, 2005; Maulik, 2009). The self-adaptation Real-Coded GAs using simulated binary crossover could be combined with Taguchi methods to exploit potential offspring. Powerful exploration capability is reached through tournament selection by creating tournaments between two solutions. It yields solutions towards global optimisation which is far better than others in terms of solution quality, handling constraints and computation time for economic dispatch (Subbaraj and Rengaraj, 2011). Crossover is another important operator of the real-coded GAs. A case study is conducted where 16 crossover operators are compared using a set of 24 benchmark functions. Statistical analysis shows significant differences among all crossover operators where the efficiency relies on distinctive properties of the fitness functions (Picek et al., 2013).

PSO is another leading method while hybrid PSO has been used for optimal multilevel thresholding. Cooperative learning and comprehensive learning are applied for extra modifications. The former is to decompose any high-dimensional swarm into several one-dimensional ones. The latter is to avoid premature convergence in each 1D swarm. The PSO capability is strengthened further in terms of better fitness values by cloning the fitter particles (Maitra and Chatterjee, 2008). PSO can also be applied to thresholding where the grey scale image is converted into a binary image. The optimal set of thresholds is obtained once the objective function is minimised. PSO has been proved to be superior to GAs for multilevel thresholding (Duraisamy, 2010). A PSO-based method is also proposed to select multiple minimum cross entropy thresholds for complex image analysis rather than bi-level thresholding. It can be extended to multilevel thresholding such that the real-time complex image analysis is feasible. The convergence rate analysis validates its potential for real-time applications. This selection method supports the search for near-optimal thresholds (Yin, 2007).

A preliminary quantitative analysis on the impact of PSO multilevel thresholding has been recently conducted with respect to systematic sets of objective metrics based on the information theory, where the number of thresholding levels is directly relevant to all information metrics being defined. The extension to multilevel and multiband image thresholding is proposed using PSO. Fuzzy entropy based optimisation has been formulated for balanced histogram thresholding. To accelerate the convergence rate and shorten computation time, the PSO based intelligent schemes are presented to split digital images into regions and to identify contours for classification with optimality (Ye and Mohamadian, 2015). A fractional-order Darwinian particle swarm optimisation based method (DPSO) is brought out to exploit vast swarms of test solutions that may exist at any time on the testing hyperspectral and multispectral image. Its fractional derivative is used to control the convergence rate of particles. It has produced a significant statistical improvement in terms of both CPU time and the fitness value. The approach can integrate with support vector machine (SVM) for classification with better accuracy. Regions and contours generated by PSO based multilevel thresholding manifest good visual attraction and colour fidelity to intrinsic information (Ghamisi et al., 2014).

On the other hand, quantitative analysis using typical metrics has been successfully conducted for image

deblurring and image fusion (Ye and Mohamadian, 2012a, 2012b). To further evaluate the impact of multilevel thresholding on true colour digital images using the leading GAs and PSO schemes, in this paper, a set of well-defined objective metrics are taken into account for comparative analysis, including the discrete entropy, discrete energy, mutual information, dissimilarity, homogeneity, correlation and contrast as well as the dynamic range.

2 Multilevel thresholding

Otsu thresholding is a conventional segmentation approach which uses the variance of region homogeneity as the measure. As a bi-level thresholding approach, the within-class variance of two sets of pixels is to be minimised. Extension to multilevel thresholding is needed in most practices which groups image pixels into several classes. The 8-bit greyscale specifies the grey scale system with pixels on the display screen using an 8-bit value. The true colour specifies the colour system with pixels on the display screen using a 24-bit value. The latter is formulated by a composite of three independent RGB (red, green, and blue) light beams. In order to obtain optimal multilevel thresholding, artificial intelligence is applied to search for a set of global optimal thresholds at a fast convergence rate.

The fitness is represented by an objective function in evolution computing. For image thresholding, it is defined as the sum of multilevel discrete entropies with respect to the occurrence of each intensity level. The occurrence probability for each intensity level is estimated from its histogram, which is a percentage of the total count of individual intensity levels formulated as equation (1):

$$p(i) = \frac{h(i)}{\sum_{j=0}^{L-1} h(i)} \quad i = 1,\ 2,\ \ldots,\ (L-1) \tag{1}$$

$$\sum_{i=0}^{L-1} p(i) = 1, \tag{2}$$

where $h(i)$ is the histogram function and $p(i)$ is the probability function. For greyscale and true colour images, L is chosen to be 2^8 for the grey scale component and for each of three primary colour (R, G, B) components, respectively. The histogram is applied to display the image intensity content. It is also applied to formulate the cooccurrence matrix of relative frequencies. Here, multilevel thresholding is constructed as an N-Dimensional optimisation problem. N optimal thresholds (T_1, T_2, …, T_N) are solved by maximising the objective function shown in equations (3) and (4).

$$f\ [H_1, H_2, ..., H_N] = \sum_{i=0}^{N} H_i = \max(H_1 + H_2 + ... + H_N) \tag{3}$$

$$T^* = [T_1,\ T_2,\ ...,\ T_N] = \arg\max\left(\sum_{i=0}^{N} H_i\right) \tag{4}$$

where H_i ($i = 0$ to N) is the discrete entropy defined as equation (5). S_i ($i = 0$ to N) is the sum of the probability functions between any two thresholds, defined as equation (6).

$$H_0 = \sum_{j=0}^{T_1-1} \frac{p(i)}{S_0} \log_2 \frac{S_0}{p(i)} = -\sum_{j=0}^{T_1-1} \frac{p(i)}{S_0} \log_2 \frac{p(i)}{S_0}$$
$$\text{and } S_0 = \sum_{j=0}^{T_1-1} p(i)$$

$$H_1 = \sum_{j=T_1}^{T_2-1} \frac{p(i)}{S_1} \log_2 \frac{S_1}{p(i)} = -\sum_{j=T_1}^{T_2-1} \frac{p(i)}{S_1} \log_2 \frac{p(i)}{S_1}$$
$$\text{and } S_1 = \sum_{j=T_1}^{T_2-1} p(i) \tag{5–6}$$

$$H_K = \sum_{j=T_K}^{T_{K+1}-1} \frac{p(i)}{S_K} \log_2 \frac{S_K}{p(i)} = -\sum_{j=T_K}^{T_{K+1}-1} \frac{p(i)}{S_K} \log_2 \frac{p(i)}{S_K}$$
$$\text{and } S_K = \sum_{j=T_K}^{T_{K+1}-1} p(i)$$

$$H_N = \sum_{j=T_N}^{L-1} \frac{p(i)}{S_N} \log_2 \frac{S_N}{p(i)} = -\sum_{j=T_N}^{L-1} \frac{p(i)}{S_N} \log_2 \frac{p(i)}{S_N}$$
$$\text{and } S_N = \sum_{j=T_N}^{L-1} p(i).$$

