Xiaoqun Qin,
Ye Chen

Hunan International Economics University, Changsha, Hunan, China, e-mail: 153569367@qq.com

OPTIMIZING FEED-FORWARD NEURAL NETWORK WEIGHT BASED ON ORTHOGONAL GENETIC ALGORITHM

Сяоцюнь Цінь,
Є Чень

Хунаньський університет міжнародної економіки, м. Чанша, провінція Хунань, КНР, e-mail: 153569367@qq.com

ОПТИМІЗАЦІЯ ВАГИ НЕЙРОННОЇ МЕРЕЖІ ПРЯМОГО РОЗПОВСЮДЖЕННЯ НА ОСНОВІ ОРТОГОНАЛЬНОГО ГЕНЕТИЧНОГО АЛГОРИТМУ

Purpose. Artificial neural network has been successfully applied in such fields as pattern recognition, intelligent control, combinatorial optimization and prediction. The integration of neural network and other traditional methods will promote the continuous development of artificial intelligence and information processing technology. The research combined Orthogonal Genetic Algorithm (OGA) and BP neural network. We have obtained an improved neural network-OGANet, which is helpful and effective for optimizing feed-forward neural network weight.

Methodology. In order to resolve the defects of the traditional BP network learning method: the tendency to be trapped in local extremum and low learning precision, we proposed a learning method to optimize BP neural network weight based on orthogonal genetic algorithm, which not only exerts the non-linear mapping capacity of BP neural network, but also strengthens the BP neural network learning ability.

Findings. The fundamental artificial neural network and orthogonal genetic algorithm were introduced. Then, we obtained the OGANet algorithm, and proved its effectiveness. The experimental result showed that the OGANet has high precision in learning, fast convergence speed and shows better performance than other neural network learning methods.

Originality. We made a study of optimizing feed-forward neural network weight based on orthogonal genetic algorithm. The analysis of the training results of the improved network has proved that the OGANet not only has fast training speed, but also, to certain extent, overcomes the tendency to be trapped in local minimum, which is a shortcoming of traditional BP neural network.

Practical value. The non-linear and self-adaptive information processing capacity of the neural network overcomes the defects of traditional artificial intelligence methods. This method effectively combines the advantages of genetic algorithm and BP neural network and it gives an excellent compromise between the global search of orthogonal genetic algorithm and the global development of BP network.

Keywords: *orthogonal genetic algorithm, BP neural network, weight optimization, global search, local minimum, learning precision*

Introduction. The structure and working mechanism of artificial neural network (ANN) is on the background of the organizational structure (the network of brain neurons) and the activity rule of the brain and it reflects certain basic characteristics of the brain [1]. By stimulating the neural network behavior feature of the animal, it is the mathematical model, which performs distributed parallel information processing. As for the system complexity, such network adjusts the interconnecting relationship between numbers of nodes and attains the goal of information processing. In recent years, neural networks have been developed in a more in-depth manner in the way of stimulating human cognition; moreover, in combinations with fuzzy system, genetic algorithm, and evolutionary mechanism, forms the computational intelligence and becomes an important trend of artificial intelligence [2].

Psychologist and mathematical logician have proposed the formalized mathematical description and network structure method of neurons and proved that every single neuron can implement the logic function, starting an era of the research of artificial neural network. J.J. Hopfield, a physicist from California Institute of Technology, has initiated new approaches of neural network in associative memory and optimization computation and greatly propelled the research on neural network [3]. The scholars such as Rumelhart have come up with back propagation algorithm of multi-layer perceptron, which has overcome the significant obstacles in the continuous development of perceptron model. A lot of in-depth pioneering work has greatly developed the model and learning algorithm of neural network. BP (Back Propagation) network was raised by the scientists' team, which was led by Rinehart and McClelland. BP network is a multi-layer feed-forward network trained by back propagation algorithm and it is one of the most extensively-used neural network models currently. BP network is the core of feedforward network and it has been very widely applied. At present, in the practical application of artificial neural network, 80–90% of artificial neural network models use BP network and its variations [4]. Intelligent search algorithms include genetic algorithm, tabu algorithm and simulated annealing algorithm. The genetic algorithm is inspired by the biological evolution of genetic inheritance. In fact, it is an evolutionary computation, which, in essence, is a self-adaptive machine learning method. According to the changes of genetic inheritance, the genetic algorithm includes three operations: selection, crossover and mutation [5]. It is more highly-efficient than the blind search and more universal than the algorithms for specific problems, in other words, it is a solution pattern irrelevant to problems. The traditional neural network training algorithms are slow in convergence and easy to be trapped in local optimum. The genetic algorithm has better global convergence performance and it can be used to train the neural network parameters and structure. Many researchers have brought form some improved genetic algorithms, but there are still some deficiencies in them.

