Scientific Research

Optimization Scheme Based on Differential Equation Model for Animal Swarming

Takeshi Uchitane, Atsushi Yagi

Department of Information and Physical Sciences, Osaka University, Osaka, Japan
Email: atsushi-yagi@ist.osaka-u.ac.jp

ABSTRACT

This paper is devoted to introducing an optimization algorithm which is devised on a basis of ordinary differential equation model describing the process of animal swarming. By several numerical simulations, the nature of the optimization algorithm is clarified. Especially, if parameters included in the algorithm are suitably set, our scheme can show very good performance even in higher dimensional problems.

Keywords: Optimization Scheme; Differential Equation Model; Animal Swarming

1. Introduction

Let $f(x)$ be a real valued continuous function defined for $x \in \mathbb{R}^D$, where D is a positive integer. The problem of finding the minimal value

$$\min_{x \in \mathbb{R}^D} f(x)$$

is called a *D*-dimensional optimization problem for $f(x)$. Such an optimization problem is one of fundamental problems in the study of science and technology and so far various methods have already been presented.

When $f(x)$ is a C^1 function, the point $\bar{x}$ which hits the minimal value $f(\bar{x}) = \min_{x \in \mathbb{R}^D}\{f(x)\}$ is obtained in the set of solutions of the functional equation $\nabla f(x) = 0$. But, in general, it is not so easy to solve this functional equation. So, approximate solutions become more important. The steepest descent method is a method of obtaining a sequence $x_n (n = 0,1,2,\cdots)$ in $\mathbb{R}^D$ which approaches to $\bar{x}$ by using the recurrence formula $x_{n+1} = x_n - \alpha_n \nabla f(x_n)$ with some suitable coefficient $\alpha_n > 0$. This method is very convenient and it is easily observed that x_n converges to a limit x_∞ such that $\nabla f(x_\infty) = 0$. But x_∞ may hit very often some local minimal value of $f(x)$ and not the very minimal value. So, when $f(x)$ possesses many local minimal values, it is very difficult to find out the point $\bar{x}$ by this method.

Recently, a development of techniques of numerical computations has yielded a new paradigm of optimiza-tion using a collection of particles in $\mathbb{R}^D$ which interact each other and move for striking out the point $\bar{x}$. Such a method is called the particle optimization. One of the typical particle optimization was devised by Kennedy-Eberhart [1]. They consider a swarm of particles not only flying in $\mathbb{R}^D$ like bird but also behaving as an intellectual individual which can memorize its personal best through the whole past and, on the other hand, can know the swarm's global best at each instant. In each step, processing such personal and swarm information, they move to a suitable position. The phase space in which the particles move about is therefore an abstract multi-dimensional space which is collision-free. They called their method the particle swarm optimization. Afterword, the particle swarm optimization has been developed extensively and applied to various problems. We will here quote only some of them [2-4]. A survey of the method has recently been published by Eslami-Shareef-Khajeh-zadeh-Mohamed [5].

In this paper, we want to compose another type particle optimization which is inspired more directly from the animal swarming like fish schooling, bird flocking, or mammal herding. Our particles move truly according to the animal's behavioral rules for forming swarm. We first set a *D*-dimensional physical space $\mathbb{R}^D$ in which the particles move about. We then assume among individuals two kinds of interactions, attraction and collision avoidance. These interactions will be formulated by generalized gravitation laws. Regarding $f(x)$ as an environ-

mental potential, we assume also that the particles are sensitive to the gradient of $f(x)$ at their positions and have tendency to move toward the most descending direction of the value of $f(x)$. But they cannot know which mate has the global best position of the swarm at any momemt. We will also incorporate uncertainty of their information processing and executing. The phase space of particles is therefore given by $\mathbb{R}^D \times \mathbb{R}^D$, here $(x_i, v_i) \in \mathbb{R}^D \times \mathbb{R}^D$ means that at that moment the i-th particle is at position x_i with velocity v_i. As for animal's behavioral rules, we are going to follow those presented by Camazine-Deneubourg-Franks-Sneyd-Theraulaz-Bonabeau ([6], Chapter 11). That is:

1) The swarm has no leaders and each individual follows the same behavioral rules;

2) To decide where to move, each individual uses some form of weighted average of the position and orientation of its nearest neighbors;

3) There is a degree of uncertainty in the individual's behavior that reflects both the imperfect information-gathering ability of an animal and the imperfect execution of its actions.