3 Genetic algorithms (GAs)

Genetic algorithms start with a randomly generated population of a fixed size M. Initialisation of M individual populations is implemented between lower and upper limits of p_i ($p_{i,\min}$ and $p_{i,\max}$). Initialisation is conducted so as to select parents as equation (7), where μ_i is a uniformly distributed random number between 0 and 1.

$$p_i = p_{i,\ \min} + \mu_i \left(p_{i,\max} - p_{i,\min}\right) \quad i = 1,\ 2,\ \ldots,\ M \tag{7}$$

For binary coding, the whole population is substituted by the binary strings with finite binary bits to represent chromosomes. The actual length determines the maximum possible precision. For real coding, precision is determined by the fixed decimal point representation of real numbers. Real coding outperforms binary coding for constrained optimisation issues, where each chromosome is encoded as a string of floating point numbers. The selection process mimics the natural survival of those fittest creatures. Roulette wheel selection and tournament selection are the two main schemes to produce offspring. The one-armed wheel is spun multiple times in the roulette wheel algorithm to make a desired selection according to the probability distribution. Tournament selection is faster and less sensitive than roulette wheel selection which is chosen in context so as to avoid the premature convergence. Tournament sets are randomly selected with a population

size whose objective functions are then examined among each other to obtain the best, second best, third best fitness, and so on. Corresponding probabilities of fitness functions are simply chosen to be θ, $\theta(1–\theta)$, $\theta(1–\theta)^2$, and so on, without loss of generality.

Simulated binary crossover is applied subsequently to generate two offspring from parents. The spread factor α_i is defined as the ratio (8) of the absolute mismatch of offspring values to that of parent values. It is calculated as equation (9) by equating the area under the probability curves and a generated random number, where η_c is the crossover index.

$$\alpha_i = \left|(p_{i,2}^{k+1} - p_{i,1}^{k+1}) / (p_{i,2}^{k} - p_{i,1}^{k})\right| \tag{8}$$

$$\alpha_i = \begin{cases} (2*\text{rand}_i)^{\frac{1}{1+\eta_c}} & (0 \le \text{rand}_i \le 0.5) \\ (2-2*\text{rand}_i)^{\frac{-1}{1+\eta_c}} & (0.5 < \text{rand}_i \le 1) \end{cases} \tag{9}$$

The two offspring are computed by equation (10).

$$\begin{aligned} p_{i,1}^{k+1} &= (1+\alpha_i)p_{i,1}^{k} + (1-\alpha_i)p_{i,2}^{k} \\ p_{i,2}^{k+1} &= (1-\alpha_i)p_{i,1}^{k} + (1+\alpha_i)p_{i,2}^{k} \end{aligned} \tag{10}$$

The generated offspring are subject to polynomial mutation.

$$q_j^{k+1} = p_j^{k+1} + (p_j^U - p_j^L)\beta_j \tag{11}$$

In equation (11), p_j^U and p_j^L are the upper and lower boundaries of p_j, and β_j is a probability distribution defined as equation (12).

$$\beta_j == \begin{cases} (2\times\text{rand}_j)^{\frac{1}{1+\eta_m}} - 1 & (0 \le \text{rand}_j \le 0.5) \\ 1-(2-2\times\text{rand}_j)^{\frac{1}{1+\eta_m}} & (0.5<\text{rand}_j \le 1) \end{cases}, \tag{12}$$

where η_m is the non-negative mutation index. A random number between 0 and 1 also needs to be generated for mutation. Polynomial mutation differs from a uniform mutation in that it leads to a slight perturbation to prevent one from premature convergence. The probability of creating an offspring identical to the parents should be higher than that of non-identical one. As the generation update proceeds, the probability of creating an offspring closer to the parents gets higher and higher. The polynomial probability distribution is listed in equation (13).

$$P(\beta_j) = 0.5(1+\eta_m)(1-\left|\beta_j\right|)^{\eta_m}. \tag{13}$$

4 Particle swarm optimisation (PSO)

In an N-Dimensional search space, PSO can be introduced to solve a multi-dimensional optimisation problem. All particles travel through a search space to reach an optimal solution by gradually communicating and sharing information with the neighbours. Define the position and velocity vectors of the ith particle of the swarm as $X_i = (X_{i1}, X_{i2}, \ldots, X_{id})$ and $V_i = (V_{i1}, V_{i2}, \ldots, V_{id})$. The objective function encompasses the local best fitness position for each particle and global best fitness position across the entire swarm at each iteration. At cycle k, vectors of position, velocity, local best location reached by particle i, global best location reached by other neighbours are denoted as X_i^k, V_i^k, $P_{\text{best}i}^k$, $G_{\text{best}i}^k$, respectively. The position and velocity of each particle in the subsequent iteration will be updated by equations (14) and (15)

$$\begin{aligned} V_i^{k+1} = V_i^k + \{&w_1 \times \text{rand}_1^k (P_{\text{best}i}^k - X_i^k) \\ &+ w_2 \times \text{rand}_2^k \times (G_{\text{best}i}^k - X_i^k)\} \end{aligned} \tag{14}$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1}, \tag{15}$$

where w_1 and w_2 are two positive weights referred to as the cognitive parameter and social parameter, respectively. The former pulls each particle towards local best position while the latter pushes the particle towards the global best position. rand_1 and rand_2 are two uniformly distributed random variables between an interval [0,1]. The fitness in evolution computing is evaluated by the objective function (3), which is the sum of multilevel discrete entropies on a basis of occurrence of each intensity level corresponding to each pixel.