This paper firstly introduces the fundamental principle of artificial neural network and orthogonal genetic algorithm. Then, based on the above research, it proposes the application step of OGANet. Finally, the comparison of the network output curve chart and the sample curve chart of two functions trained by OGANet and proves the effectiveness of the algorithm suggested by the authors.

Principle of artificial neural network. ***Artificial neuron model.*** Make $x_i(t)$ represent the input information received by neuron j from neuron i at the time of t and $o_j(t)$ the output information of neuron j at the time of t, then the state of neuron j can be expressed as

$$o_j(t) = f\{[\sum_{i=1}^{n} w_{ij} x_i (t - \tau_{ij})] - T_j\}.$$

In the above formula, τ_{ij} is the synapse delay between the input and output, T_j is the threshold of neuron j, w_{ij} is the synapse link coefficient from neuroni to neuron j and $f(\bullet)$ is the neuron transfer function [6].

BP neural network. Multi-layer BP network is a kind of multi-layer neural network with three or more layers. Every layer is comprised of certain neurons. The neurons in the left and right layers are fully connected, meaning that every neuron in the left layer is connected to every neuron in the right layer while the upper and lower neurons are not connected.

The excitation mode (input vector) units in each source node in the input layer of the multi-layer feed-forward neural network of BP neural network model constitute the input signals used in the neurons (computation nodes) in the 2nd layer. The output signal in the 2nd layer is the input of the 3rd layer and the rest are similar to the above. The topology of BP neu-ral network includes the input layer, the hidden layer, and the output layer (Fig. 1).

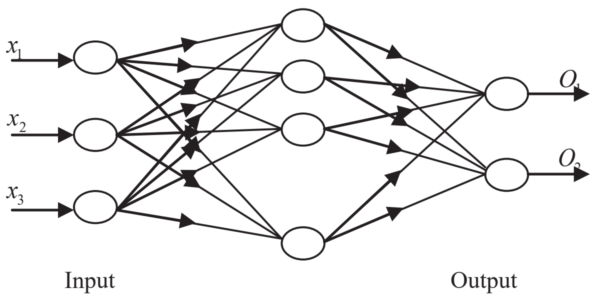

Fig .1. BP neural network model

The steps of BP learning algorithm are classified as follows:

1. Initialize the input according to the network requirements and set the input variables and parameters. The input and output are parallel analog signals.

2. Perform learning on the input sample with BP network and conduct feed-forward computation on the input and output signals of the neurons in every layer of BP network. The input-output relationship of the network is determined by the weight factor connecting various layers.

3. Calculate the error from the expected output and the obtained actual output and judge whether it meets the requirements.

4. Perform pattern classification on the samples to be recognized with well-trained BP network. The more hidden layers, the higher the network output precision.

Generally, BP network consists of more than one neuron and is trained in the form of supervised learning. Its common behavior not only depends on the features of every single neuron, but it may also be determined by the interaction and interconnection between the neurons. Without fixed algorithm, the weight factor adjusts through the learning signal; the damage of any weight factor will not have significant influence on the network output. The enlargement of the layers can further decrease the error and improve the precision, however, it will make the network more complicated and subsequently increase the training time of the network weight [7].

The basic idea of the orthogonal genetic algorithm. The basic idea of the genetic algorithm starts from the initial population and searches from the intermediate set of the solutions instead of the single solution. It selects the individuals with the natural rule of survival of the fittest. Searching string set, it has a large covering range. It is good for global optimization and requires only such universal information as fitness value and string encoding. In fact, the initial string set carries lots of information, which is greatly irrelevant to the optimal solution, and new generation of the population is produced through the crossover and mutation. The optimal solution search is performed with random methods. The selection shows the approximation towards the optimal solution, crossover demonstrates the production of the optimal solution while mutation indicates the coverage of the global optimal solution. A generation by generation is evolved this way until the goal is achieved [8].

