Improved Differential Evolution Algorithm Based On Elite Group

XiaoBo GAO [1, a], YouCai WANG [2, b *], GuangZhao YANG [3, c]

[123]High-tech Institute, Fan Gong-ting South Street on the 12th, Qing Zhou, Shan Dong, China

[b]email:wangyoycai@126.com, [*]corresponding author

Keywords: DE, information entropy, average-distance-amongst-points, elite group.

Abstract. By introduce the information entropy and the average-distance-amongst-points to analysis the population distribution in the process of evolution, and figured out the cause of the DE/best/* premature convergence is the control function of the current optimal individual to decrease the population diversity of the algorithm. Based on the number of base vectors, improved the DE algorithm by setting up the elite group, the elite differential evolution algorithm is proposed. Finally, several typical test functions are used to test the performance. The results show that the elite differential evolution algorithm has a good performance in the search success rate and the global search capability.

Introduction

As a new evolutionary computation technique [1], the Differential Evolution algorithm(DE) is widely used in Intelligent computing area by virtue of its good performance, such as neural network, mechanical design, Power system, robot, signal and image processing, system identification and fault diagnosis. But the DE algorithm has slow convergence rate, low efficiency, and sometimes cannot find the optimal solution during solve the problem of noise optimization.

Many scholars have carried out a lot of research to overcome shortcomings and deficiencies of the DE, Mainly focused on how to carry out Parameter setting [2], improve the convergence rate of the algorithm [3], overcome premature convergence [4], and solve the constrained optimization problems [5]. In this paper, aim at solving the problem of search success rate low and the shortcoming of easy to fall into the local optimum, improve the DE from the number of base vector selection, take the current optimal solution and partial sub optimal solution to form Elite group. Make full use of sub optimal individual information, improve the utilization of information and select Mutation basis vectors from the elite group to improve the DE algorithm.

DE Algorithm Analysis

The DE algorithm is a random search algorithm from the point of view of Mathematics. From an engineering point of view, it is a kind of adaptive iterative optimization process. The basic idea of the DE algorithm is to start from a randomly generated initial population. It generates new individuals by weighting the vector difference between any two individuals in a population and Sum a Third individuals in certain rule, then the new individual and the target individual are operated Crossover to get the test individual. If the fitness value of the test individual is better than target individual, the test individual is used to replace the target individual in the next generation. Otherwise the target individual is still preserved. By constantly iterative computing, retain excellent individuals and eliminate the poor quality individuals, It guides the search process to the optimal solution. The DE algorithmic process [6] specific includes population initialization, mutation, crossover and selection operation, and continuous evolutionary update to determine whether the termination conditions are satisfied.

The Mutation operation is the basis of DE algorithm, the selection of suitable mutation strategy is a problem that must be faced in the process of solving practical problems by using DE algorithm. There are ten different types of mutation strategies that can be selected for the mutation operation in DE. According to the choice of base vector, there are two basic solutions: (1) the Uniform random selection scheme that is DE/rand/*/* algorithm; (2) the Current optimal selection scheme that is DE/best/*/* algorithm.

The DE/rand/*/* algorithm does not need any fitness information, it randomly selects the individual as the basis vector in the whole population range, and has strong reliability and high Search success rate, but it need more objective function evaluation times and lower convergence rate; the DE/best/*/* algorithm selects current optimal individual as the base vector according to the individual fitness information, it has a fast convergence speed. But sometimes premature convergence occurs that lead to lower search success rate. If the search success rate can be further improved, even up to 100%, with its fast search speed, it is bound to have positive reality significance in practical application.

DE Premature Convergence Analysis

In order to analyze the reason of premature convergence of DE/best/*/*, it need to study the change of population distribution in the process of evolution. Here the average-distance-amongst-points [7] is used to describe the degree of dispersion among populations.

Set *L* to be search space diagonal maximum length, *NP* is the size of the population size, *D* is the dimension of the solution space, and $p_{i,j}$ represents the *J-th* dimensional coordinate value of the *I-th* individual, $\overline{p}_j$ represents the average value of the *J-th* dimensional coordinate of all individuals. The average-distance-amongst-points is defined as follows.