At each cycle, the global best position is updated among the entire swarm. Implicit interaction among particles is adopted in the PSO scheme to update local and global information. Thus the velocity and position of particles keep updating each time. The process continues step by step until a stagnation condition is met when an objective function reaches maxima. Sometimes, it is necessary to maintain the particle population in a swarm at a certain amount, thus the upper and lower bounds should be added. For this reason, a swarm with a very limited population should be deleted. However, it is possible that the particle travels at high velocity, which results in the velocity explosion of the particle. Thus DPSO is thus introduced with a constriction factor (λ) to limit the velocity which leads to the sub-optimal problem formulated as equations (16) and (17). Assumptions have been made in DPSO. The lifespan of swarm is subject to either extension or reduction when more or less fitness solutions could be located. The lifespan of swarm also simulates the natural survival case that is associated with the actual chance of producing offspring. Since DPSO has one parameter exclusively, it is less difficult to control than the PSO. DPSO also enhances the population diversity.

$$v_i^{k+1} = \lambda[v_i^k + \text{rand}^k \times (P_{\text{best}i}^k - x_i^k) + (1-\text{rand}^k) \times (G_{\text{best}i}^k - x_i^k)] \tag{16}$$

$$x_{id}^{i+1} = x_{id}^i + v_{id}^{i+1}. \tag{17}$$

5 Case studies on multilevel thresholding

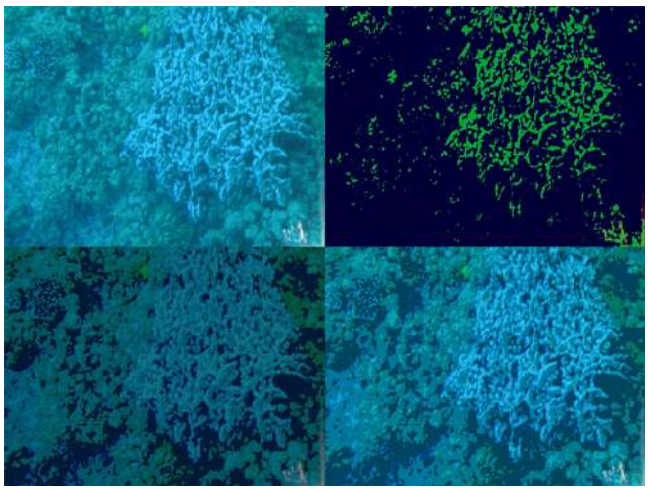
Both GAs and PSO are applied to multilevel thresholding of a set of true colour images. Comparisons of performance are

necessary for potential real-time processing. Both qualitative and quantitative analyses will be conducted. Without loss of generality, four types of true colour images are chosen including the low dynamic range still image, high dynamic range still image, low dynamic range kinetic image, and high dynamic range kinetic image. The kinetic image simply refers to the one generated by the fluctuation of attenuation caused by motion. A still image does not involve physical motion. A kinetic image involves physical movement of objects caused by the relative motion. Multilevel thresholding is conducted for each of the three primary colour components (red, green, blue). The additive outcome of three primary colour outputs after multilevel thresholding will be demonstrated for comparative study. In order to evaluate the impact of multilevel thresholding using real-coded GAs and DPSO, qualitative comparisons among all four sets of images selected will be made in this session. The corresponding quantitative study will be conducted in next session using a set of well-defined metrics.

5.1 Multiple level segmentation using real-coded gas

In Figures 1–4, real-coded GAs multilevel thresholding in terms of the low dynamic range still image, high dynamic range still image, low dynamic range kinetic image, and high dynamic range kinetic images are shown. The source images and additive outcomes based on two-level thresholding are listed on top-right, while additive outcomes based on three-level thresholding and five-level thresholding are listed on the bottom-left and bottom-right. It shows that the real-coded GAs can be well performed for multilevel thresholding.

Figure 1 Thresholding of low dynamic range still image (real-coded GAs) (see online version for colours)



Additive outcomes generated from lower level thresholding (2, 3) could barely reflect all detail features within a digital image (e.g., corner, edge, curve, boundary, connectivity). Higher level thresholding (5) produces better outcomes but information loss still exists. In general, except for special colour distortion cases, additive outcomes generated from multilevel thresholding give rise to a broader dynamic range and larger contrast than those of source images. Lower level thresholding could, however, produces sharper images than higher level thresholding.

Figure 2 Thresholding of high dynamic range still image (real-coded GAs) (see online version for colours)

Figure 3 Thresholding of low dynamic range kinetic image (real-coded GAs) (see online version for colours)

Figure 4 Thresholding of high dynamic range kinetic image (real-coded GAs) (see online version for colours)

5.2 Multilevel segmentation using DPSO

DPSO could also be used for multilevel thresholding and similar comparisons should be made. DPSO enhances the population diversity which has only one control parameter to adjust. In Figures 5–8, DPSO multilevel thresholding in terms of the low dynamic range still image, high dynamic range still image, low dynamic range kinetic image, and high dynamic range kinetic image are shown. The source images and additive outcomes based on two-level thresholding are listed on top-right, while additive outcomes based on three-level thresholding and five-level thresholding are listed on the bottom-left and bottom-right. It manifests that DPSO generally surpasses GAs on multilevel thresholding. The additive outcomes based on lower level thresholding (2, 3) barely reflect all detail features. Higher level thresholding (5) gives rise to significant improvement but information loss still exists.

