■ 614

Image Denoising Based on Artificial Bee Colony and BP Neural Network

Junping Wang, Dapeng Zhang*
Information Engineer Department, Henan Vocational and Technical Institute,
Zhengzhou 450046, Henan, China
*Corresponding author, e-mail: 55410445@qq.com

Abstract
Image is often subject to noise pollution during the process of collection, acquisition and transmission, noise is a major factor affecting the image quality, which has greatly impeded people from extracting information from the image. The purpose of image denoising is to restore the original image without noise from the noise image, and at the same time maintain the detailed information of the image as much as possible. This paper, by combining artificial bee colony algorithm and BP neural network, proposes the image denoising method based on artificial bee colony and BP neural network (ABC-BPNN), ABC-BPNN adopts the "double circulation" structure during the training process, after specifying the expected convergence speed and precision, it can adjust the rules according to the structure, automatically adjusts the number of neurons, while the weight of the neurons and relevant parameters are determined through bee colony optimization. The simulation result shows that the algorithm proposed in this paper can maintain the image edges and other important features while removing noise, so as to obtain better denoising effect.

Keywords: Image Denoising, Artificial Bee Colony, BP Neural Network

1. Introduction

The main media for information transmissionis voice and image. The quantity of information and intuitive contained in an image are unmatched by sound and words. However, image is easily interfered by all kinds of noise in the process of generation and transmission, the quality of image will be damaged, which is very unfavorable to the subsequent higher-level image processing [1]. Therefore, at the pre-processing stage of image , it is quite necessary to conduct image denoising, in order to improve the SNR (signal to noise ratio) of image and highlight the desired features of image [2].

Today, the study on the theory and application of image denoising is still a very active research angle in the image processing field. Many researchers at both home and abroad have analyzed the statistical models and frequency distribution of the noise signal according to the signal characteristics, summarized and put forward many methods for image denoising. Divided according to the application domain, the image denoising algorithm can be divided into spatial domain method and transform domain method [3]. Divided according to the concrete applied theory, the image denoising algorithm can be divided into algorithm based on multi-resolution analysis, algorithm based on probability statistics theory, method based on nonlinear filtering theory, method based on partial differential equation theory, etc. A dilemma of denoising is how to the maintain image details as much as possible while lowering image noise. And the research focus is to explore whether the algorithm can sparsely represent the image information effectively by dealing with different image features while denoising [4].

Artificial neural network is a forefront subject as well as a hot research topic in image processing which has gained the international recognition currently. It has attracted more and more attention with its application in fields of image denoising, image compression, image edge detection, image integration, etc. In the field of swarm intelligence, the artificial bee colony is a random optimization search algorithm based on the heuristic method of population, specializing in finding solutions to spacial optimization problems [5]. This paper, by combining artificial bee colony and BP neural network, proposes the ABC-BPNN image denoising method. In this paper, the principle of image denoising is first described, then on basis of the analysis of bee colony algorithm and BP neural network, the BP neural network training process optimized by

artificial bee colony is put forward, and finally the simulation experiment and analysis are carried out.

2. Source of Noise

The quality of digital image will be affected by noise, and the noises are mainly generated during two processes: acquisition and transmission. In the image acquisition process, the imaging sensor will be affected by various factors, such as the external environment and the quality of the sensor itself, etc. Figure 1 shows the digitization process of the simulation image, a grayscale image can be represented by a two-dimensional function $f(x, y)$ with the size of $m \times n$ after being digitized, wherein, (x, y) represents the coordinate of the pixel in the original image, $f(x, y)$ represents the grayscale value of coordinate pixel, at this time, x、 y and amplitude f are discrete. The noise mainly comes from the collection, transmission and management stages during the image digitization process.
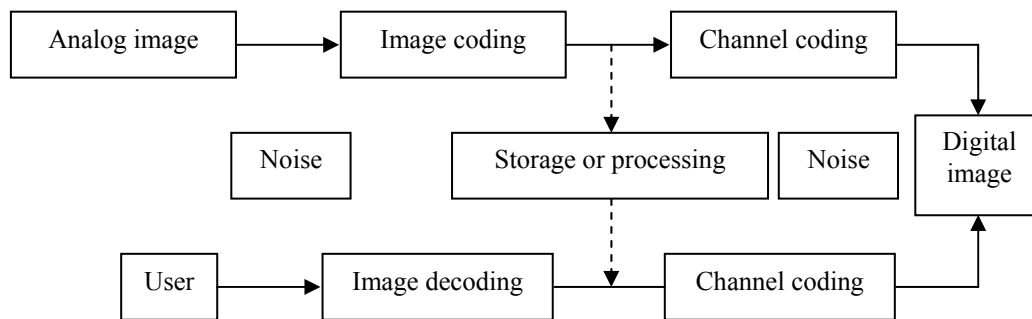