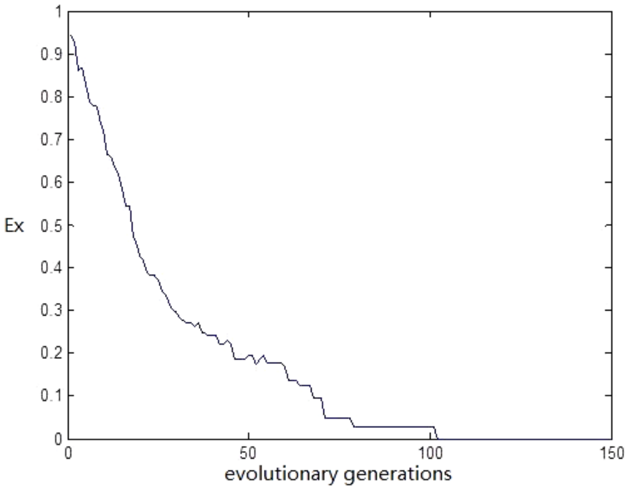
$$D_{is}\left(t\right) = \frac{1}{NP \times L}\sum_{i=1}^{NP}\left(\sum_{i=1}^{D}\left(p_{i,j} - \overline{p}_j\right)^2\right)^{1/2} \tag{1}$$

In Eq. 1, the population is more concentrated with smaller $D_{is}\left(t\right)$, but it is not perfect for the description of population diversity [8]. So the information entropy is introduced to describe the population diversity.

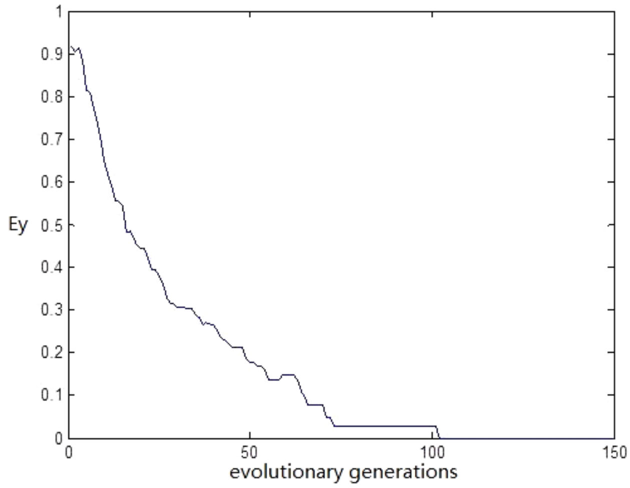
Assuming that the *G* generation group has *Q* subsets, make $S_{G1}, S_{G2}, \cdots, S_{GQ}$ each subset contains the number of individuals referred to as $\left|S_{G1}\right|$, $\left|S_{G2}\right|$, ⋯, $\left|S_{GQ}\right|$, And for any $p,q \in \{1,2,\cdots,Q\}$, $S_{Gp} \cap S_{Gq} = \Phi$, $\bigcup_{q=1}^{Q} S_{Gq} = S_G$, S_G is a collection of the *G-th* generation groups. The information entropy of the *G-th* generation group is as follows.