Figure 5 Thresholding of low dynamic range still image (DPSO) (see online version for colours)

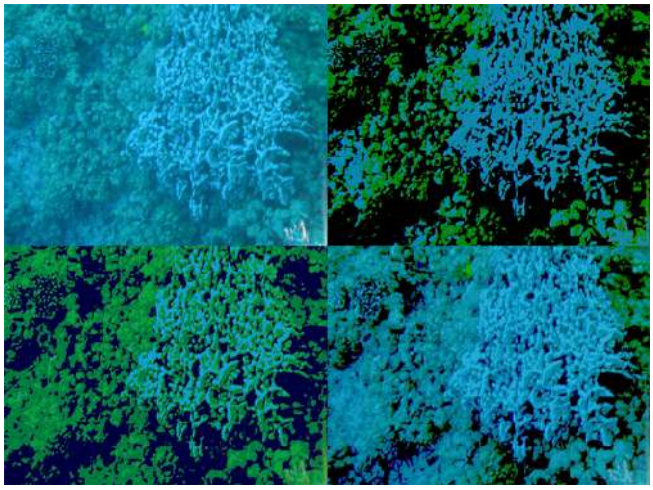


Figure 6 Thresholding of high dynamic range still image (DPSO) (see online version for colours)

Colour distortion seldom occurs when DPSO schemes are applied from visual appearance. It also shows that DPSO multilevel thresholding will produce outcomes with broader dynamic range and larger contrast than those of source images. At the same time, lower level thresholding always produces sharper images than higher level thresholding.

Figure 7 Thresholding of low dynamic range kinetic image (DPSO) (see online version for colours)

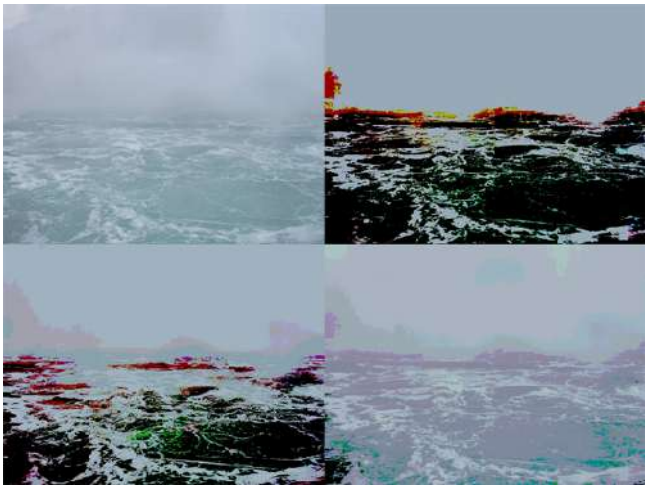


Figure 8 Thresholding of high dynamic range kinetic image (DPSO) (see online version for colours)

5.3 Multiple level segmentation using OTSU's method

GAs and PSO are widely recognised to represent two best multilevel thresholding approaches by far, which is, in fact, the scope of this research paper. GAs and PSO produce much better outcomes than all those classical approaches. Here the fundamental Otsu's thresholding has been selected as a typical example for multilevel thresholding. The simulation results of Otsu's multilevel thresholding for the low dynamic range still image and low dynamic range kinetic image are shown in Figures 9 and 10. The source images and additive outcomes based on two-level thresholding are listed on top-right, while additive outcomes based on three-level thresholding and five-level thresholding are listed on the bottom-left and bottom-right. The visual appeal in each multilevel case already shows that severe inconsistency in pixels does exist. The strength of the

classical Otsu's method actually lies in not multilevel thresholding but bi-level thresholding. For multilevel cases (e.g., 3, 4, 5, 6, …), Otsu's search-based optimisation (exhaustive search) is time-consuming. Also for multilevel cases, it barely guarantees global optimal results because its role is limited to local optimisation. A quantitative approach is even not necessary at all to compare Otsu with intelligent approaches (GAs or PSO), because the qualitative approach is already quite enough.

Figure 9 Thresholding of low dynamic range still image (Otsu's thresholding method) (see online version for colours)

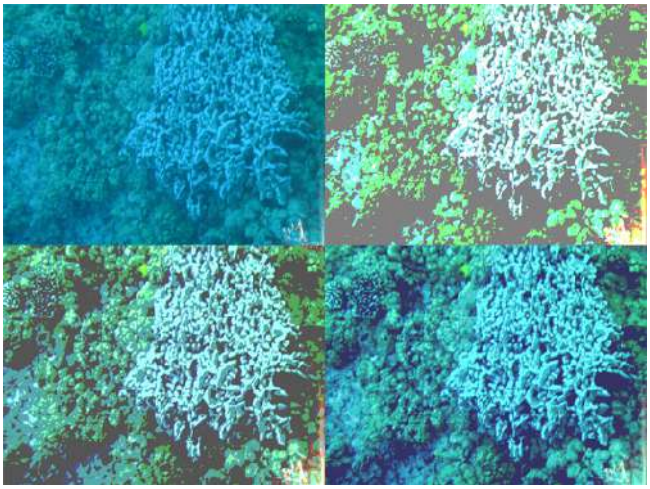
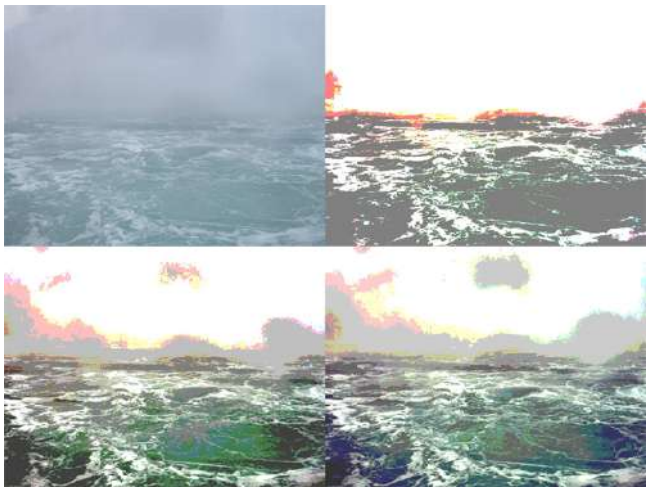


Figure 10 Thresholding of low dynamic range kinetic image (Otsu's thresholding method) (see online version for colours)



Comparing results generated from GAs and DPSO, there are quite some similarities. For instance, the outcomes from lower level thresholding can hardly manifest all important features (e.g., corner, edge, curve, boundary, connectivity). Higher level thresholding results in much better outcomes despite the fact that it still shows information loss. In the meanwhile, multilevel thresholding outcomes cover broader dynamic range and larger contrast. In particular, lower level thresholding generally produces sharper images with higher dynamic range than higher level thresholding. On the other hand, mismatches exist between outcomes from two schemes of GAs and PSO. GAs based multilevel thresholding generates colour distortion in several cases while colour distortion almost never occurs when DPSO based multilevel thresholding is applied. Even though there are numerous minor differences between the segmented images from GAs and DPSO, it is not entirely convincing to claim which one is the best scheme based on the qualitative analysis exclusively. Therefore, it is necessary to conduct quantitative analysis in order to figure out the best choice for potential real-time processing of complex true colour images.