Orthogonal experiment design method has been introduced in the genetic algorithm with multi-point orthogonal crossover operation to handle various function optimization problems. It arranges the exchange operation of the genetic algorithm through the orthogonal table and selects the offspring individuals with big fitness into the next evolution. The crossover operator randomly produces new offspring individuals near the selected parental individuals. These operations can be deemed as a sampling experiment. It accelerates the convergence speed of the algorithm and guarantees the diversity of the population. It not only effectively overcomes the defects of the standard genetic algorithm, but also greatly improves its computation speed, precision and the algorithm stability [9].

The orthogonal design is a common method in experiment design and its basic tool is the orthogonal table, which is a kind of mathematical table constructed by combinational mathematical theory according to the idea of equilibrium distribution. The equilibrium distribution is the core of the orthogonal table.

For a group of given training sample sets $U=[u_1,u_2,\ldots,u_n]\in R^{m\times n}$, u_i is the ith sample. Assume that $S_i\in R^{n\times 1}$ is the orthogonal reconstruction coefficient of u_i and the ith element is 0. Calculate every orthogonal reconstruction coefficient through orthogonal projection.

$$\min_{a_i} |s_i|;$$
$$s.t.\|u_i-Us_i\|_2<\varepsilon, 1=e^T s_i .$$

Here, ε is the threshold value of the orthogonal reconstruction error. In order to maintain the orthogonal reconstruction relationship among the training samples in the exchange space, the orthogonal reconstruction projection defines the minimum error objective function in the projection space, namely

$$\min_v \sum_{i=1}^{n}\left\|v^T u_i - v^T U s_i\right\|^2 ,$$
$$s.t. v^T U U^T v = 1 .$$

The purpose of the iterative orthogonal reconstruction projection is to find a group of orthogonal projection matrixes, which can maintain the orthogonal reconstruction relationship of the original samples. The matrix vectors are in a pair wise orthogonal relationship and they do not change the constraint conditions of the original orthogonal reconstruction. Because the constraint conditions are quite complicated, the iteration is needed to obtain every projection vector [10]. The procedures of the orthogonal genetic algorithm are presented in Fig. 2.