$$E_G = -\sum_{j=1}^{Q} p_j \ln p_j \tag{2}$$

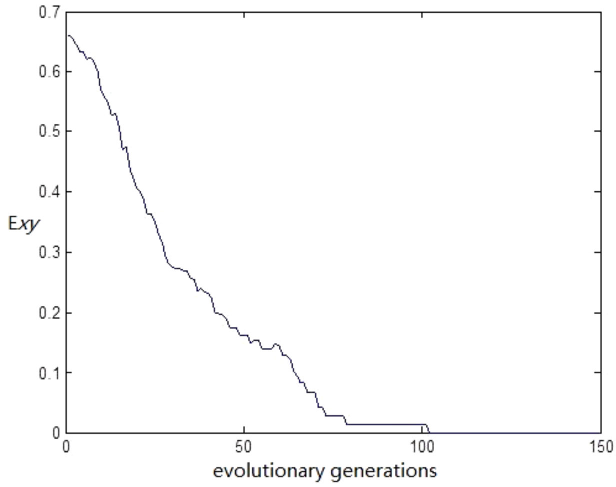
In Eq. 2, $p_j = \left|S_{Gj}\right| / N$, and *N* is the population size. It shows that the more types of individuals in a group the distribution is more uniform, and the greater the information entropy. The average-distance-amongst-points and information entropy are introduced in the DE algorithm. Numerical experiments are carried out with the Schaffer function and DE6 algorithm [9]. For Schaffer functions, define three kinds of information entropy: E_x is the information entropy of *x*, E_y is the information entropy of *y*, E_{xy} is the joint information entropy of *x* and *y*. The variations are shown in the following figures.



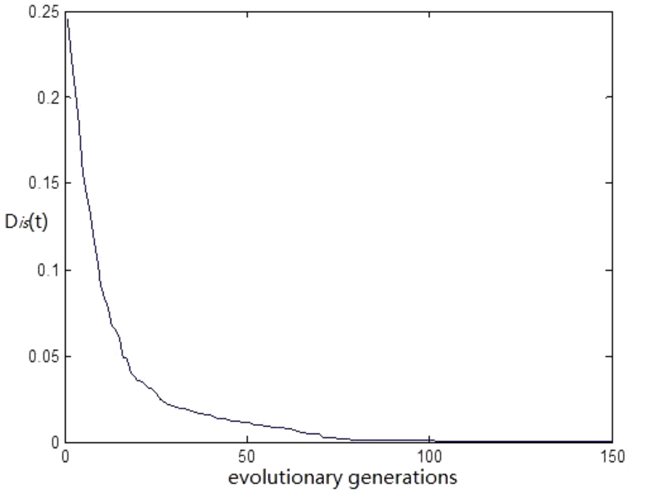
(a) The curve of *Ex*



(b) The curve of *Ey*



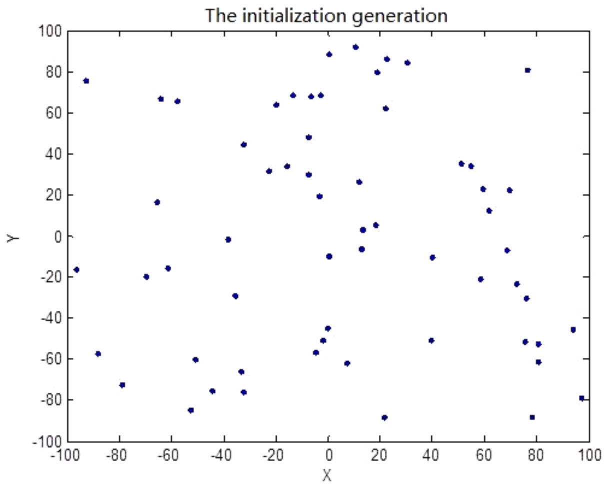
(c) The curve of *Exy*



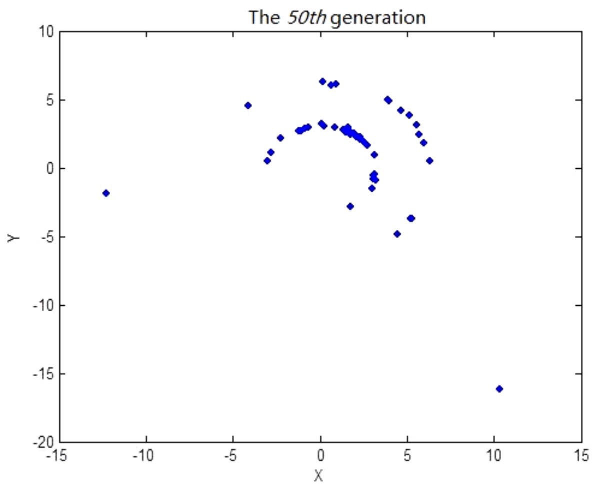
(d) The curve of *Dis*(t)