Figure 1. The digitization process of the simulation image

(1) Image information collection stage: at the image collection phase, the state of the system sensor is affected by factors such as the quality of components, the working environment, etc. The noise at this stage is mainly salt&pepper noise and bipolar noise.

(2) Image information transmission stage: at the image transmission stage, the transmission channel is interfered and noise is thus generated. Bipolar noise and Gaussian noise are common at this stage.

(3) Image information management stage: image management includes storage, delete, copy, etc. Due to aging of components, self-excitation of circuit and poor filtering in the management system, the resistance thermal noise, thunder-ball noise, etc. are generated [6].

3. Artificial Bee Colony

Artificial bee colony simulates the actual bee honey gathering mechanism and processes the function optimization problem, and divides the artificial bee colony into three categories: leading bee, followed bee and scouter. The basic idea of such algorithm starts from one randomly generated initial population, searches around one half individual with best fitness value and preferentially retains the individual by adopting the one-to-one competition survival strategy, and such operation is named the leading bee search. Then select optimal individual by using the roulette selection method and conduct the greedy search around to generate the other half individual, and such a process is named the followed bee search. New population is formed by individuals generated by leading bee and followed bee to avoid the loss of population diversity. Conduct the analogous scouter variation search to form the iterative population. By constant iterative calculation, such algorithm retains excellent individuals, eliminates inferior individuals, and gets closer to the global optimal solution.

Then, take the solution of the nonlinear function minimization for example to elaborate the concrete operation process of ABC algorithm in details.

The nonlinear function minimum value is expressed as $\min f(X)$, $X^L \le X \le X^U$, X^L and X^U are respectively the upper and lower bounds of the variable $X = (x_1, x_2, \cdots x_n)$ value, and n is the variable dimension. During the solution of the nonlinear function minimum value with ABC algorithm, first of all, the initial population including NP individuals is generated within the value scope, and each individual has a corresponding candidate solution in the feasible solution space and the dimension of the individual variable D equals the dimension n of the objective function decision variable X. Suppose the algorithm's maximum iterative number is G, the number i individual in number t population can be expressed as $x_i^t = (x_i^t(1), \cdots, x_i^t(n))$, $i = 1, 2 \cdots NP/2$. The following is the description of key steps of ABC algorithm:

(i) Population initialization

Set the initial evolution algebra $t = 0$, the initial population consisting of NP X individuals that are randomly generated by formula (1) and satisfies constraint conditions within the feasible optimization solution space is formed.

$$X_i^0 = X_i^L + rand() \times (X_i^U - X_i^L), i = 1, 2 \cdots NP \tag{1}$$

In which, X_i^0 shows the number i individual in number 0 population and $rand()$

(ii) Leading bee search

The half individual with small fitness forms the leading bee population, and the other half forms the followed bee population.

For one goal individual X_i^t of the current number t leading bee population, randomly select the individual $r_1 \in [1, 2 \cdots, NP/2]$ to conduct the crossover search by dimension to generate the new individual V. See the formula (2) for details.

$$V(j) = x_i^t(j) + (-1 + 2rand) \times (x_i^t(j) - x_{i1}^t(j)) \tag{2}$$

Figure 2 is the crossover search diagram when the objective function is two.