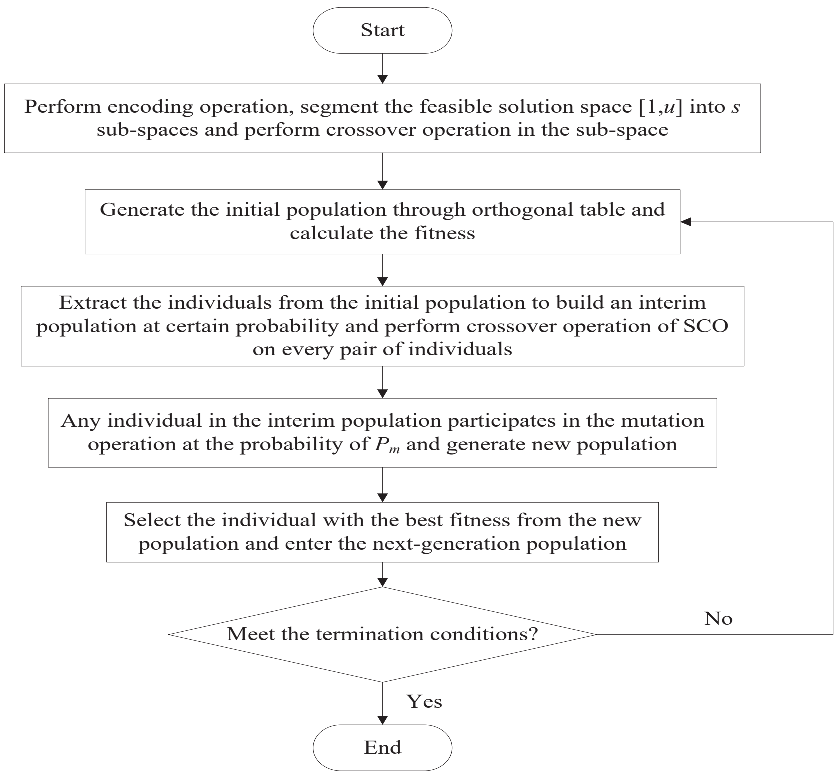

Fig. 2. The procedures of the orthogonal genetic algorithm

Neural network weight training based on the orthogonal genetic algorithm (OGA). The OGANet can converge to a global optimal solution and it has strong robustness. The integration of the OGANet and a feed-forward neural network can not only exert the non-linear mapping capacity of BP neural network but also make the BP neural network have fast convergence speed and strong learning ability.

The entire process of OGA is the same as that of the traditional genetic algorithm and the only difference is that this algorithm uses orthogonal design to produce the initial population and the offspring individuals produced by the crossover operators.

Step 1: In the computational experiment, the parameter selection of the mixed algorithm is as follows: population size is $P=60$, the number of individual to have the operation is $O=0.1P$, the selection probability of optimal individual is $q=0.2$, the crossover probability is $p_c=0.80$ and the mutation probability is $p_m=0.03$. Since the optimization complexity of every test function is different, the optimization termination condition for mixed algorithms of different functions is also different. The algorithm terminates when the current number of iterations l reaches the evolutional generation L.

Step 2: As for the input position of every weight in the offspring chromosome, produce NP mutation individuals $v_1^{l+1}, v_2^{l+1}, \ldots, v_{NP}^{l+1}$ through the mutation operators of GA weight. The mutation operator randomly selects a value at the probability of P_m in the initial probability distribution and then adds it with the weight of that input position.

Step 3: As for the input position of every weight in the offspring chromosome, the mutation operator randomly selects several crossover positions from the two parental chromosomes and produces NP crossover individuals $u_1^{l+1}, u_2^{l+1}, \ldots, u_{NP}^{l+1}$ through mixed crossover methods. Calculate the corresponding fitness value of every crossover individual (weight vector). Perform crossover operation on the generation of the chromosome in the crossover position, in this way; the offspring chromosome will have the gene of two parental generations.

Step 4: After calculating the fitness of each individual, select the individual with bigger fitness and pass genetically to the next generation to make the solution closer and closer to the optimal solution space. Produce NP individuals of next-generation weight vector, namely $w_1^{l+1}, w_2^{l+1}, \ldots, w_{NP}^{l+1}$. The selection probability of each individual is in proportion to its fit-ness. The selection probability P_s to the individual with the fitness F_t of is

$$P_s = F_t \Big/ \sum_{l=1}^{N} F_l .$$

Here, N is the population size. In the practical learning, the individual with the maximum fitness is passed genetically to the next generation.

Step 5: Make $l = l + 1$, turn to Step 2 and keep training the weight, otherwise, the result is the individual of the maximum fitness and determine the initial connection weight and threshold of BP neural network.

Step 6: Train BP neural network, output the global error of the result calculated through feed-forward propagation, judge whether it reaches the established precision. If it does, terminate the network training.

Step 7: If it fails to reach the established precision, turn to Step 6 until BP neural network is converged or until the algorithm has performed the given maximum number of iterations L, end the operation and output the calculation result.

Training result and analysis. In order to evaluate the performance of the hybrid training algorithm OGANet in optimizing the weight and threshold of feed-forward neural network, we tested the function curve simulation recognition problems with the data from the neural network learning database.

The test functions are as follows

$$y = f_1(x) = \frac{\sin(5x)}{5x}, x \in (0,16] ;$$

$$y = f_2(x) = [1 - 7\cos(3x^2)]\exp\left(-\frac{x^2}{2}\right), x \in [-5,5] .$$

The parameters are set as follows: the population size is $NP = 60$, the maximum number of iterations is L. Both the scale factor F and the crossover probabilit CR use random parameters, namely $F = 0.6 + rand(0,1) \times 0.15$; $CR = 0.5 + rand(0,1) \times 0.2$. Here, $rand(0,1)$ is the random number within (0, 1).