Fig.1 the curve of information entropy and average-distance-amongst-points

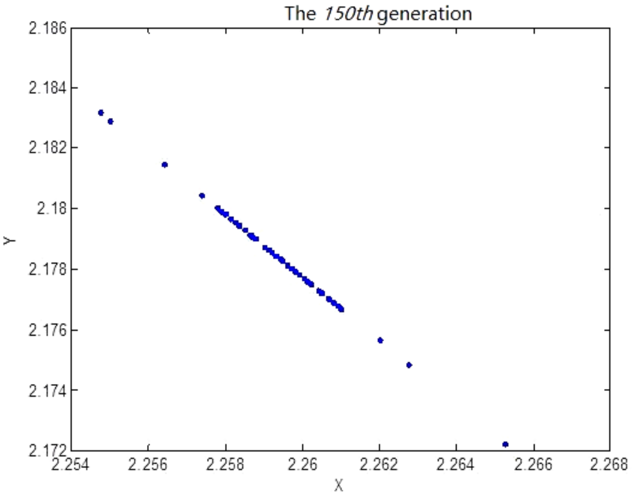
Fig.1 show that the optimization changing curves of the former *150th* generations. Which indicates the diversity of the population decrease dramatically with the evolution continues, and loses the diversity when evolution to a certain stage, the corresponding average-distance-amongst-points and information entropy close to zero.



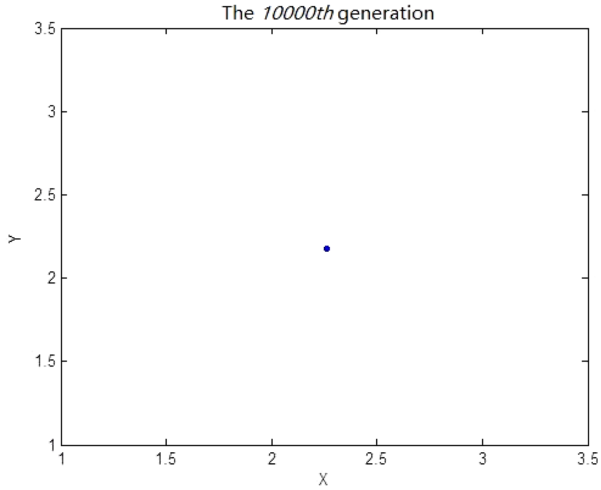
(a) The distribution of the initialization generation



(b) The distribution of the *50th* generation



(c) The distribution of the *150th* generation



(d) The distribution of the *10000th* generation

Fig.2 The change of Population distribution

Fig.2 show that the initialization population, the *50th*, the *150th*, the *10000th* generation Population distribution change. With the continuous deepening of the evolution process, the random distribution of the population in the plane [-100,100; -100,100] is gradually approaching to x=2.2593. The differences between individuals are becoming less and less, the population is rapidly moving towards the current optimal solution $x_{gbest,G}$, and all individuals tend to $x_{gbest,G}$=(2.2593,2.1785) that indicate premature convergence occurs.

Above the analysis, we can know that the rapid decrease of population diversity cause DE6 premature convergence. Due to the control of the current optimal individual, the population is rapidly moving towards the current optimal solution, Algorithm cannot jump out of the current optimal individual, and unable to search for other parts of the search space efficiently. Premature convergence occurs when the current optimal individual is the local optimum.Therefore, in order to overcome the premature convergence of DE/best/*/*, It is necessary to improve the diversity of the population and reduce the control effect of the current optimal individual on the algorithm.

Improved Algorithm Based On Elite Group

The idea of elite group has been successfully applied in particle swarm optimization algorithm, it can effectively increase the diversity of the population of particle swarm optimization algorithm and avoid premature phenomenon [10].