(We refer also a similar idea due to Reynolds [7].) The authors have in facttried to model in the previous paper [8] these mechanisms by stochastic differential Equations (2.1) and (2.2) below.

Our optimization scheme is actually composed on the basis of the continuous model Equations (2.1), (2.2) but just ignoring velocity matching (*i.e.*, taking $\beta = 0$) and setting the external force as the sum of resistance and gradient of the potential function to be optimized. At each step, the particles have a velocity determined by the sum of centering with nearby mates and acceleration by the external force. Its position is renewed by the sum of the velocity and a noise which reflects the imperfectness of information-gathering and execution of actions. At the first stage, the particles move striking out the global minimal value on one hand keeping a swarm and on the other hand keeping a territorial distance with other mates. The noise helps the swarm to escape from the traps of local minimal values and to reach into a neighborhood of the global one. At the second stage, the movement of particles slows down. The particles go to an equilibrium state in which some particle attains at the swarm's best value.

In order to investigate swarm behavior of particles, we shall apply our optimization scheme to a few benchmark problems. When $f(x)$ has many local minimal values, the territorial distance must be chosen in a suitable length to find a good approximate solution. Optimal strength of noise is also required. If it is too small, the swarm is easily trapped by a local minimum; to the contrary, if too large, the particles cannot keep on swarming because of the strong dispersion. If these are suitably set, then our

scheme can show very high performance even in 12-dimensional problems.

2. Continuous Model

We being with reviewing the continuous model presented by [8].

We consider motion of I animals in the physical space $\mathbb{R}^3$. The position of i-th animal is denoted by $x_i = x_i(t)$ $(i = 1, \cdots, I)$, and its velocity by $v_i = v_i(t)(i = 1, \cdots, I)$. The model equations are then written as

$$dx_i = v_i dt + \sigma_i dB_i \tag{2.1}$$

$$dv_i = \left\{ -\alpha \sum_{j \neq i} \left[\left(\frac{r}{x_{ij}} \right)^p - \left(\frac{r}{x_{ij}} \right)^q \right] x_{ij} \right.$$
$$\left. -\beta \sum_{j \neq i} \left[\left(\frac{r}{x_{ij}} \right)^p - \left(\frac{r}{x_{ij}} \right)^q \right] v_{ij} + f_i(t, x_i, v_i) \right\} dt, \tag{2.2}$$

where $x_{ij} = x_i - x_j$ and $v_{ij} = x_i - v_j$. The first Equation (2.1) is a stochastic equation for x_i, here $\{B_i(t), t \geq 0\}(i = 1, \cdots, I)$ denotes a system of independent 3-dimensional Brownian motions defined on a complete probability space with filtration $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ (see [9]). The term $\sigma_i dB_i$ therefore denotes a noise resulting from the imperfectness of information-gathering and action of the i-th animal, σ_i being some coefficient. In the meantime, (2.2) is a deterministic equation for v_i. The first term in the right hand side denotes the centering and the collision avoidance of animal. The animals have tendency to stay nearby their mates and at the same time avoid colliding each other. As p and q are such that $1 < p < q < \infty$, if $\|x_i - x_j\| > r$, then the i-th animal moves toward the j-th; to the contrary, if $\|x_i - x_j\| < r$, then it acts in order to avoid collision with the other. The number $r > 0$ therefore denotes a critical distance. If p is large, then, as the distance $\|x_i - x_j\| \geq r$ increases, its power $\left(r / \|x_i - x_j\| \right)^p$ decreases quickly. Hence, the larger p is, the shorter the range of centering is. The second term of (2.2) denotes the effect of velocity matching with nearby mates. Finally, the term $f_i(t, x_i, v_i)$ denotes an external force imposed to the i-th animal at time t which is a given function for x_i and v_i. In the subsequent section, we take a function of the form

$$f_i = -cv_i - \gamma \nabla f(x_i), \tag{2.3}$$

where $-cv_i$ denotes a resistance for motion and $-\gamma \nabla f(x_i)$ denotes an external force determined by a $\mathbb{C}^1$ potential function $f : \mathbb{R}^3 \to \mathbb{R}$.