In the training, all the test problems use 3-level feed-forward neural network. The neurons in the hidden layer use Sigmoid excitation function (logsig) and the neurons in the output layer adopt linear excitation function (purelin). The test operation results of the above 2 functions are indicated in Fig. 3, 4 where x axis represents the input vector and the y axis is the output vector.

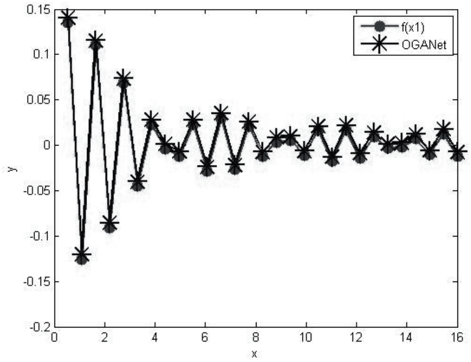

Fig. 3. The comparison between the sample curve of function $f_1(x)$ and the training curve of OGANet

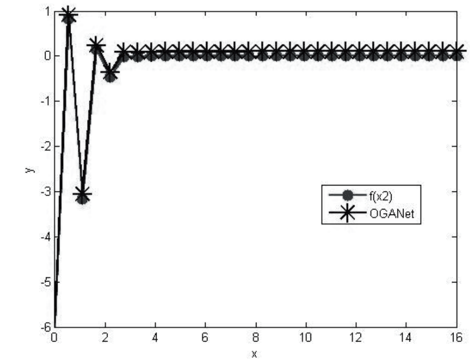

Fig. 4. The comparison between the sample curve of functio $f_2(x)$ and the training curve of OGANet

It can be seen that in the 30 independent trainings, the optimization performance of OGANet is excellent. As for the 2 function test problems, the mean square error of the training results of OGANet is smaller. The training result of OGANet algorithm is very close to the actual curve of this function.

Conclusion. In order to resolve the defects of the traditional BP network learning method: the tendency to be trapped in local extremum and the low learning precision, we proposed a learning method to optimize BP neural network weight based on the orthogonal genetic algorithm. This method effecttively combines the advantages of the genetic algorithm and BP network, and it gives an excellent compromise between the global search of the orthogonal genetic algorithm and the global development of BP network. The experimental result has shown that the method has high learning precision, fast convergence speed and shows better performance than other neural network learning methods.

References / Список літератури

1. Pławiak, P., 2014. An estimation of the state of consumption of a positive displacement pump based on dynamic pressure or vibrations using neural networks. *Neurocomputing*, vol.144, no.20, pp. 471–483.

2. Renbo Luo, Wenzhi Liao and Youguo Pi., 2014. Discriminative supervised neighborhood preserving embedding feature extraction for hyperspectral image classification. *TELKOMNIK Indonesian Journal of Electrical Engineering*, vol. 12, no.6, pp. 4200–4205.

3. Buse Melis Ozyildirim and Mutlu Avci, 2014. Logarithmic learning for generalized classifier neural network. *Neural Networks*, vol.60, no.12, pp. 133–140.

4. Hrvoje Krstić, Željko Koški, Irena Ištoka Otković, Martina Španić, 2014. Application of Neural Networks in Predicting Airtightness of Residential Units. *Energy and Buildings*, vol. 84, no.12, pp. 160–168.

5. R. Krishnamoorthi and S. Sathiya Devi, 2013. A simple computational model for image retrieval with weighted multi-features based on orthogonal polynomials and genetic algorithm. *Neurocompu-ting*, vol.116, no.20, pp. 165–181.

6. Susmita Mall, S. Chakraverty, 2014. Chebyshev neural network based model for solving Lane–Emden type equations. *Applied Mathematics and Computation*, vol.247, no.15, pp. 100–114.

7. Tomás Rodríguez García, Nicoletta González Cancelas, Francisco Soler-Flores, 2014. The artificial neural networks to obtain port planning parameters. *Procedia-Social and Behavioral Sciences*, vol.162, no.19, pp. 168–177.

8. Donghong Zhao, 2014. Total variation differential equation with wavelet transform for image restoration. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol.12, no.6, pp. 4747–4755.