In order to make full use of the information of sub optimal individuals, improve the standard DE from the number of base vectors, the Algorithm idea is to select a certain number of sub optimal individuals and the current optimal individual components to form the elite group Ω_k , k is the number of elite group individuals. When the mutation operation is carried out, one member is randomly selected from the elite group as the base vector for each individual of the population, and generates mutation individuals, then to do crossover operator and select operation. Each member of the elite group has the opportunity to participate in the mutation operation as the current optimal individual of the population, and guides the search direction of algorithm.

Assumed that the optimization problem is to find the minimum value, if $x_i\left(i=1,2,\cdots,NP\right)$ meets constraint condition $f\left(x_1\right)\le f\left(x_2\right)\le\cdots\le f\left(x_{NP}\right)$. The elite group design as follows

$$\Omega_k=\left\{x_1,x_2,\cdots,x_k\right\},\quad k\le NP \tag{3}$$

According to ascending order sorts the fitness of all individuals in each iteration, and takes the former k individuals to form the elite group Ω_k as its basis vectors to generate mutation individual for each individual $x_{i,G}$. According to the requirements, the elite DE algorithm procedure is designed as follows.

Step1：Initialize the Initial population size *NP*, mutation factor *F*, crossover factor *CR*, and the Number of elite group *k*, Initialize population according to the equation as follows.

$$x_{ij} = rand[0,1] \cdot \left(x_j^{\ u} - x_j^{\ l}\right) + x_j^{\ l} \tag{4}$$

In Eq. 4, $i = 1,2,\cdots,NP$； $j = 1,2,\cdots,D$； $x_j^{\ u}$ is the upper bound of the *j-th* dimension variable, $x_j^{\ l}$ is the lower limit.

Step2：Calculate all individual fitness, and figure out the optimal fitness and optimal individual.

Step3： Determine whether the optimal individual objective function value reaches the accuracy requirement or reach the maximum evolutionary generations. If reaches the requirement, then exit; Otherwise, move to the next step.

Step4：According to the fitness ascending order, sorts all individuals and form the elite group Ω_k .

Step5：Randomly selects one individual from the elite group Ω_k as the basis vector of $x_{i,G}$, and to do mutation operation generate mutation individual.

Step6：Generate test individuals $v_{i,G+1}$ by crossover operation.

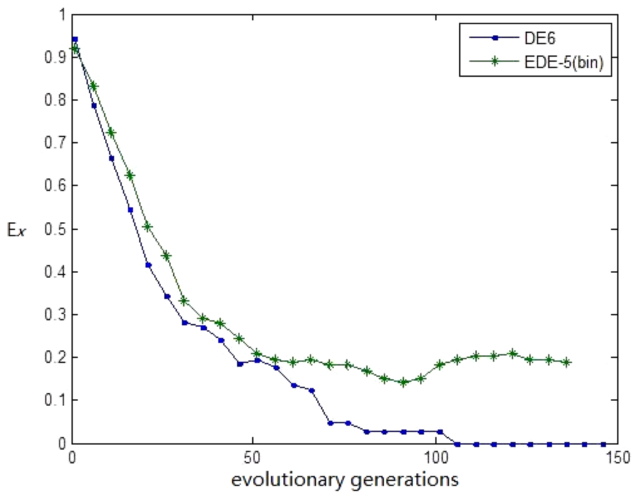
Step7：Generate the *(G+1)-th* generation individual by Select operations.