3. Optimization Scheme

Let $f(x)$ be a real valued $\mathbb{C}^1$ function defined for $x \in \mathbb{R}^D$ with positive integer D. Consider the optimi-

zation problem

$$\min_{x \in \mathbb{R}^D} f(x_i). \qquad (3.1)$$

On the basis of (2.1), (2.2), we introduce an optimization algorithm for (3.1). Let $x_i(t), i = 1, \cdots, I$, denote the positions of I particles moving about in the space $\mathbb{R}^D$, and let $v_i(t), i = 1, \cdots, I$, denote their velocities. We ignore velocity matching of particles, namely, we set $\beta = 0$, and take the function $f_i(t, x_i, v_i)$ in (2.2) in the form (2.3) above. Our scheme is then described by

$$x_i(t + \Delta t) = x_i(t) + v_i(t + \Delta t)\Delta t + \sigma \xi_i(\Delta t), \qquad (3.2)$$

$$v_i(t + \Delta t) = w(\Delta t)v_i(t)$$
$$- \left\{ \alpha \sum_{j \neq i} \left[\left(\frac{r}{x_{ij}(t)} \right)^p - \left(\frac{r}{x_{ij}(t)} \right)^q \right] x_{ij}(t) - \gamma \nabla f[x_i(t)] \right\} \Delta t, \qquad (3.3)$$

where $x_{ij}(t) = x_i(t) - x_j(t)$. In addition, $\xi_i(\Delta t) = B_i(t + \Delta t) - B_i(t), i = 1, \cdots, I$, is a family of independent stochastic functions defined on a complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ with values in $\mathbb{R}^D$ whose distributions are a normal distribution with mean 0 and variance Δt. And $w(\Delta t)$ is given by $w(\Delta t) = 1 - c\Delta t$.

Set initial positions and initial velocities

$$\boldsymbol{x}_0 \equiv (x_1(0), \cdots, x_I(0)) \in \mathbb{R}^{DI},$$
$$\boldsymbol{v}_0 \equiv (v_1(0), \cdots, v_I(0)) = \boldsymbol{0} \in \mathbb{R}^{DI},$$

respectively. Then, the algorithm (3.2), (3.3) defines a discrete trajectory

$$\boldsymbol{x}_n \equiv (x_1(t_n), \cdots, x_I(t_n)) \in \mathbb{R}^{DI},$$

where $t_n = n\Delta t, n = 0, 1, 2, \cdots, N$. In each step n, we compute the minimal value $f_n \equiv \min_{1 \leq i \leq I} f[x_i(t_n)]$ and memorize its value together with the point $x_n \in \mathbb{R}^D$ hitting it, i.e., $f(x_n) = f_n$. Repeating the iteration N-times, $\overline{f} = \min_{0 \leq n \leq N} f_n$ is an approximation value of (3.1) and $\overline{x}$ such that $f(\overline{x}) = \overline{f}$ is an approximation solution of our scheme.

As $\overline{x}$ is one of members of swarm $x_1, \cdots, x_I$ having interactions one and another, the approximate solution $\overline{x}$ may not satisfy the condition $\nabla f(\overline{x}) = 0$. This means that there is a point $\overline{x}'$ in a neighborhood of $\overline{x}$ which hits a local minimum of $f(x)$, i.e., $\nabla f(\overline{x}') = 0$. In this case, $\overline{x}'$ which can easily be obtained by classical methods (e.g., the steepest descent method) gives a better approximate solution of (3.1) than $\overline{x}$.

4. Numerical Experiments

We show some numerical experiments to expose the par-

ticle's behavior of our scheme. It is expected that the particles crowd around a point $\overline{x}$ giving the global minimum of $f(x)$ or are dispersed into a number of neighborhoods of points giving local minimums. The behavior may change heavily depending on the choice of parameters p, q, r and σ.