6 Quantitative comparisons

In most cases, the qualitative analysis could be far from enough to completely characterise the merits and drawbacks of GAs and PSO based multilevel thresholding. Thus diversified metrics are introduced together for further quantitative analysis, including discrete entropy, discrete energy, mutual information, dissimilarity, homogeneity, correlation, contrast, and dynamic range. To depict the colour balance of three specific channels (red, green, blue) based on outcomes of multilevel thresholding via GAs and PSO, each of three colour channels is subject to quantitative analysis individually similar to the grey scale digital images, using the occurrence frequency of pixel counts from each of the 256 intensity levels.

6.1 Discrete entropy

The discrete entropy represents the average amount of information conveyed from each individual image. It is defined as the sum of products of the probability of outcomes and logarithm of the inverse of the probability (18), taking into account all possible outcomes in the event $\{x_1, x_2, \ldots, x_k\}$, where k is a count of intensity levels, $p(i)$ is the probability distribution.

$$H(x) = \sum_{i=1}^{k} p(i) \log_2 \frac{1}{p(i)} = -\sum_{i=1}^{k} p(i) \log_2 p(i). \tag{18}$$

6.2 Discrete energy

The discrete energy is also related to randomness which shows how the intensity level of each primary colour channel is distributed which are defined as equation (19), where $E(x)$ refers to the discrete energy with 256 bins and $p(i)$ refers to the probability distribution function for each channel (R, G, B) on a basis of the histogram counts. Segmentation itself involves energy minimisation process. For an arbitrary image with constant intensity, the discrete energy reaches the maximum value of one. The larger energy amount corresponds to a lower total number of intensity levels while the smaller one is corresponding to a higher total number of intensity levels.

$$E(x) = \sum_{i=1}^{k} p(i)^2. \tag{19}$$

6.3 Mutual information

Mutual information is a symmetric function that is formulated as equation (20), where $I(X; Y)$ represents the mutual information; $H(X)$ and $H(X|Y)$ are the entropy and conditional entropy values. It is interpreted as the information that Y can tell about X is equal to the uncertainty reduction of X due to the existence of Y. Zero mutual information means two images are independent.

$$\begin{aligned} I(X;Y) &= H(X) - H(X \mid Y) \\ &= \sum_{X,Y} p_{XY}(X,\ Y) \log_2 \frac{p_{XY}(X,\ Y)}{p_X(X) p_Y(Y)} \end{aligned} \tag{20}$$

6.4 Dissimilarity

Dissimilarity counts on local distance representation which is expressed as equation (21). It is described as the distance between two pixels in the cooccurrence matrix, where $g(i, j)$ is an element in the matrix at the coordinates i and j; M and N show the total numbers of pixels in the row and column of the digital image.

$$\Delta = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} g(i, j)\ \ | i - j |. \tag{21}$$

6.5 Homogeneity

Homogeneity is a direct measure of local variations which is formulated as equation (22). Lower values reflect more structural variations. Higher values of homogeneity reflect less structural variations.

$$S = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \frac{1}{1 + (i - j)^2} g(i, j). \tag{22}$$

6.6 Correlation

Correlation is used to analyse a linear dependency of intensity levels of neighbouring pixels which is defined as equations (23) and (24). In fact, it indicates the amount of local variations over an entire digital image being analysed.

$$R\ = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \frac{(i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} g(i, j) \tag{23}$$

$$\sigma_i = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (i - \mu_i)^2 g(i, j); \mu_i = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [i \times g(i, j)] \tag{24}$$

$$\sigma_j = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (j - \mu_j)^2 g(i, j); \mu_j = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [j \times g(i, j)],$$

where i and j are coordinates of the cooccurrence matrix; μ_i and σ_i are the horizontal mean and variance while μ_j and σ_j are the vertical mean and variance.

6.7 Contrast

Contrast is defined to measure variations of the intensity distribution in each colour channel which is formulated as equation (25), where g_{AVG} is the average intensity. The contrast is relevant to the diverse information content of objects that are visually distinguished from each other. It clarifies the intensity differences between individual features and background within the same scope.

$$T = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \frac{[g(i, j) - g_{AVG}]^2}{M \times N}. \tag{25}$$

6.8 Dynamic range

Definitions of dynamic range vary across diverse contexts. In regard to real-world scenes, it is referred to as the contrast ratio of the intensity values between the brightest region and darkest region. Sometimes, it can be also interpreted as the total number of distinctive pixels being covered by a digital image in the spatial domain. When the light intensity range is described by powers of two (log2), it will lead to measuring unit (f-stop) in the dynamic range. On the other hand, considering great variations, the dynamic range can also be formulated on a decade logarithmic scale (log10). There is no essential difference between definitions from the analysis point of view. In this paper, dynamic range is defined as the ratio of the largest measurable intensity level (pixel saturation) to smallest detectable intensity level (read-out noise) within the individual image, using binary logarithmic (log2) scale. It is simply formulated as equation (26), where I_{sat} refers to the saturation intensity and I_{min} refers to the dimmest non-zero detectable intensity.

$$DR = \log_2 \left(I_{\mathrm{sat}} / I_{\mathrm{min}} \right). \tag{26}$$

The RGB true colour model is efficient in producing diverse colours. The HSI colour model (hue, saturation, intensity) is able to decouple the intensity and colour information instead, such that the intensity component is independent of the colour components of hue and saturation. RGB and HSI models are interchangeable. For an image with RGB colour format, the corresponding hue, saturation and intensity components in the HSI space are expressed as equations (27)–(29). An exception in equation (27) is to use (360°–H) to substitute H when $B > G$. Here, the intensity is applied to compute the dynamic range.

$$H = \cos^{-1} \left\{ \frac{[(R - G) + (R - B)] / 2}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\} \tag{27}$$

$$S = 1 - \frac{3}{(R + G + B)} \{\min(R,\ G,\ B)\} \tag{28}$$

$$I = \left(R + G + B \right) / 3. \tag{29}$$

Based on four sets of digital images being selected, comprehensive quantitative comparisons will be made. Tables 1–4 are placed in an order of the low dynamic range still image, high dynamic range still image, low dynamic range kinetic image, and high dynamic range kinetic image, respectively. The quantitative metrics are listed with respect

to the source images and the addictive outcomes at two-, three-, and five-level thresholding via GAs and PSO.