Step8： *G=G+1*，Return to Step2。

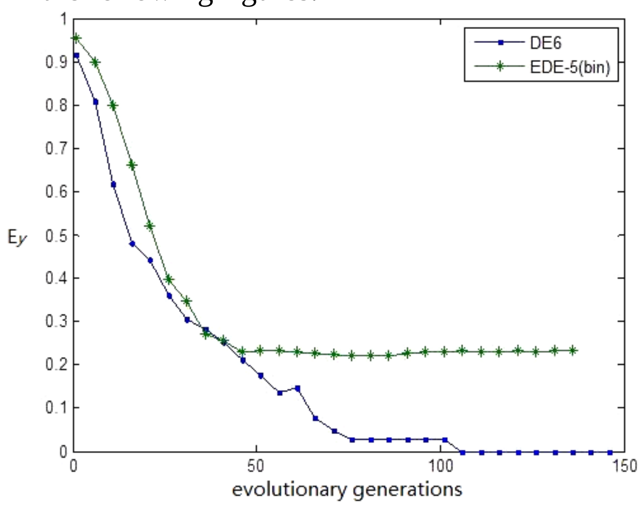
Can be seen from the above, when $k = 1$, the effect of elite group be shielded and the Elite DE algorithm just become DE / best /*; when $k = NP$, it is DE / rand /*. So DE / best /*and DE / rand /* are two special cases of the elite differential evolution algorithm.

Algorithm Test And Performance Analysis

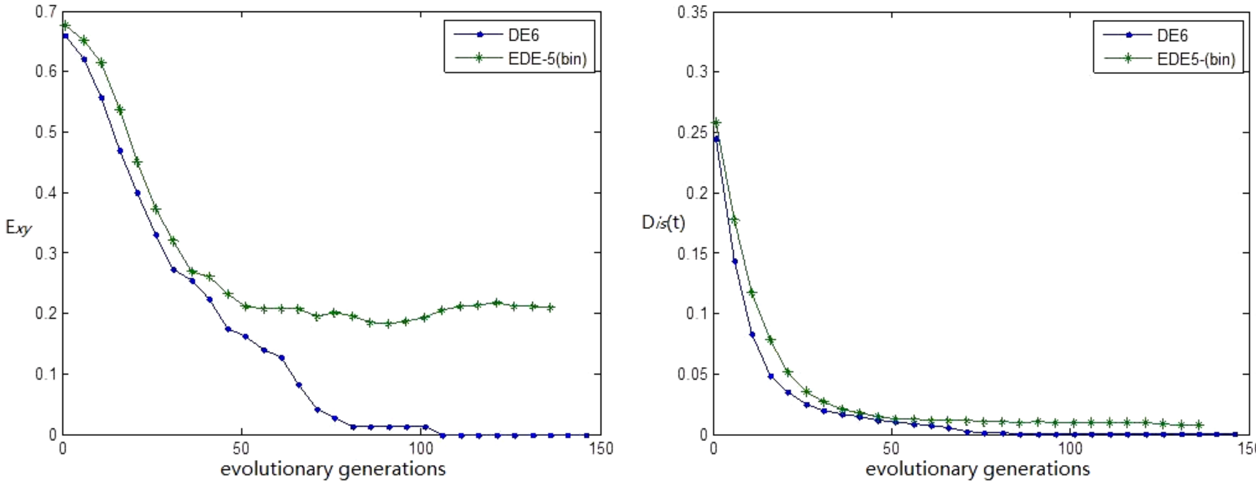
In order to verify the effectiveness of the elite differential evolution algorithm, select three typical test functions [11] to test and compare with the DE algorithm which are the Rosenbrock function (f_1), the Rastrigin function (f_2) and the Griewank function (f_3). The experimental parameters are set as follows：The test function dimension is 5, the Population size is set to 60, the maximum evolution generation is 2000, the mutation factor $F = 0.5$, cross factor $CR = 0.6$, the Number of individuals in the elite group $k = 5$, the algorithm terminates when the error is less than 10^{-12}. To facilitate the representation, DE6 represents DE/BEST/1/*, EDE represents the elite differential evolution algorithm in all figures. As shown in the following figure, EDE-5 (bin) algorithm finished evolution at the 136*th* generation and found the solution to meet the accuracy requirements. And DE6 algorithm trapped into local optimum. The results are shown in the following figures.