From the Figure 2 crossover search diagram, we can see that the generated difference vector is uncertain in the direction and size. The goal leading individual bee and randomly selected leading individual bee add such difference vector to the base vector, which equals that the random disturbance within one regulated scope is added to the base vector, thus the population diversity is enlarged [7].
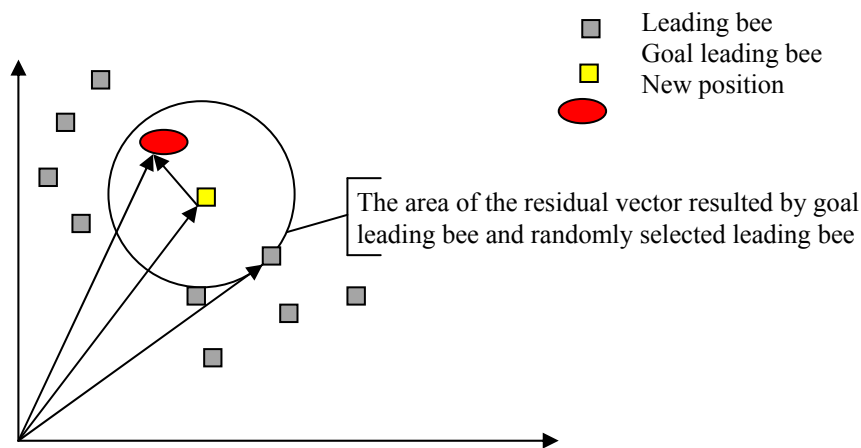

Figure 2. Illustration of the crossover process with D=2

Like other evolutionary algorithms, ABC algorithm adopts the "survival of the fittest" idea of Darwin's evolution theory retain the individual with the purpose to ensure that the algorithm

evolves constantly to the global optimal solution. Evaluate the fitness of newly generated individual V and goal individual X_i^t, and then compare their fitness values, and select the individuals with better fitness value by formula (3) to form the leading bee population.

$$X_i^{t+1} = \begin{cases} V, f(V) < f(X_i^t) \\ X_i^t, f(X_i^t) \geq f(V) \end{cases} \tag{3}$$

```
                        ┌──────────────┐
                        │    Start     │
                        └──────┬───────┘
                               ▼
              ┌────────────────────────────────┐
              │  Initialize related parameter,  │
              │       NP, G, limit, t=0         │
              └────────────────┬───────────────┘
                               ▼
              ┌────────────────────────────────┐
              │   Produce initial population    │
              └────────────────┬───────────────┘
                               ▼
```

Calculate the individual fitness value and take half numbers of the individual with best fitness value as leading bee population and the other half as followed bee population

Leading bee searches to produce new leading bee population

Followed bee searches to produce new followed bee population

Combine the leading bee population and followed bee population to form the iterative population

Whether the scouter is produced — Yes

The scouter searches and updates the iterative population
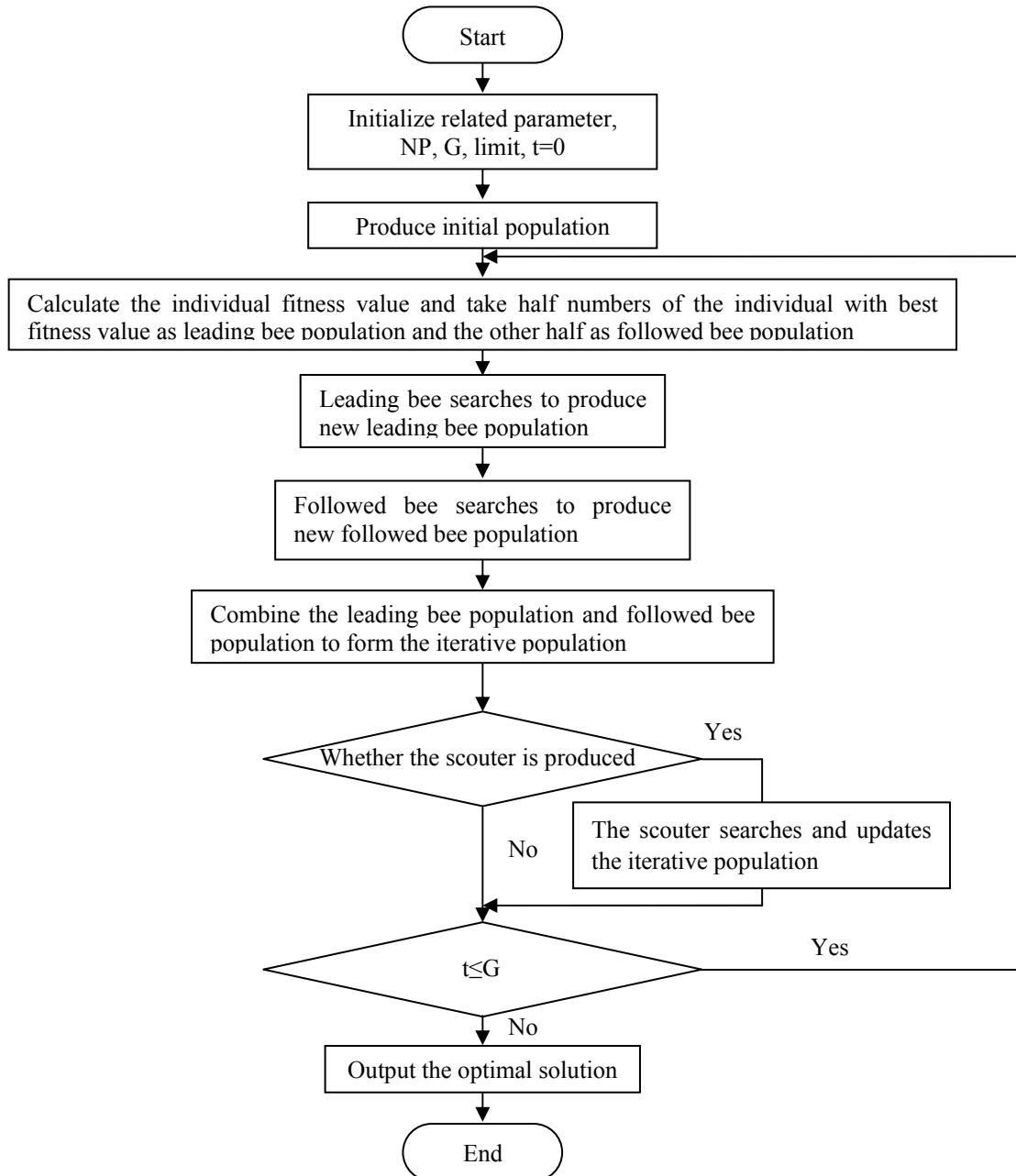
No