Table 1 Metrics of low dynamic range still image

Metrics: R/G/B	*Source image*	*GA 2-Level*	*GA 3-Level*	*GA 5-Level*	*DPSO 2-Level*	*DPSO 3-Level*	*DPSO 5-Level*
Discrete entropy	4.9149	3.6526	3.8748	4.4765	1.7454	1.9890	3.6285
	6.7515	5.2097	5.9188	6.6497	2.3765	2.4412	6.3212
	6.9265	5.2155	6.1515	7.1592	2.1826	4.4814	7.3468
Discrete energy	0.0505	0.2452	0.1851	0.1022	0.5131	0.4047	0.2429
	0.0104	0.0901	0.0496	0.0145	0.4955	0.3556	0.0298
	0.0095	0.1553	0.0318	0.0131	0.5115	0.2525	0.0094
Mutual INFO		1.2623	1.0401	0.4384	3.1695	2.9258	1.2863
		1.5418	0.8327	0.1018	4.3750	4.3103	0.4303
		1.7111	0.7751	0.2327	4.7439	2.4451	0.4203
Dissimilarity	0.0462	0.1284	0.1276	0.1005	0.0061	0.0021	0.0960
	0.0948	0.2277	0.2454	0.1994	0.0411	0.0312	0.1957
	0.1062	0.2472	0.2579	0.2279	0.0812	0.1525	0.2489
Homogeneity	0.9769	0.9366	0.9366	0.9501	0.9970	0.9990	0.9524
	0.9526	0.9003	0.8878	0.9058	0.9795	0.9844	0.9104
	0.9469	0.8870	0.8774	0.8908	0.9594	0.9248	0.8808
Correlation	0.8927	0.5374	0.5497	0.6777	0.8801	0.9646	0.6832
	0.9523	0.9067	0.9397	0.9433	0.9143	0.9398	0.9478
	0.9439	0.5997	0.8707	0.9594	0.6808	0.9038	0.9382
Contrast	0.0466	0.1367	0.1318	0.1041	0.0061	0.0022	0.0998
	0.0951	0.3703	0.3499	0.2547	0.0418	0.0312	0.2783
	0.1065	0.3551	0.3215	0.2755	0.0812	0.1631	0.3012
Dynamic range	2.5472	6.4757	5.5935	2.7564	8.8202	7.0112	6.0321

Based on data of quantitative information metrics, the source images always possess largest discrete entropy, smallest discrete energy, highest correlation, and lowest contrast as well as shortest dynamic range. At each colour channel, it shows that the intensity levels of source images are more smoothly and evenly distributed than those of all thresholded images. Multilevel thresholding broadens dynamic range, reduces correlation, and enhances the contrast of digital images but the information loss (entropy, energy) occurs. Low-level segmentation suffers from most significant information loss with the lowest correlation, but it generates maximal expansion of dynamic range and contrast which is in favour of visual perception. Both dissimilarity and homogeneity show similar attributes where source images contain the largest homogeneity but smallest dissimilarity. The thresholding level itself takes a trivial role in both homogeneity and dissimilarity. The uncertainty reductions from two-, three-, and five-level thresholding additive outcomes are also computed and compared with source images. The higher the level of thresholding is, the smaller the mutual information is. At each thresholding level, the resulting mutual information via GAs is less than that via DPSO. It is shown that higher level thresholding has a relatively smaller dependency on the source image than lower level thresholding. In general, outcomes from DPSO cover greater dynamic range and higher contrast than those from GAs regardless of the actual thresholding level in all four cases. DPSO also produces better colour balance at the same time. Thus, DPSO is more suitable than GAs to show information content, to expand dynamic range, and to avoid colour distortion in real-time processing.

Table 2 Metrics of high dynamic range still image

Metrics: R/G/B	*Source image*	*GA 2-Level*	*GA 3-Level*	*GA 5-Level*	*DPSO 2-Level*	*DPSO 3-Level*	*DPSO 5-Level*
Discrete entropy	7.6027	6.3130	7.0613	7.4881	5.9827	6.7127	7.4360
	7.5200	6.2654	7.0454	7.4712	5.8911	6.4921	7.2723
	7.5417	5.8735	6.8370	7.4269	5.1534	6.6570	7.4143
Discrete energy	0.0059	0.0298	0.0133	0.0067	0.0398	0.0176	0.0073
	0.0063	0.0279	0.0106	0.0065	0.0301	0.0187	0.0086
	0.0063	0.0731	0.0242	0.0075	0.1531	0.0215	0.0077
Mutual INFO		1.2897	0.5414	0.1146	1.6200	0.8900	0.1667
		1.2546	0.4746	0.0488	1.6289	1.0279	0.2476
		1.6681	0.7047	0.1148	2.3882	0.8847	0.1273
Dissimilarity	0.4650	0.6869	0.6776	0.5358	0.3725	0.5656	0.5245
	0.4633	0.7011	0.6721	0.5372	0.4034	0.6576	0.5152
	0.4620	0.6479	0.6582	0.5475	0.3077	0.4828	0.5360
Homogeneity	0.7794	0.7427	0.7124	0.7529	0.8390	0.7542	0.7548
	0.7802	0.7413	0.7175	0.7533	0.8287	0.7121	0.7586
	0.7802	0.7525	0.7160	0.7464	0.8525	0.7921	0.7486
Correlation	0.8472	0.7443	0.7667	0.8184	0.7787	0.7861	0.8203
	0.8336	0.7393	0.7649	0.8158	0.7431	0.7723	0.8208
	0.8544	0.7192	0.7620	0.8138	0.8134	0.8247	0.8258
Contrast	0.586	1.6015	1.1990	0.7475	0.6255	0.9392	0.6969
	0.5838	1.6822	1.2201	0.7590	0.7083	1.0696	0.6783
	0.5756	1.4527	1.1164	0.7515	0.3712	0.8205	0.7032
Dynamic range	8.5793	9.5603	9.3923	8.9658	9.5793	9.5661	9.2360