We here use three well known benchmark problems, namely, Problem (3.1) with Sphere function, Rastrigin function and Rosenbrock function. Problem (3.1) with the Sphere function

$$f(x) = \sum_{d=1}^{D} x_d^2 \qquad (4.1)$$

is the simplest problem. The optimal point is obviously given by $0 = (0, \cdots, 0)$. The following function

$$f(x) = 10D + \sum_{d=1}^{D} \left[x_d^2 - 10\cos(2\pi x_d) \right] \qquad (4.2)$$

is called the Rastrigin function. Its optimal point is also given by $0 = (0, \cdots, 0)$ with the global minimum $f(0) = 0$. It is easily seen that this function possesses many local minimums; indeed, at every lattice point, $f(x)$ has a local minimum. Therefore, Problem (3.1) with function (4.2) is very difficult to treat especially in a high dimension D. Finally, the Rosenbrock function is formulated by

$$f(x) = \sum_{d=1}^{D-1} \left[100(x_{d+1} - x_d^2)^2 + (x_d - 1)^2 \right]. \qquad (4.3)$$

This function takes its global minimum 0 at the lattice point $e = (1, \cdots, 1)$.

As shown below, behavior of the particles depends on the model parameters. We set $I = 9, 25, 49; w = 0.0, 0.9; \sigma = 3000, 10000, 30000; r = 0.5, 1.0, 2.0;$ and $(p, q) = (3, 5), (7, 9)$. The step size is fixed by $\Delta t = 1.0 \times 10^{-6}$ and the total step number is $N = 5000$. As for D, we consider 2-dimensional and 12-dimensional problems.

When $D = 2$, the initial position $\boldsymbol{x}_0$ is taken in, $[8, 21]^{2I}$, where no optional point exists in $[8, 12]^2$, moreover, $\boldsymbol{x}_0$ is far away from the optimal point. In addition, for 2-dimensional Rosenbrock problem, we will take $\boldsymbol{x}_0 \in [-12, -8]^{2I}$, too, because the Rosenbrock function defined by (4.3) is not symmetric with respect to the transformation $x \mapsto -x$.

When $D = 12$, the initial position $\boldsymbol{x}_0$ is taken randomly in $[-2, 2]^{12}$, but it is very difficult to find the optimal point in a higher dimensional search space.

4.1. Sphere Function

Figure 1 shows the numerical results for Sphere function. The parameters are set as $I = 25$, $w = 0.9$; $\sigma = 3000$; $r = 0.5$, and $(p, q) = (3, 5)$. For $D = 2$, the computed

result is given by **Figure 1(a)**; similarly for $D = 12$, by **Figure 1(b)**.

In the first stage $(0 \leq n \leq 1500)$, $f(x_n)$ decreases rapidly, which means that all the particles strike out for the optimal point 0 in group influenced form the force due to the gradient of $f(x)$. In the second stage, since the interaction among particles becomes dominant rather than the potential force, decreasing of the value $f(x_n)$ becomes slow. This means that the particles make local searches in a neighborhood of the global optimum. But, the interaction may prevent the optimal particle from attaining exactly the global minimum. So, our method makes it possible to search both wide range and small range by using the same scheme. On the other hand, it is an important problem to find a better combination of parameters in order to have a suitable balance between the global search and the local search.

4.2. Rastrigin Function

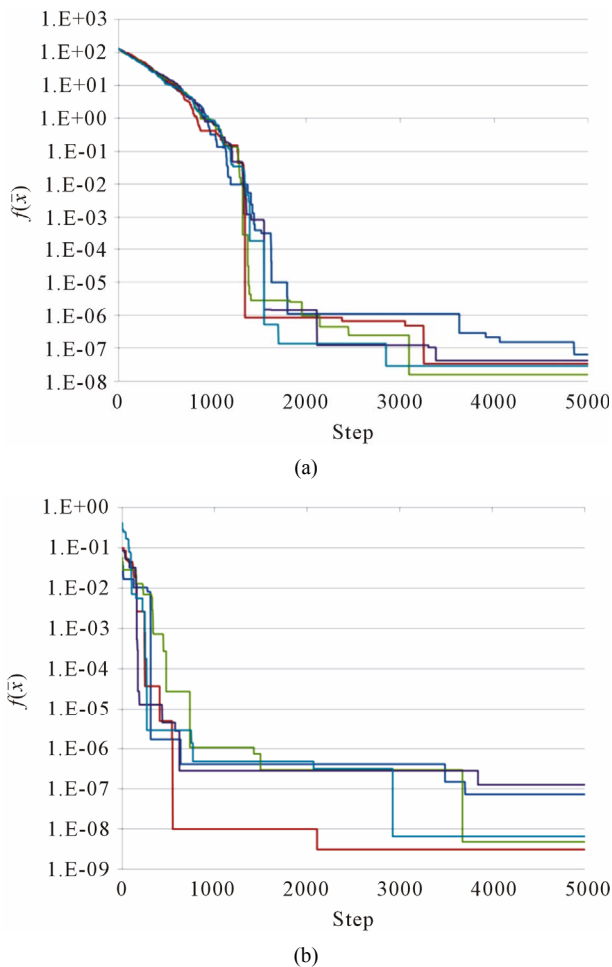**Figure 2** shows the numerical results for the Rastrigin