(a) The *Ex* curve of EDE-5(bin) and DE6



(b) The *Ey* curve of EDE-5(bin) and DE6



(c) The *Exy* curve of EDE-5(bin) and DE6 (d) The *Dis*(t) curve of EDE-5(bin) and DE6

Fig. 3 The evolution curve of EDE-5(bin) and DE6

Fig. 3 (a - c) show that the information entropy and joint information entropy of *X, Y* basically have the same variation trend. In the early stage of evolution, the information entropy curves of DE and EDE-5 are both decline fast. But compare with DE, the EDE-5 (bin) algorithm decreases gently with the evolution conducting, basically maintains turbulence near the fixed value, and loses the downward trend in the end. When the evolution is over, the Information entropy and XY joint information are maintained at stable level, it shows that the elite group can maintain certain diversity of population in the entire search space.

Fig. 3 (d) shows that the average-distance-amongst-points of DE6 and EDE-5 (bin) are both decrease. When the evolution to the *110th* generation, the average-distance-amongst-points of the DE6 algorithm basically to be zero. Although the EDE-5 (bin) algorithm reduce in the late, the rate of decrease is slow, and $D_{is}(t)$ is not zero in the end, which shows that the diversity of the population is maintained at a certain level.

Conclusion

Analyzed the cause of DE/best/* algorithm premature convergence by introducing the information entropy and the average-distance-amongst-points as the measure of population diversity, it is concluded that the main reason for premature convergence is the control function of the current optimal individual to decrease the population diversity of the algorithm. Using the idea of elite group to maintain population diversity, the elite differential evolution algorithm is proposed. The performance of the algorithm is analyzed by the typical test function, and the comparison with standard DE algorithm is carried out. The compute results show that the elite differential evolution algorithm can greatly improve the algorithm quality and search success rate.

References

[1] Rainer Storn and Kenneth Price. Differential Evolution-A simple and efficient adaptive scheme for global optimization over continuous spaces [R]. Technical Report, TR-95-012, ICSI, March 1995.

[2] Mahamed G.H.Omran, Andries P Engelbrecht. Self-Adaptive Differential Evolution Methods for Unsupervised Image Classification. IEEE 2006.

[3] Shahryar Rahnamayan, Hamid R.Tizhoosh, Magdy M.A Salama, "Opposition-based Differential Evolution for Optimization of Noisy Problem", 2006 IEEE Congress on Evolutionary Compution, pp 1865-1872.

[4] M.Fatih Tasgetiren, P.N Suganthan. A multi-populated Differential Evolution Algorithm for Solving Constrained Optimization Problem [C]. IEEE Congress on Evolutionary Computation, 2006:33-40.

[5] Janez Brest, Viljem Zmer, Mirjam Sepesy Maucec. Self-adaptive differential evolution algorithm in constrained real-parameter optimization [C]. IEEE Congress on Evolutionary Computation, 2006:215-222.

[6] Rainer Storn, Kenneth Price. Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces[J]. Journal of Global Optimization, 1997, 11(4) : 341- 359.

[7] Liu Xinying, Wang Shuhong, Qiu Jie. Improved particle swarm optimization with its application to parameter optimization of high temperature superconducting cable [J]. JOURNAL OFXI'ANJIAOTONG UNIVERSITY, 2007, 2(41):219-222,240.

[8] Jakob S.Vesterstrom, Jacques Riget and Thiemo Krink. Division of Labor in Particle Swarm Optimisation. 2002 IEEE, 1570-1575.

[9] ZHAO Guangquan, PENG Xiyuan, SUN Ning. A Modified Differential Evolution Algorithm with Hybrid Optimization Strategy [J].ACTA ELECTRONICA SINICA,2006,34(12A):2402-2405.

[10] Liu Yu, Qin Zheng, Shi Zhewen. Compact Particle SwarmOptimization Algorithm[J]. JOURNAL OF XI′ AN JIAOTONG UNIVERSITY, 2006, 40(8):883-887

[11] Zhao Guangquan. DIFFERENTIAL EVOLUTION ALGORITHM WITH GREEDY STRATEGY AND ITS APPLICATIONS:[D]. Harbin: Harbin Institute of Technology ,2007.