Table 3 Metrics of low dynamic range kinetic image

Metrics: R/G/B	*Source image*	*GA 2-Level*	*GA 3-Level*	*GA 5-Level*	*DPSO 2-Level*	*DPSO 3-Level*	*DPSO 5-Level*
Discrete entropy	6.1059	3.2676	4.4352	5.3057	1.9963	2.5464	3.3711
	5.7784	3.3688	4.2705	4.8143	2.8315	3.2157	3.4277
	5.9951	3.2253	4.3577	5.2966	1.9834	2.5347	3.5097
Discrete energy	0.0170	0.2856	0.1689	0.0825	0.5740	0.4408	0.2756
	0.0208	0.2696	0.1674	0.1007	0.2922	0.2830	0.2316
	0.0175	0.2907	0.1726	0.0825	0.5808	0.4480	0.2674
Mutual INFO		2.8383	1.6707	0.8002	4.1096	3.5595	2.7347
		2.4096	1.5079	0.9641	2.9468	2.5627	3.3507
		2.7698	1.6374	0.6985	4.0118	3.4604	2.4855
Dissimilarity	0.0322	0.1461	0.1694	0.0936	0.0586	0.0789	0.0455
	0.0335	0.1738	0.2019	0.089	0.0385	0.0612	0.0092
	0.0320	0.1530	0.1756	0.0991	0.0500	0.0688	0.054

Table 3 Metrics of low dynamic range kinetic image (continued)

Metrics: R/G/B	*Source image*	*GA 2-Level*	*GA 3-Level*	*GA 5-Level*	*DPSO 2-Level*	*DPSO 3-Level*	*DPSO 5-Level*
Homogeneity	0.9839	0.9469	0.9338	0.9553	0.9707	0.9606	0.9773
	0.9833	0.9403	0.9238	0.9583	0.9812	0.9695	0.9954
	0.984	0.9466	0.9342	0.9528	0.9750	0.9656	0.9730
Correlation	0.9660	0.9213	0.9329	0.9549	0.9371	0.9377	0.9587
	0.9327	0.9076	0.9125	0.9253	0.9174	0.9227	0.9284
	0.9654	0.9141	0.9232	0.9442	0.9266	0.9293	0.9542
Contrast	0.0322	0.3684	0.3689	0.1148	0.0589	0.0789	0.0455
	0.0335	0.4875	0.4747	0.1173	0.0431	0.0621	0.0092
	0.0320	0.4172	0.4192	0.1231	0.0502	0.0689	0.0541
Dynamic range	0.9366	3.9019	1.2688	1.1242	9.3443	9.3080	9.2021

Table 4 Metrics of high dynamic range kinetic image

Metrics: R/G/B	*Source image*	*GA 2-Level*	*GA 3-Level*	*GA 5-Level*	*DPSO 2-Level*	*DPSO 3-Level*	*DPSO 5-Level*
Discrete entropy	7.6592	3.8207	4.5751	5.7938	3.8588	4.7088	6.3974
	7.6884	3.3046	4.6226	5.8926	3.3065	4.5875	5.1918
	7.6125	3.1308	4.7960	6.0123	3.5139	4.7629	6.3141
Discrete energy	0.0054	0.3820	0.1222	0.0602	0.3531	0.1205	0.0360
	0.0051	0.2699	0.1257	0.0790	0.2678	0.1248	0.0691
	0.0057	0.3156	0.1381	0.0591	0.2657	0.1306	0.0390
Mutual INFO		3.8386	3.0841	1.8654	4.8004	2.9504	1.2619
		4.3838	3.0658	1.7958	4.3819	3.1009	2.4966
		4.4817	2.8165	1.6002	4.0987	2.8496	1.2985
Dissimilarity	0.1526	0.2086	0.1700	0.2008	0.1651	0.1994	0.2103
	0.1522	0.1970	0.1744	0.2071	0.1891	0.2125	0.2600
	0.1494	0.1557	0.1635	0.2215	0.1840	0.1820	0.2033
Homogeneity	0.9249	0.9173	0.9239	0.9026	0.9399	0.9103	0.9036
	0.9250	0.9230	0.9226	0.8998	0.9320	0.9039	0.8796
	0.9265	0.9356	0.9265	0.8926	0.9310	0.9177	0.9075
Correlation	0.9717	0.9022	0.9442	0.9624	0.9357	0.9381	0.9518
	0.9716	0.9066	0.9450	0.9603	0.9308	0.9390	0.9406
	0.9754	0.9515	0.9530	0.9627	0.9538	0.9561	0.9595
Contrast	0.1643	0.4316	0.2592	0.2308	0.4033	0.3001	0.2983
	0.1640	0.4171	0.2726	0.2404	0.4732	0.3148	0.3568
	0.1611	0.2927	0.2463	0.2547	0.4264	0.2697	0.2953
Dynamic range	8.5793	9.2312	9.1649	8.9629	9.3264	9.1948	9.0279

7 Conclusions

GAs and DPSO have been presented for digital image multilevel thresholding. By qualitative analysis, both schemes could produce satisfactory outcomes from visual appeals which surpass the classical schemes significantly. However, qualitative analysis is not quite enough for detail comparisons between 2 leading schemes of GAs and PSO. Since quantitative analysis will generate more accurate analytical results, typical quantitative metrics of the discrete entropy, discrete energy, mutual information, dissimilarity, homogeneity, correlation, contrast and dynamic range are introduced to interpret outcomes from multiple points of view. It has been observed that the DPSO scheme slightly outperforms the GAs which conveys more information content, generates a broader dynamic range, and produces better colour balance with less occurrence of severe colour distortion. The segmented images at various individual thresholding levels will produce broader dynamic range scenes than those of source images. The segmented images also show higher contrast with less information content, especially at the low level. It is demonstrated clearly that evolutionary intelligence can be successfully implemented on nonlinear optimisation problems of real-time multilevel thresholding.