t≤G — Yes

No

Output the optimal solution

End

Figure 3. Flow chart of the standard ABC algorithm

(iii) Followed bee search

Followed bee selects the optimal goal individual $X_k^{t+1}, k \in [1, \cdots, NP/2]$ in the new leading bee population according to the probability formula (4) and in the roulette selection mode and

conducts the search together with randomly selected individual based on the formula (2) to generate new individual $X_k^{t+1}, k \in [NP/2+1,\cdots,NP]$ and form the followed bee population.

$$P_i = \frac{fit_i}{\sum_{i=1}^{NP/2} fit_i} \qquad (4)$$

The search way of followed bee population in the artificial bee colony is the key why it is different from other evolutionary algorithms. Its essence is to select preferentially individually to conduct the greedy search, which is the key factor of the algorithm's fast convergence. But, its search way itself introduces some random information, which thus does not reduce the population diversity to a large extent [8].

(iv) Scouter search

After the combined search by leading bee population and followed bee population, new population in the same size with the initial population is formed. In order to avoid the excessive loss of the population diversity as the population evolves, the artificial bee colony simulates the biological behavior of the scouter searching for potential honey sources to puts forward specific scouter search way. Assume a certain individual does not change continuously for "limit" generations, the corresponding individual transfers into the scouter, generates new individual according to formula (1) search and makes one-to-one comparison with original individual according to formula (3) to preferentially retain individuals with better fitness.

Through above searches of leading bee population, followed bee population and scouters, make the population evolve to the next generation and recycle until the algorithm iteration number t reaches the preset maximum iteration number G or the population optimal solution reaches the preset error accuracy.

In order to further understand the principle of the ABC algorithm, Figure 3 shows the operation flow chart.

4. Back-Propagation Network (BP Network)

Back-Propagation Network (BP Network) is the multilayer network generalizing W-H learning rule and conducting the weight training towards the nonlinear differential function. The adjustment of weights adopts the back propagation learning algorithm which is a kind of multilayer forward feedback neural network, and its neurons transformation function is S function, its output quantity is the continuous amount from 0 to 1, and it can achieve the any nonlinear mapping from input to output.

4.1. BP Network Features

(i) The input and output are parallel analog quantity.

(ii) The input and output relationship of the network is determined by weight factors connected by each layer, and there is no fixed algorithm.

(iii) The weight factor is adjusted by studying signal, and the more you learn, the smarter the network is.

(iv) The more hidden layer is, the higher the network output precision is, and the damage of some individual weight factor will not exert large impact on the network output.

