A Novel Algorithm for Multimodal Function Optimization Based on Evolution Strategy

Chang-Hwan Im, Hong-Kyu Kim, Hyun-Kyo Jung, *Senior Member, IEEE*, and Kyung Choi, *Member, IEEE*

Abstract—In this paper, a novel algorithm for multimodal function optimization is proposed, based on the concept of evolution strategy. A new concept, named restricted evolution, shows more improved characteristics than conventional approaches that have been used for multimodal function optimization. The efficiency and usefulness of the proposed method will be verified by the application to various cases including practical optimization problems.

Index Terms—Evolution strategy (ES), multimodal function optimization, restricted evolution.

I. INTRODUCTION

RECENTLY, algorithms to find more than one optimum of a function, usually referred to as a multimodal function optimization, have been widely studied. Most of the studies were based on the genetic algorithm (GA). Various concepts have been introduced to realize multimodal function optimization in GA, such as sharing [1], [2], deterministic crowding (DC) [3], niching method [4], restricted competition selection (RCS) [5], [6], and so on [7]–[9]. Basically, however, algorithms based on the GA tend to converge slowly, compared to other heuristic algorithms such as simulated annealing (SA) and evolution strategy (ES).

Nowadays, there have been some attempts at applying the ES to multimodal function optimization [10]–[12]. Most of the approaches adopted some concepts of GA, such as crossover and sharing. The common feature of the approaches was to cluster neighboring solutions around a peak. However, excessive application of the GA concept may lose the powerful advantage of the ES, the deterministic algorithm-like characteristic that yields faster convergence than the GA does.

In this paper, a novel algorithm for multimodal function optimization is developed, based on the concept of the ES. A new concept, named restricted evolution, is proposed to realize the multimodal function optimization scheme. The concept prevents solutions from clustering with their neighbors and allows only one solution to survive for each peak. Hence, the proposed concept is more efficient and practical than the conventional ones because a smaller population is required.

Moreover, it has additional advantages that it is easy to implement and shows fast convergence because it keeps most of the basic features of the ES. The efficiency and usefulness of the proposed method will be verified by the application to some mathematical functions and a practical electromagnetic optimization problem.

II. PROPOSED ALGORITHM

As stated briefly in the Introduction, the main feature of the proposed algorithm is the separation of neighboring solutions, which is very different from the conventional ES applications presented in [10]–[12]. In the case of the GA, the concept of the RCS is somewhat similar to that of the proposed method [5], [6]. However, in the case of the RCS-GA, the niche radius should be determined before the optimization process, which is very hard to do because the shape of an objective function cannot be estimated *a priori*. Whereas, for the proposed method, the evolution range can be modified during the optimization process by checking the convergence rate. The "self-adaptation" is the unique feature of the ES. In this study, the evolution range is used instead of the niche radius. The proposed method adopts the concept of an *elite set*, which stores superior solutions with some distance between each other. The members of the elite set are replaced with much superior or improved solutions during the evolution. The total process of the proposed algorithm is as follows:

1) Step 0—Initialization: Initialize $\alpha_{\min,i}, \alpha_{\max,i}$, and $\alpha_{\text{init},i}$ for each design variable.

α_i Evolution range for ith design variable. If the ith design variable is p_i, the child generation is generated within $[p_i - \alpha_i, p_i + \alpha_i]$, where $\alpha_{\min,i} < \alpha_i < \alpha_{\max,i}$. This factor is modified during the evolution and every elite set has different ranges except at the starting time.

$\alpha_{\min,i}$ Minimum distance between two elite solutions. This represents minimum distance between two closest extrema.

$\alpha_{\max,i}$ Maximum distance between two elite solutions.

$\alpha_{\text{init},i}$ Initial value for α_i.

2) Step 1—Generation of Initial Elite Set: Create the initial population and elite set. The size of the initial population is $\mu \times \lambda$. Among them, μ solutions are selected as members of an elite set. The elite set is determined by following rules:

a) put a currently best solution into the elite set and remove other solutions inside the evolution range of the best solution;
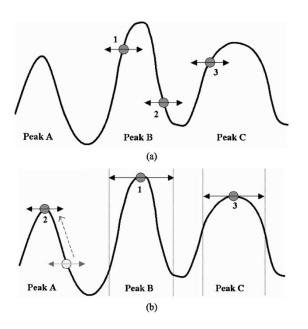
Fig. 1. Schematic illustrations to explain the performance of multimodal function optimization algorithm. (a) Initial solutions and evolution ranges. (b) Optimized result from proposed multimodal approach.

   b) find another best one except the removed solutions, and repeat the previous step until μ members are found.