(a)

(b)

Figure 1. Graph of $f(x_n)$ **at Steps** $0 \leq n \leq 5000$ **for 5 trials. (a) Sphere function for** $D = 2$; **(b) Sphere function for** $D = 12$.
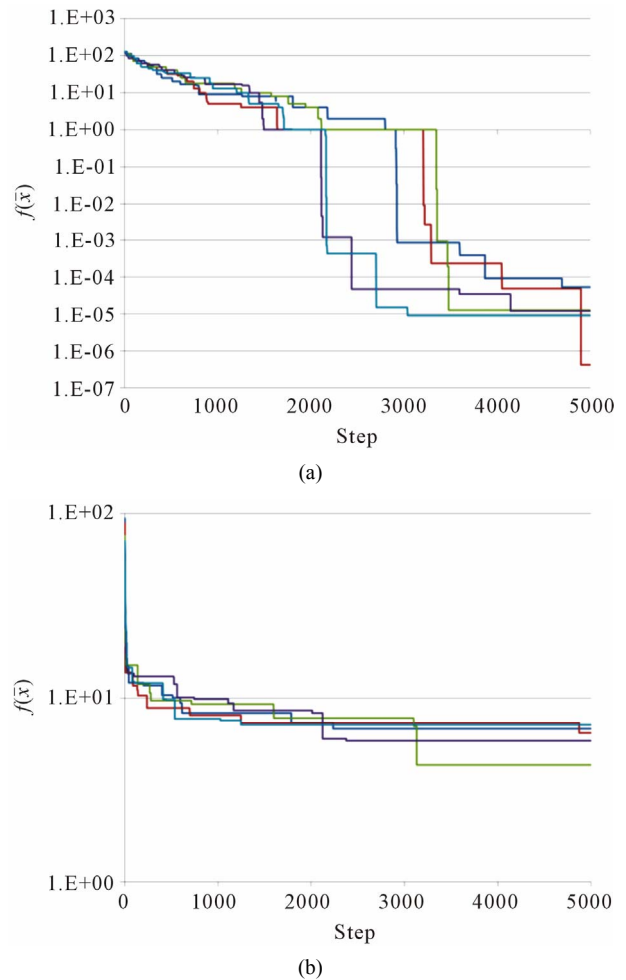

(a)

(b)

Figure 2. Graph of $f(x_n)$ **at Steps** $0 \leq n \leq 5000$ **for 5 trials. (a) Rastrigin function for** $D = 2$; **(b) Rastrigin function for** $D = 12$.

function. For $D = 2$, we set $I = 49$, $w = 0.9$; $\sigma = 10000$; $r = 1.0$, and $(p,q) = (7,9)$. For $D = 12, r$ and (p,q) are replaced by $r = 2.0$ and $(p,q) = (3,5)$, respectively, and the rests are the same as for $D = 2$. Such a combination yielded the best result in our numerical simulations.

Figures 2(a) and **(b)** illustrate the graphs of $f(x_n)$ for 5 trials in the cases $D = 2$ and $D = 12$, respectively. **Figure 3** then shows the positions of particles at Step $N = 5000$. (When $D = 12$, only the first two coordinates are illustrated.) In the 2-dimensional case, many particles reach a neighborhood of the optimal point $(0,0)$ without being trapped by lattice points giving local minima. In the 12-dimensional case, however, no particles can exactly get to the optimal point. But, we see that, for the optimal point $\overline{x}$ of the scheme, most of its coordinates coincide optimally with 0.