8 Future work

Multilevel thresholding represents a challenging practice of real-world problem solving involving numerous optimisation approaches. The optimal criteria vary across methods which are variance-based, gradient-based, Hessian-based, moment-based, cross-entropy-based, entropy-based, and so on. Basically, the methodologies can be primarily categorised into either classical approaches or intelligent approaches. A fundamental thresholding approach (i.e., classical Otsu's method) has been initially proposed for bi-level thresholding rather than multilevel thresholding. For multilevel thresholding, classical approaches tend to be very time consuming via exhaustive search which can barely guarantee global optimal results because the role is limited to local optimisation. Accordingly, severe inconsistency in image pixels does exist even from visual appeals. In general, qualitative analysis is already enough to make comparisons between classical approaches (e.g., Otsu, gradient descent, simulated annealing, moment preserving) and intelligent approaches (e.g., GAs, PSO, ACO) on multilevel thresholding. However, the quantitative approach turns out to be necessary to compare among all existing artificial intelligence (AI) based approaches (e.g., genetic algorithms, genetic programming, evolution strategies, PSO, ACO, bacteria foraging algorithms, artificial bees algorithm, invasive weed optimisation, memetic search, differential evolution search, artificial immune systems, stochastic tunnelling, Tabu Search, gravitational search algorithm), because qualitative analysis is no longer accurate and effective to observe the very slight differences among all AI approaches. At the same time, sub-optimality could be achieved using intelligent approaches, which gives rise to potential real-time applications. A more comprehensive overview to compare diversified multilevel thresholding approaches will be conducted in the near future. Especially, the quantitative approach will be used to determine the potential best multilevel thresholding approach in real time.

References

Duda, R., Hart, P. and Stork, D. (2000) *Pattern Classification*, 2nd ed., John Wiley & Sons, Hoboken, NJ, USA.

Duraisamy, S. (2010) 'A new multilevel thresholding method using swarm intelligence algorithm for image segmentation', *J. Intell. Learn. Syst. Appl.*, Vol. 2, pp.126–138.

Engelbrecht, A. (2007) *Computational Intelligence: An Introduction*, 2nd ed., John Wiley & Sons Press, Hoboken, NJ, USA.

Ghamisi, P., Couceiro, M., Martins, F., Fernando, M. and Benediktsson, J. (2014) 'Multilevel image segmentation based on fractional-order Darwinian particle swarm optimization', *IEEE Trans. Geosci. Remote Sens.*, Vol. 52, pp.2382–2394.

Gonzalez, R. and Woods, R. (2007) *Digital Image Processing*, 3rd ed., Prentice-Hall, New York, NY, USA.

MacKay, D. (2005) *Information Theory, Inference and Learning Algorithms*, University of Cambridge Press, Cambridge, UK.

Maitra, M. and Chatterjee, A. (2008) 'Hybrid cooperative–comprehensive learning based PSO algorithm for image segmentation using multilevel thresholding', *Expert Syst. Appl.*, Vol. 34, pp.1341–1350.

Maulik, U. (2009) 'Medical image segmentation using genetic algorithms', *IEEE Trans. Inf. Technol. Biomed.*, Vol. 13, pp.166–173.

Otsu, N. (1979) 'A threshold selection method from gray-level histograms', *IEEE Trans. Syst. Man Cybern.*, Vol. 9, pp.62–66.

Picek, S., Jakobovic, D. and Golub, M. (2013) 'On the recombination operator in real-coded genetic algorithms', *Proceedings of the 2013 IEEE Congress on Evolutionary Computation (CEC)*, 20–23 June, Cancun, Mexico.

Subbaraj, P. and Rengaraj, R. (2011) 'Enhancement of self-adaptive real-coded genetic algorithm using Taguchi method for economic dispatch problem', *Appl. Soft Comput.*, Vol. 11, pp.83–92.

Ye, Z. (2005) 'Artificial intelligence approach for biomedical sample characterization using Raman spectroscopy', *IEEE Trans. Autom. Sci. Eng.*, Vol. 2, pp.67–73.

Ye, Z. and Mohamadian, H. (2012a) 'Comparative and quantitative study of fundamental approaches on aerial image deblurring', *Int. J. Model. Optim.*, Vol. 2, pp.378–383.

Ye, Z. and Mohamadian, H. (2012b) 'Generating higher dynamic range scene via fusion integration based on DWT and SVM', *Proceedings of the 2012 IEEE International Conference on Computer Science and Engineering*, 14–17 July, Melbourne, Australia, pp.1173–1177.

Ye, Z. and Mohamadian, H. (2013) 'Enhancing decision support for pattern classification via fuzzy entropy based fuzzy C-means clustering', *Proceedings of the 2013 52nd IEEE Conference on Decision and Control*, 10–13 December, Florence, Italy, pp.7432–7436.

Ye, Z. and Mohamadian, H. (2015) 'Intelligent control based multilevel and multiband digital image thresholding using fuzzy entropy and particle swarm optimization', *Int. J. Adv. Comput. Eng. Netw.*, Vol. 3, pp.24–30.

Ye, Z., Ye, Y. and Mohamadian, H. (2007) 'Biometric identification via PCA and ICA based pattern recognition', *Proceedings of the 2007 IEEE International Conference on Control and Automation (ICCA 2007)*, 30 May–1 June, Guangzhou, China, pp.1600–1604.

Yin, P.Y. (2007) 'Multilevel minimum cross entropy threshold selection based on particle swarm optimization', *Appl. Math. Comput.*, Vol. 184, pp.503–513.