3) Step 2—Generating Children, Restricted Evolution: Create λ new children within mutation range of each elite solution. This process is the same as $(1 + \lambda)$ ES. The restricted evolution ultimately finds a local optimum solution for each elite member, like deterministic search algorithms.

4) Step 3—Mutation: If the objective function for a child generation is improved compared to its parent, replace the parent (a member of elite set) with the child. After finishing the mutation, check if each elite solution is located in the mutation range of other elite solutions. If it is "true" and its objective function has a worse value than other solution, eliminate it (remove it from elite set). The number of removed solutions is denoted by ξ.

5) Step 4—Shaking: The $\xi + \rho$ solutions are randomly generated in the whole search space outside the mutation ranges of existing elite solutions (ρ is the number of shaking solutions). During the generation process, they should not invade other solutions' evolution range α_i.

6) Step 5—Annealing: Form a new elite set. The ξ removed solutions are replaced by the new solutions generated by the *shaking* process (the best ξ solutions are selected among the $\xi + \rho$ solutions). Existent elite solutions that are worse than remaining ρ solutions are also replaced with new solutions. If an elite solution is improved within the evolution range, the evolution range is increased by dividing it by 0.85. If the elite solution is not improved, the mutation range is decreased by multiplying it by 0.85. The initial evolution range $\alpha_{\text{init},i}$ is given to the newly generated solutions.

7) Step 6—Convergence Check: Repeat steps 2–6, until most solutions are not improved any more.

Fig. 1 shows the schematic illustrations to briefly explain the performance of the proposed multimodal optimization algorithm. Fig. 1(a) shows initial solutions and evolution ranges.

When a general evolution strategy is used, solutions 2 and 3 are eliminated and finally all the solutions are gathered around the global peak B. On the contrary, when the proposed algorithm is used, each solution searches its neighboring local peaks. Hence, solutions 1 and 3 converge to local peaks B and C, respectively. During the evolution, solution 2 approaches to the same peak (B) along with the solution 1 and is eliminated from the elite set. When the removed solution 2 is newly generated, it has more chance to be located around the empty peak A, as shown in Fig. 1(b).[1]

III. NUMERICAL TESTS AND RESULTS

The proposed algorithm was applied to some multimodal mathematical functions and a practical electromagnetic optimization problem.

A. Tests for Mathematical Functions

Two mathematical functions were tested: a test multimodal function and sinc function. The formula of the test multimodal function is

$$f(x) = 900 - \sum_{i=1}^{2} [(x_i - 5)^2 - 10 \times \cos(2\pi(x_i - 5))] \quad (1)$$

where $1 \leq x_i \leq 10$. The formula of the two-dimensional (2-D) sinc function is

$$f(x) = \sin \sum_{i=1}^{2} |x_i - 10| \bigg/ \sum_{i=1}^{2} |x_i - 10| \quad (2)$$

where $1 \leq x_i \leq 20$. Fig. 2(a) and (b) show the shapes of the test functions.

Table I shows the basic conditions to execute the proposed algorithm. Figs. 3 and 4 show the intermediate solutions for a test multimodal function and Sinc function, respectively. From the figures, we can see that the results show very fast convergence speed and the proposed method is very efficient for finding multiple peaks.

To verify the performance of the proposed algorithm more robustly, the number of variables for the test multimodal function was increased. The formula of the test multimodal function for N variables is

$$f(x) = 900 - \sum_{i=1}^{N} [(x_i - 5)^2 - 10 \times \cos(2\pi(x_i - 5))] \quad (3)$$

where $1 \leq x_i \leq 10$. The exact peak positions are already known as $x_i = 1, 2, \ldots, 10$. $N = 3, 4$ were tested repeatedly 100 times. The number of members in the elite set was 500 for $N = 3$ and 1000 for $N = 4$. The number of children per each elite member was 10 for $N = 3$ and 20 for $N = 4$. After 20 iterations, the number of found peaks was investigated. In the case of $N = 3$, 491.24 peaks were found among 500 elite members. For the case of $N = 4$, 975.71 peaks were found among 1000 elite members (491.24 and 975.71 are averaged values for 100 repeated simulations).