In order to measure the total optimality of swarm, we want to consider the total energy defined by
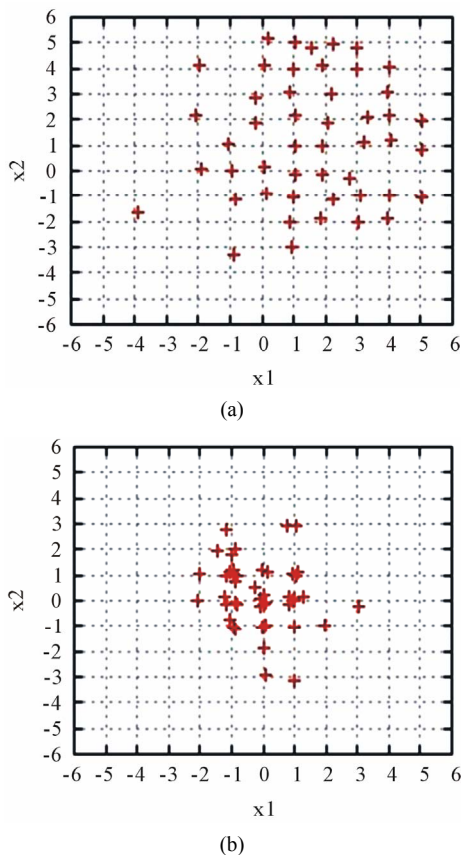
(a)

(b)

Figure 3. Positions of particles at Step *N* = 5000. (a) Rastrigin function for *D* = 2; (b) Rastrigin function for *D* = 12.

$$E_n = \sum_{i=1}^{I} f\left[x_i\left(t_n\right)\right].$$

We can say that, if E_n is small, then most of swarm particles stay in rather better positions and the swarm is in a steady state. To the contrary, if E_n is large, then many swarm particles move actively from local minima to others to reach a steady state. **Table 1** shows some summation results of E_{5000}. As seen by this table, if σ is too small, namely, the noise is too small, then the swarm is trapped in a steady state of high total energy. To the contrary, if σ is sufficiently large, then particles can escape from the trap of local minima and the swarm moves into a state of lower total energy. But if σ is too large, then the noise effect disperses swarm particles in every direction to disconnect the interaction among swarm mates. In this simulation, the value $\sigma = 10000$ is the best one.

4.3. Rosenbrock Function

Figure 4 shows the numerical results for the Rosenbrock function. For $D = 2$, we set $I = 49$, $w = 0.9$, $\sigma = 3000$, $r = 0.5$, and $(p, q) = (3, 5)$. The initial positions are set in either $K_+ = [8, 12]^{98}$ or $K_- = [-12, -8]^{98}$. For $D = 12$,

Table 1. Total energy of particles at Step *N* = 5000.

σ		1test	2test	3test	4test	5test
0	sum	938	973	973	1001	935
	min	9.02	8.00	10.99	8.02	8.03
3000	sum	1082	1106	1091	1147	1053
	min	10.09	9.58	10.83	9.89	11.26
10000	sum	2508	2655	2593	2311	2523
	min	21.89	24.36	16.79	16.51	22.57
30000	sum	11732	18830	18892	19219	19046
	min	147.63	222.27	198.41	195.95	181.32

we set $I = 49$, $w = 0.9$, $\sigma = 3000$, $r = 1.0$, and $(p, q) = (7, 9)$. Such a combination provided the best result in our numerical simulations.

Figures 4(a) and **(b)** illustrate the graphs of $f\left(x_n\right)$ for the 2-dimensional problems with initial positions $x_0 \in K_+$ and $x_0 \in K_-$, respectively. **Figure 4(c)** illustrates the same for the 12-dimensional problem. **Figure 5** then shows the positions of particles at Step $N = 5000$. (When $D = 12$, only the first two coordinates are illustrated.) We observe that particles line along a curve drawn on $\mathbb{R}^D$ by the local minima of $f(x)$ defined by (4.3).