Only when you want to limit the output of the network, for example between 0 and 1, there should exist S type activation function in the output layer. In general, S type activation function is usually adopted in hidden layer, while, the output layer adopts the linear activation function [9],[10].

4.2. Introduction to Multilayer BP Network

Multilayer BP network is the multilayer neural network with three layers or over three layers, and each layer is composed of a number of neurons as shown in Figure 4, and each neuron between its left and right layers achieves the full connection, namely, each neuron in the

left layer and the right layer is connected, but there is no connection between up and down neurons [11].
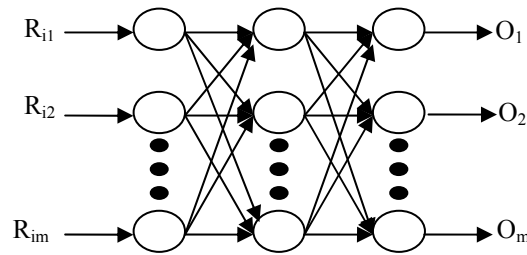
Figure 4. Multilayer BP network

BP network conducts the training according to the learning way with teachers. When a pair of learning model is offered to the network, the activation value of its neuron will spread from the input layer to the output layer through the middle layer, and each neuron in the output layer corresponds with the network response of the input mode, then, passes through the middle layer from the output layer according to the principle of reducing expected output and actual output errors, and finally returns to the input layer to correct the individual connection weight one layer by one layer. Due to this correction process is conducted from the output to input level, so it is called "error back propagation algorithm". With this kind of error back propagation training is conducted, the response accuracy rate of the network to the input mode will also be unceasingly enhanced [12].

Because BP network has the hidden layer which is in the middle position, and there are corresponding learning rules to follow to train such network to make it able to identify the nonlinear model.

4.3. Three-layer BP Network

In order to endow the BP network with a certain function to complete a task, the interlayer connection weights and node threshold value must be adjusted to ensure that errors of all samples' actual and expected output stabilize within a smaller value. In the process of training BP network algorithm, the error back propagation algorithm is one most effective and commonly used method [13]. Figure 5 shows three-layer BP network structure chart.

The learning process of BP network mainly has following four parts.

Input mode direct propagation (the input mode calculation is spread to the output layer through the middle layer from the input layer).

Output error back propagation (the output error is spread to the input layer to the output layer through the middle layer).

Cycle memory training (the calculation process of the mode direct propagation and error back propagation is repeated alternately).

Study result discretion (determine whether the global error is closer to the minimum).
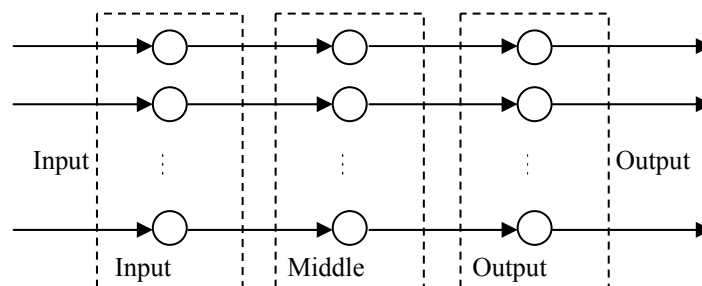
Figure 5. Three-layer BP network structure chart

5. Artificial Bee Colony Optimized BP Neural Network and Training
Artificial bee colony algorithm doesn't need the derivative but the value of the objective function and it has excellent optimization effects and low requirement for software and hardware; therefore, it has obvious advantages to use it to train the BP neural network. Integrate the artificial bee colony algorithm and the BP neural network and form ABC-BPNN.

5.1. Pixel Pollution Judgment Based on ABC-BPNN
Firstly, judge whether the image pixels have been polluted by impulse noise and divide the pixels into two types: polluted and un-polluted according to the judgment result. Preprocess the data before it enters into ABC-BPNN, including the extraction and normalization of median filter and eigenvalue. After the normalized data enters into ABC-BPNN, train it with the training method indicated in Figure 6.

Since ABC-BPNN is used to classify the polluted and unpolluted pixels, its expected output is defined by the following formula:

$$S(i,j) = \begin{cases} 1, |g(i,j) - f(i,j)| > 0 \\ 0, |g(i,j) - f(i,j)| = 0 \end{cases} \tag{5}$$

In this formula, g stands for the original and unpolluted image while f represents the image polluted by the impulse noise. If pixel (i,j) has been polluted, namely $|g(i,j) - f(i,j)| > 0$, then the expected output of ABC-BPNN is 1, otherwise, it is 0. Accordingly, if the output of the trained ABC-BPNN is close to 1, then the corresponding pixel may have been polluted, if close to 0, they may not be polluted.