9. Mohammad Khanzadeh and Moslem Malekshahi, Ali Rahmati, 2013. Optimization of loss in orthogonal bend waveguide: Genetic algorithm simulation. *Alexandria Engineering Journal*, vol.52, no.3, pp. 525–530.

10. Bernadete M.M. Neta, Gustavo H.D. Araújo, Frederico G. Guimarães, Renato C. Mesquita and Petr Ya. Ekel, 2012. A fuzzy genetic algorithm for automatic orthogonal graph drawing. *Applied Soft Computing*, vol.12, no.4, pp. 1379–1389.

Мета. Штучна нейронна мережа успішно застосовується в таких областях як розпізнавання образів, інтелектуальне управління, комбінаторна оптимізація та прогнозування. Спільне використання нейронних мереж із традиційними методами сприятиме подальшому розвитку технологій штучного інтелекту та обробки інформації. У дослідженні були поєднані ортогональний генетичний алгоритм (OGA) та навчання нейронних мереж методом зворотного поширення помилки. Отримана модифікована нейронна мережа OGANet, що ефективно застосовується для оптимізації ваги нейронних мереж прямого поширення.

Методика. З метою усунення недоліків традиційного навчання нейронних мереж методом зворотного поширення помилки, а саме схильності до попадання в “пастку” локального оптимуму та низької точності навчання, запропоновано метод навчання на основі OGA-алгоритма, що оптимізує вагу нейронної мережі, навченої методом зворотного поширення помилки. Він позитивно впливає на здатності нелінійного картування й покращує навченість розглянутих нейронних мереж.

Результати. Були поєднані базова штучна нейронна мережа та ортогональний генетичний алгоритм. Отримано OGANet-алгоритм і підтверджена його ефективність. Результат експерименту показав, що OGANet має високу точність навчання, швидку збіжність і перевершує інші методи навчання нейронних мереж за продуктивністю.

Наукова новизна. Обґрунтована оптимізація ваги нейронних мереж прямого поширення, що ґрунтується на OGA-алгоритмі. Аналіз результатів навчання покращеної нейронної мережі підтвердив, що OGANet не лише має велику швидкість навчання, але й долає, у деякій мірі, такий недолік традиційного навчання нейронних мереж методом зворотного поширення помилки, як схильність до попадання в “пастку” локального оптимуму.

Практична значимість. Здібність розглянутої нейронної мережі до нелінійної саморегульованої обробки інформації дозволяють здолати недоліки традиційних методів штучного інтелекту. Даний метод ефективно поєднує переваги генетичного алгоритму та навчання нейронних мереж методом зворотного поширення помилки, пропонує гарний компроміс між глобальним пошуком ортогонального генетичного алгоритму та навчанням нейронної мережі.

Ключові слова: *ортогональний генетичний алгоритм, навчання нейронних мереж методом зворотного поширення помилки, оптимізація ваги, глобальний пошук, локальний мінімум, точність навчання*

Цель. Искусственная нейронная сеть успешно применяется в таких областях как распознавание образов, интеллектуальное управление, комбинаторная оптимизация и прогнозирование. Совместное использование нейронных сетей с традиционными методами будет способствовать дальнейшему развитию технологий искусственного интеллекта и обработки информации. В исследовании были совмещены ортогональный генетический алгоритм (OGA) и обучение нейронных сетей методом обратного распространения ошибки. Получена модифицированная нейронная сеть OGANet, которая эффективно применяется для оптимизации веса нейронных сетей прямого распространения.

Методика. С целью устранения недостатков традиционного обучения нейронных сетей методом обратного распространения ошибки, а именно склонности к попаданию в “ловушку” локального оптимума и низкой точности обучения, предложен метод обучения на основе OGA-алгоритма, оптимизирующий вес нейронной сети, обученной методом обратного распространения ошибки. Он положительно влияет на способности нелинейного картирования и улучшает обучаемость рассматриваемых нейронных сетей.

Результаты. Были совмещены базовая искусственная нейронная сеть и ортогональный генетический алгоритм. Получен OGANet-алгоритм и подтверждена его эффективность. Результат эксперимента показал, что OGANet имеет высокую точность обучения, быструю сходимость и превосходит прочие методы обучения нейронных сетей по производительности.