---

[1]The evolution ranges of solutions 1 and 3 generally increase before convergence, as shown in Fig. 1(b). Note that a new solution is generated outside the evolution ranges of existing elite members.
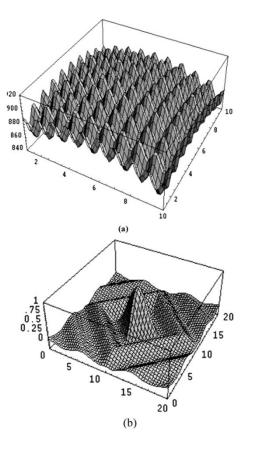
(a)

(b)

Fig. 2.   Mathematical multimodal functions. (a) Test multimodal function. (b) sinc function.
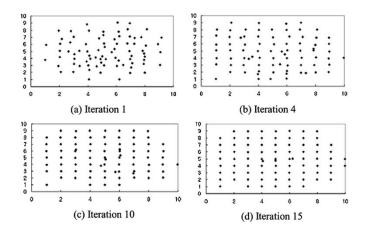

Fig. 3.   Optimization result for a test function ($\mu = 80, \lambda = 5, \rho = 10$).

Fig. 4.   Optimization result for a test function ($\mu = 80, \lambda = 5, \rho = 10$).

TABLE I
CONDITIONS TO EXECUTE THE PROPOSED ALGORITHM

Number of variables	2
Number of total iterations	100
Number of members in elite set (μ)	80
Number of children per each member (λ)	5
Number of initial population ($\mu \times \lambda$)	400
Number of shaking populations (ρ)	10

Although the proposed algorithm could find most of the peaks well, as shown in the previous simulations, the algorithm still has a problem. The problem stems from the unique feature of the ES—self-adaptation. When some elite members converged to their local optima, their evolution ranges decrease to $\alpha_{\min,i}$. When the value $\alpha_{\min,i}$ is a quite small value and most peaks were already found, newly generated solutions may not find new peaks. In other words, even when the number of peaks and the number of elite members are equal, every peak may not found. That is why some clustered solutions appear in Figs. 3 and 4. However, in practice, every peak does not need to be found and most peaks are already found before the freezing point. To confirm this property, additional simulations were performed. For the test multimodal function (1), two cases ($\mu = 121, \lambda = 5$, and $\mu = 50, \lambda = 5$) were tested and the results are demonstrated in Fig. 5. When the number of elite members was the same as that of the peaks, some solutions could not find empty peaks. However, when the number of elite members was smaller
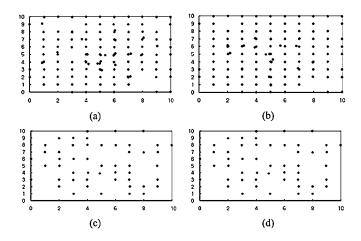

Fig. 5.   Comparison of performances when the number of elite members are different. (a) $\mu = 121$, iteration 10, 99 peaks were found. (b) $\mu = 121$, iteration 20, 103 peaks were found. (c) $\mu = 50$, iteration 10, 50 peaks were found. (d) $\mu = 50$, iteration 20, 50 peaks were found.

than that of peaks, every member found multiple peaks and converged very rapidly. The simulation results show that the proposed algorithms demonstrate much faster and more robust convergence characteristics when the number of elite members is smaller than that of the peaks, whereas conventional multimodal algorithms generally use a larger population than the number of peaks needed to be found, which is somewhat impractical.
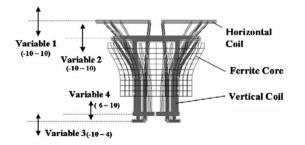
Fig. 6. Test deflection yoke model and design variables. This figure shows the original state, i.e., all values of the design variables are zero. Then, the design variables 1–4 implies the variance of the heights of each coil (Arrow implies the directions of the variance).

TABLE II
RANGES OF DESIGN VARIABLES

	Variable 1	Variable 2	Variable 3	Variable 4
Min. (mm)	-10	-10	-10	-6
Max. (mm)	10	10	4	10

TABLE III
CONDITIONS TO EXECUTE THE PROPOSED ALGORITHM

Number of variables	4
Number of total iterations	20
Number of members in elite set (μ)	10
Number of children per each member (λ)	5
Number of initial population ($\mu \times \lambda$)	50
Number of shaking populations (ρ)	2

B. Application to a Practical Optimization Problem

As a practical optimization example, a deflection yoke (DY) for color display tubes was selected. The objective functions to be evaluated were the beam shadow neck (BSN), trilemma, and sensitivity of horizontal coils, which are defined well in [13]. The optimum values of the BSN and trilemma are 5 and −1 mm, respectively, and the sensitivity should be minimized. Among the objective functions, the BSN and trilemma can be easily evaluated using simple numerical simulations during the optimization process. However, because the sensitivity cannot be calculated automatically, it cannot be included in the optimization process [13]. Fortunately, there are some bounds on the specification of the BSN and trilemma (BSN: 4–6 mm, trilemma: −2–0 mm). In this study, after finding multiple solutions that satisfy the design margins of the BSN and trilemma, sensitivities for each possible solution were compared. Fig. 6 presents initial state of the DY and definitions of four design variables that affect the objective function most dominantly. If the variables change, automatic winding modeler generates a new winding structure and the value of the BSN and trilemma are updated. Tables II and III show the ranges of four design variables and conditions to execute the proposed optimization algorithm, respectively. The objective function to be maximized was defined as