5.2. ABC-BPNN Training Process
The number of neurons in the BP neural network has a relatively significant impact on the approximation accuracy. The more neurons there are, the approximation accuracy is higher. In practical application, more neurons are not better since the computation complexity is needed to be taken into consideration. Therefore, for the BP neural network, the dynamic adjustments of the number of neurons will greatly increase its application scope and get the optimal cost performance in a certain condition.

In order to automatically adjustment the network structure in the training, the training of ABC-BPNN has added the updated module of neural network structure and constituted the "dual cycle" training process of ABC-BPNN with the updated module of neural network coefficient. The neural network structure update mainly refers to the increase of neurons while the neural network coefficient is mainly updated by the bee colony algorithm. The following is the further description of ABC-BPNN training:

(i) Abstract and define the individual attribute of the bee colony.
It can be known every neuron in the neural network has 4 variables and there are 4K coefficients to be determined in the BP neural network with K neurons and every coefficient corresponds to one dimension of the search space. In other words, the search scope of the bee colony is 4K-dimensional.

(ii) Initialize the bee colony.
Initialize the bee colony before the training. Additionally, re-initialize the bee colony when there are new neurons since change has occurred in the dimensions of the search space. For the latter, if the number of the neurons increases from K to K+1, the individual dimensions of the bee colony also increases from 4K to (4K+4). Adopt a random assignment in the initialization.

(iii) Calculate the neural network output.
Calculate the output value according to Formula (6).

$$o_k = \psi(net_k) = \psi(\sum_{i=1}^{q} w_{ki} y_i + a_k) = \psi\left[\sum_{i=1}^{q} w_{ki}\phi(\sum_{j=1}^{M} w_{ij}x_j + \theta_i) + a_k\right] \tag{6}$$

(iv) Training stop conditions

Two conditions: a) When it reaches the maximum iterations. b) When it reaches the expected output error. Iteration will stop if one of the two conditions is satisfied.

(v) Adjust the network structure conditions.

Judge whether adjustments need to be made in the network structure. To increase the neurons need to investigate the preset neural network adjustment rules in advance. Increase the number of the neurons if satisfying the rules and the adjustment rules are as follows:

a) The number of the initial neurons is 1.

b) Set the error threshold Γ to accept the existing structure and the maximum tolerance algebra G_p to achieve this threshold.

Structure adjustment rule:

Accept the setting of the current neurons if the average error in the approximate generation of G_p is smaller than Γ after G_p generation evolution since the initialization, otherwise, increase the number of the neurons.

Actually, what the error Γ gives is the expected convergence accuracy of ABC-BPNN in the current structure while G_p gives the expected convergence speed. If the current structure fails to converge at the set accuracy and speed, it can only adjust ABC-BPNN structure and increase the neurons.