5. Conclusions

After reviewing the stochastic differential equation model for animal swarming, we have introduced our optimization algorithm directly based on animal swarming behavior, centering, collision avoidance and imperfectness of information-gathering and acting. The optimization function was treated as a potential function and each particle has tendency to move into the most descending direction of the values of the potential function. By these, on one hand, the particles keep on swarming; on the other hand, they search the optimal point globally and locally without being trapped by points giving local minima.

Numerical experiments were performed in order to clarify particle's behavior of the proposed scheme. We used Sphere function, Rastrigin function and Rosenbrock function. For getting good performance, the parameters of algorithm must be suitably chosen.

Especially, (p, q) controls an effective range of how far the interaction among particles is attainable. Meanwhile, r decides the standard distance of any two particles; that is, if the distance of two particles is more than r, then they are attractive each other, on the contrary, if less than r, then they are repulsive. The parameter σ controls vitality of the swarm. If σ is too small, then the
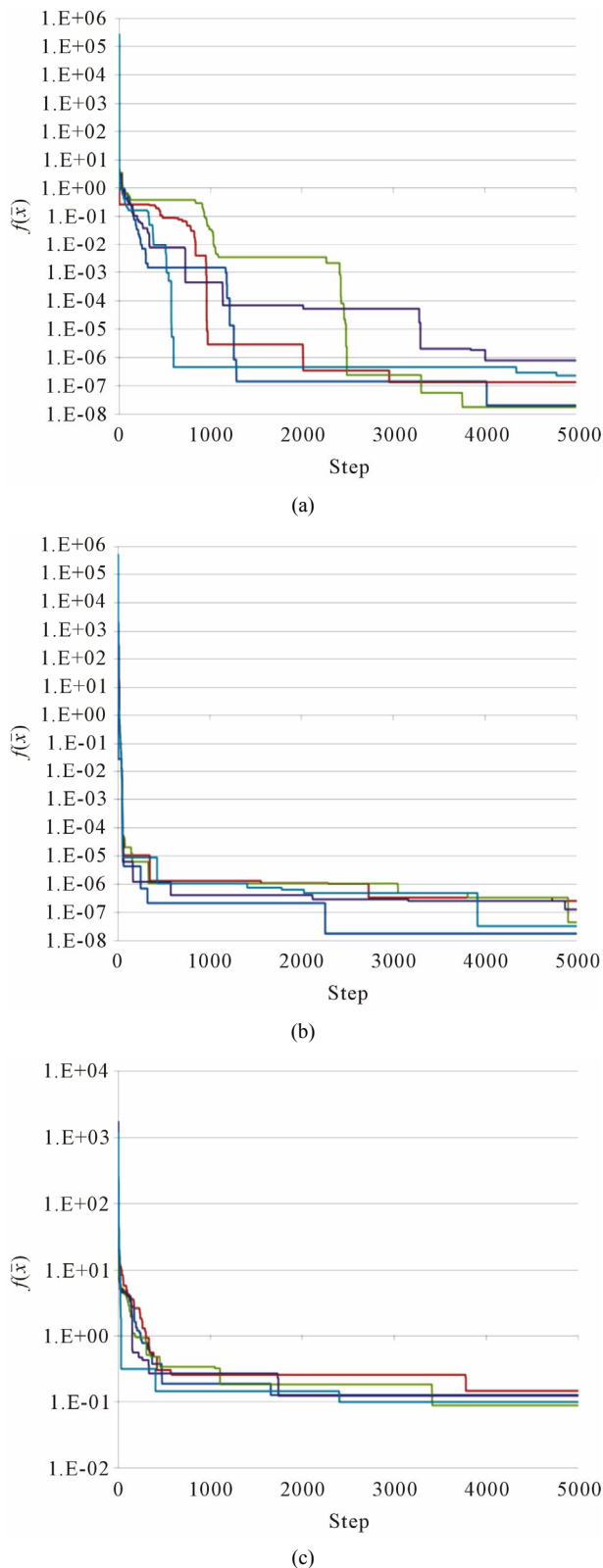
T. UCHITANE, A. YAGI

(a)

(b)