Научная новизна. Обоснована оптимизация веса нейронных сетей прямого распространения, основываю-

щаяся на OGA-алгоритме. Анализ результатов обучения улучшенной нейронной сети подтвердил, что OGANet не только имеет большую скорость обучения, но и преодолевает, в некоторой степени, такой недостаток традиционного обучения нейронных сетей методом обратного распространения ошибки, как склонность к попаданию в "ловушку" локального оптимума.

Практическая значимость. Способности рассматриваемой нейронной сети к нелинейной саморегулирующейся обработке информации позволяют преодолеть недостатки традиционных методов искусственного интеллекта. Данный метод эффективно совмещает преимущества генетического алгоритма и обучения нейронных сетей методом обратного распространения ошибки, предлагает хороший компромисс между глобальным поиском ортогонального генетического алгоритма и обучением нейронной сети.

Ключевые слова: *ортогональный генетический алгоритм, обучение нейронных сетей методом обратного распространения ошибки, оптимизация веса, глобальный поиск, локальный минимум, точность обучения*

Рекомендовано до публікації докт. техн. наук В.І. Корнієнком. Дата надходження рукопису 15.03.15.

Cheng Zhang,
Hao Peng

Hunan International Economics University, Changsha, Hunan, China, e-mail: 289962526@qq.com

IMAGE EDGE DETECTION BASED ON HYBRID ANT COLONY ALGORITHM

Чен Чжан,
Хао Пен

Хунаньский університет міжнародної економіки, м. Чанша, провінція Хунань, КНР, e-mail: 289962526@qq.com

ВИДІЛЕННЯ КОНТУРІВ ЗОБРАЖЕННЯ НА ОСНОВІ ГІБРИДНОГО МУРАШИНОГО АЛГОРИТМУ

Purpose. With the scientific and technological development as well as the plenty of image information and exchange, the edge detection and automatic identification of the target image have become increasingly intensive. Its practical application problems have raised higher and higher requirements on the image edge detection techniques. The research deals with the problems of detecting the ideal edges and determining every parameter of this algorithm.

Methodology. We combined Ant Colony Algorithm (ACA) and Differential Evolution Algorithm (DE) and used it in the image edge detection. By analyzing the convergence time and optimization capacity of these two algorithms, we found the best method to combine and apply them to the image edge detection.

Findings. We found a way of image edge detection based on ACA and DE. We firstly made a theoretical analysis of image edge detection and found the best combination point of ACA and DE according to their own features. Mainly, it integrated these two algorithms according to the set conditions and included the operating steps of DE in the early phase of image processing followed by the operations of ACA.

Originality. We made a study of image edge detection based on the hybrid ant colony algorithm. We discussed how to detect the ideal edges of an image based on ACA and DE. The research on this aspect has not been found at present.

Practical value. We integrate these two algorithms to extract the complete edge contour in order to make the detected edges continuous and the edge localization accurate. The experimental result shows that the hybrid algorithm not only enhances the adaptive ability and capacity of global optimization but also has excellent edge detection effect, greatly reducing the computation workload and time.

Keywords: *image edge detection, differential evolution, ant colony algorithm, convergence time, optimization capacity, best combination point*

Introduction. The edge is the most fundamental image feature and it contains most of the image information. The research of human visual system demonstrates that the image edge is of particular importance since an object can be identified by only a few rough contours, which are the image edges [1]. Edge detection techniques have been extensively applied in such image analysis and processing field as feature description and image segmentation and it has become a research hotspot in the image processing and analysis technology. Generally, the image recognition and classification techniques include the following phases: image pre-processing, image edge detection, feature extraction and automatic classifier design, among which, the contents and methods of edge detection are extremely important. In practice, the background, the views, the lighting and the size may interrupt edge detection; therefore, it is a significant problem in application research how to extract the feature parameters with excellent edge detection performance and few noise interruptions [2].

As the preliminary phase of vision, image edge detection has a long research history where new theories and methods keep emerging, stressing the importance of the image edge detection research. Together with the development of computer vision and image processing techniques, it is badly in need to make some breakthroughs in the early vision stage