(vi) Implement the update of the bee colony individuals according to the established artificial bee colony optimization.
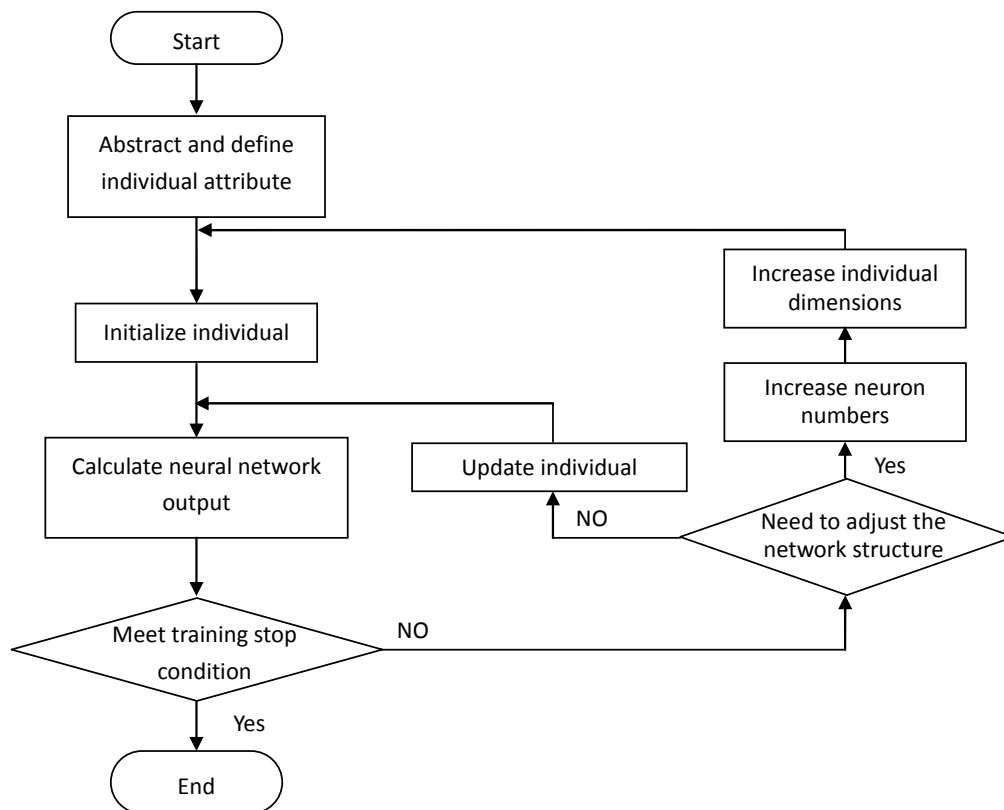
ABC-BPNN training process is shown in Figure 6.

Figure 6. ABC-BPNN training process

6. Simulation and Experimental Analysis

In order to test the performance of the algorithm proposed in this paper, the following takes the common 256*256 original grayscale image Cameraman as an example, in Matlab test

environment, add the Gaussian noise $\mu = 0$, $\sigma^2 = 0.02$, and make a comparison with the traditional denoising methods such as mean filtering, global threshold filtering and median filtering, the experimental result is shown in Figure 7.

It can be seen from Figure 7, after adding intense Gaussian noise to the original image Lena, the visual effect of the image after being filtered in this paper is significantly improved, the edge and image details are more clear, the image quality is improved, and it has shown prominent advantages compared with the traditional Gaussian noise filtering method. Thus, the algorithm adopted in this paper has great advantages compared with the traditional algorithm and the global threshold algorithm, the SNR of the image is increased to a large extent and the quality of the image is greatly improved.
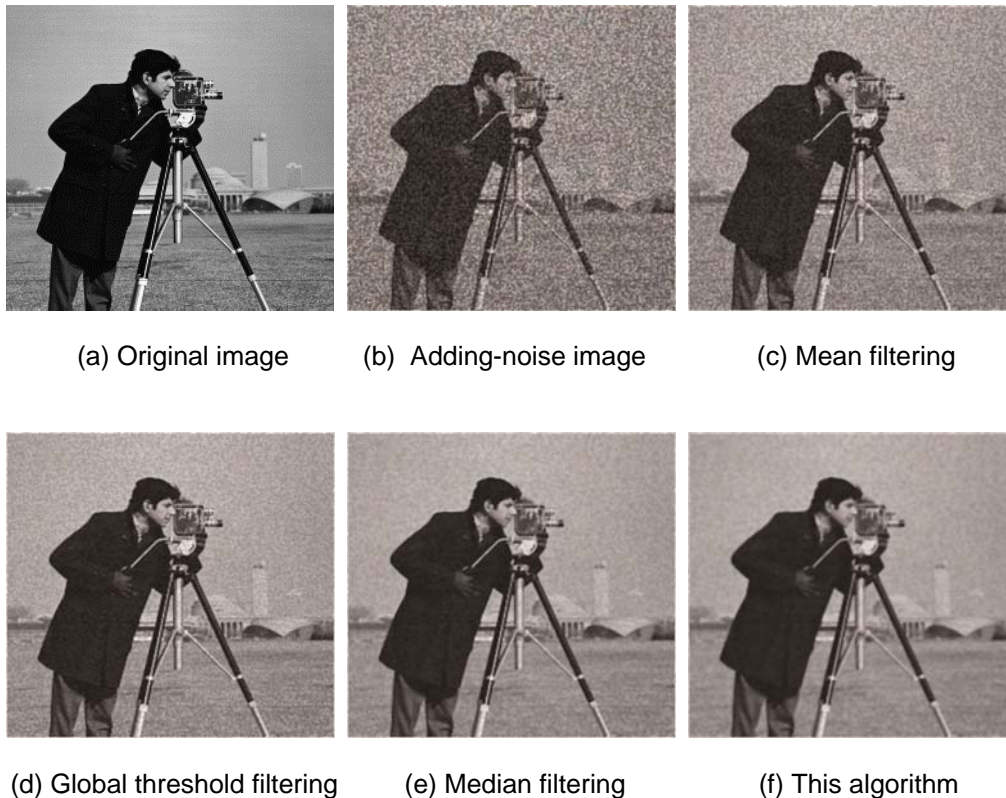
(a) Original image	(b) Adding-noise image	(c) Mean filtering
(d) Global threshold filtering	(e) Median filtering	(f) This algorithm