Figure 4. Graph of $f(x_n)$ **at Steps** $0 \le n \le 5000$ **for 5 trials.** **(a) Rosenbrock function for** $D = 2$ **with** $x_0 = K_+$; **(b) Rosenbrock function for** $D = 2$ **with** $x_0 = K_-$; **(c)Rosenbrock function for** $D = 12$.
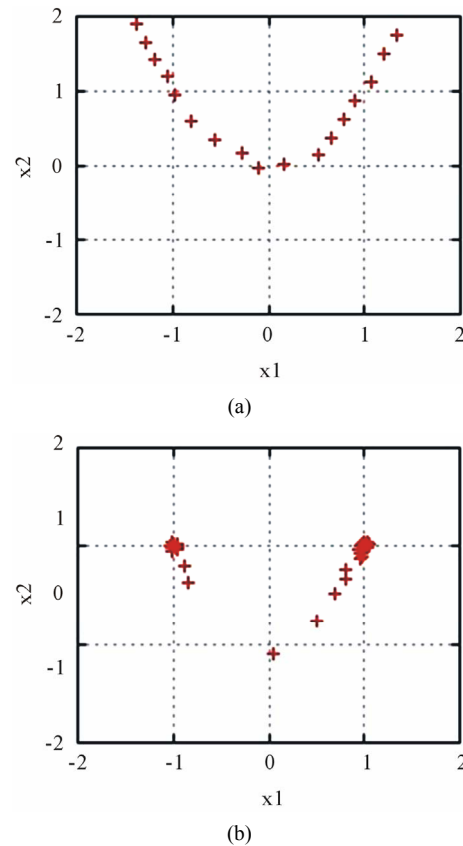

(a)

(b)

Figure 5. Positions of particles at Step $N = 5000$. **(a) Rosenbrock function for** $D = 2$; **(b) Rosenbrock function for** $D = 12$.

swarm is easily trapped in a neighborhood of points giving local minima; on the contrary, if too large, then they cannot keep on swarming and they result in searching the optimal point individually. If σ is suitably chosen, then the swarm strikes out toward the optimal point making global searching without being trapped by local minima. As shown by numerical examples in 12-dimensional space, such machinery is available even for higher dimensional problems.

Our optimization method devised on the basis of animal swarming may be applicable to various optimization problems presented from the real world. In those problems, it may be unknown whether the optimization function is of class C^1 or not, or even if so, its gradient may not be easily computed. For further such applications, we have therefore to combine our method with some methods of numerical differentiations or those of differentiating non smoothing functions numerically using functional values alone.

REFERENCES

[1] Kennedy and R. Eberhart, "Particle Swarm Optimization," *Proceedings of IEEE International Conference*

Neural Networks, Perth, 27 November-1 December 1995, 1942-1948.

[2] D. Bratton and J. Kennedy, "Defining a Standard for Particle Swarm Optimization," *IEEE Swarm Intelligence Symposium*, Honolulu, 1-5 April 2007, pp. 120-127.

[3] H. Liu, A. Abraham and W. Zhang, "A Fuzzy Adaptive Turbulent Particle Swarm Optimization," *International Journal of Computer Applications*, Vol. 1, No. 1, 2007, pp. 39-47. doi:10.1504/IJICA.2007.013400

[4] S. He, Q. Wu, J. Wen, J. Sanuders and R. Paton, "A Particle Swarm Optimizer with Passive Congregation," *Biosystems*, Vol. 78, 2004, pp. 135-147. doi:10.1016/j.biosystems.2004.08.003

[5] M. Eslami, H. Shareef, M. Khajehzadeh and A. Mohamed, "A Survey of the State of the Art in Particle Swarm Optimization," *Research Journal of Applied Science*, *Engineering and Technology*, Vol. 4, 2012, pp. 1181-1197.

[6] J. S. Camazine, J. L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz and E. Bonabeau, "Self-Organization in Biological Systems," Princeton University Press, Princeton, 2001.

[7] C. W. Reynolds, "Flocks, Herds, and Schools: A Distributed Behavioral Model," *Computer Graphics*, Vol. 21, No. 4, 1987, pp. 25-34. doi:10.1145/37402.37406

[8] T. Uchitane, T. V. Ton and A. Yagi, "An Ordinary Differential Equation Model for Fish Schooling," *Scientiae Mathematicae Japonicae*, Vol. 75, 2012, pp. 339-350.

[9] B. Øksendal, "Stochastic Differential Equations," Springer, Berlin, 2007.