$$F = 1/(1 + |\text{BSN} - 5| + |\text{Trilemma} + 1|). \qquad (4)$$

TABLE IV
OPTIMIZATION RESULTS

Solutions	1	2	3	4	5
Variable 1	1.44	-2.48	0.42	7.79	2.09
Variable 2	6.72	7.98	8.34	1.23	9.10
Variable 3	-5.33	-9.20	-8.74	-6.38	-7.66
Variable 4	7.27	5.04	9.31	8.15	9.57
BSN (mm)	5.09	5.73	5.10	5.02	5.34
Tri-lemma (mm)	-1.15	-1.25	-0.88	-0.76	-0.97
Sensitivity (mHA^2)	0.88	0.83	0.59	0.74	0.69

Table IV shows the results of the optimization. Among ten elite solutions five candidate ones that satisfy the design margins were selected. After that, the sensitivity was evaluated for each candidate solution. From the table, we can see that solution 3 minimizes the sensitivity while satisfying the design margins. Hence, we selected the solution 3 as the optimal one.

IV. CONCLUSION

In this paper, a novel algorithm for multimodal function optimization was proposed. The algorithm is based upon the combination of the evolution strategy and the restricted evolution. The proposed method was verified by applying to various cases and showed very good performances.

REFERENCES

[1] D. E. Goldberg and J. Richardson, "Genetic algorithms with sharing for multimodal function optimization," in *Proc. 2nd Int. Conf. Genetic Algorithms (ICGA)*, 1987, pp. 41–49.
[2] B. L. Miller and M. J. Shaw, "Genetic algorithms with dynamic niche sharing for multimodal function optimization," in *Proc. IEEE Int. Conf. Evolutionary Computation*, 1996, pp. 786–791.
[3] S. W. Mahfoud, "Niching methods for genetic algorithms," Ph.D. dissertation, Illinois Genetic Algorithm Lab., Univ. Illinois, Urbana, IL, 1995.
[4] B. Sareni and L. Krähenbühl, "Fitness sharing and niching methods revisited," *IEEE Trans. Evol. Comput.*, vol. 2, pp. 97–106, Sept. 1998.
[5] C.-G. Lee, D.-H. Cho, and H.-K. Jung, "Niching genetic algorithm with restricted competition selection for multimodal function optimization," *IEEE Trans. Magn.*, vol. 34, pp. 1722–1725, May 1999.
[6] J.-K. Kim, D.-H. Cho, H.-K. Jung, and C.-G. Lee, "Niching genetic algorithm adopting restricted competition selection combined with pattern search method," *IEEE Trans. Magn.*, vol. 38, pp. 1001–1004, Mar. 2002.
[7] J. P. Li, M. E. Balazs, G. T. Parks, and P. J. Clarkson, "A species conserving genetic algorithm for multimodal function optimization," *Evol. Computat.*, vol. 10, no. 3, pp. 207–234, 2002.
[8] H. Qi, "A genetic algorithm with tabu list and sharing scheme for optimal design of electrical machines," *Elect. Mach. Power Syst.*, vol. 27, no. 5, pp. 543–552, May 1999.
[9] V. Rupela and G. Dozier, "Parallel and distributed evolutionary computations for multimodal function optimization," in *Proc. 5th Biannu. World Automation Congr.*, vol. 13, June 2002, pp. 307–312.
[10] J. Zhang, X. Yuan, Z. Zeng, and S. Amer, "Niching in as ES/EP context," in *Proc. Congr. Evolutionary Computation*, vol. 2, 1999, pp. 1426–1433.
[11] O. Takahashi and S. Kobayashi, "An adaptive neighboring search using crossover-like mutation for multimodal function optimization," in *Proc. IEEE Int. Conf. Systems, Man, and Cybernetics*, vol. 1, 2001, pp. 261–267.
[12] O. Aichholzer *et al.*, "Evolution strategy and hierarchical clustering," *IEEE Trans. Magn.*, vol. 38, pp. 1041–1044, Mar. 2002.
[13] C.-H. Im, H.-K. Kim, and H.-K. Jung, "Optimization of the coil shape in deflection yoke considering practical coil winding processes," *IEEE Trans. Magn.*, vol. 38, pp. 1077–1080, Mar. 2002.