Figure 7. The denoising effect comparison of Cameraman image

7. Conclusion

In order to improve the image quality and meet the requirement of the subsequent higher-level processing, image denoising has become important work in the pre-processing process of image. The conventional image denoising methods will blur the image edge, while maintaining and enhancing the image edge, they will affect the denoising effect of the image. This paper, with its focus on image denoising, has realized the image denoising method based on artificial bee colony and BP neural network (ABC-BPNN), by comparing with other traditional denoising methods through simulation experiment, the better image quality is restored, which has verified the performance of the algorithm proposed in this paper.

References

[1] Hamid AJ, Rabha WI. Fractional Alexander Polynomials for Image Denoising. *Signal Processing*. 2015; 107(2): 340-354.

[2]　PR Hill, AM Achim, DR Bull, ME Al-Mualla. Dual-tree Complex Wavelet Coefficient Magnitude Modelling using The Bivariate Cauchy–Rayleigh Distribution for Image Denoising. *Signal Processing*. 2014; 105(12): 464-472.

[3]　Fatma L. A Novel Approach to Speckle Noise Filtering Based on Artificial Bee Colony Algorithm: An Ultrasound Image Application. *Computer Methods and Programs in Biomedicine*. 2013; 111(3): 561-569.

[4]　Hanshan L, Zhiyong L. Research on Infrared Special Facula View Measurement Method Based on Image Processing Technology. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(6): 1422-1429.

[5]　Hamid AJ, Rabha WI. Fractional Alexander Polynomials for Image Denoising. *Signal Processing*. 2015; 107(2): 340-354.

[6]　PR Hill, AM Achim, DR Bull, ME Al-Mualla. Dual-tree Complex Wavelet Coefficient Magnitude Modelling using The Bivariate Cauchy–Rayleigh Distribution for Image Denoising. *Signal Processing*. 2014; 105(12): 464-472.

[7]　Fatma L. A Novel Approach to Speckle Noise Filtering Based on Artificial Bee Colony Algorithm: An Ultrasound Image Application. *Computer Methods and Programs in Biomedicine*. 2013; 111(3): 561-569.

[8]　Kazim H, M Fatih T. Segmentation of SAR Images using Improved Artificial Bee Colony Algorithm and Neutrosophic Set. *Applied Soft Computing*. 2014; 21(8): 433-443.

[9]　G Rosline N, S Maruthuperumal. Normalized Image Watermarking Scheme Using Chaotic System. *International Journal of Information and Network Security (IJINS)*. 2012; 1(4): 255-264.

[10]　Bhargav V, Biswarup D, Rudra PM. An Improved Scheme for Identifying Fault Zone in A Series Compensated Transmission Line using Undecimated Wavelet Transform and Chebyshev Neural Network. *International Journal of Electrical Power & Energy Systems*. 2014; 63(12): 760-768.

[11]　Zhen L, Bingang X, Zheru C, Dagan F. Intelligent Characterization and Evaluation of Yarn Surface Appearance using Saliency Map Analysis, Wavelet Transform and Fuzzy ARTMAP Neural Network. *Expert Systems with Applications*. 2012; 39(4): 4201-4212.

[12]　Chia-Nan K. Identification of Nonlinear Systems with Outliers using Wavelet Neural Networks Based on Annealing Dynamical Learning Algorithm. *Engineering Applications of Artificial Intelligence*. 2012; 25(3): 533-543.

[13]　Alexander AF, Dusan H, Pavel YP, Vaclav S. New BFA Method Based on Attractor Neural Network and Likelihood Maximization. *Neuro computing*. 2014; 132(20): 14-29.