IEICE Proceeding Series

PSO with a Pseudo Gradient

Takuya Shindo, Kenya Jin'no

Vol. 2 pp. 205-208
Publication Date: 2014/03/18
Online ISSN: 2188-5079

PSO with a Pseudo Gradient

Takuya Shindo† and Kenya Jin'no†

†Graduate school of Electrical Engineering, Nippon Institute of Technology
4–1 Gakuendai, Miyashiro, Minami-saitama, Saitama, Japan
Email: e2092009@estu.nit.ac.jp, jinno@nit.ac.jp

Abstract—A particle swarm optimization (PSO) is an algorithm that the particle which has the information of its position and velocity searches for the solution in swarm. Especially, based on the numerical simulation results, changing the particle's velocity in the search process influences searching ability of the solution. We pay attention to the velocity of each particle in order to improve the searching ability. Then, particles with good evaluation value are chosen stochastically, and the method to search the best solution by using the position of the particle is proposed.

1. Introduction

Searching for an optimal value of a given evaluation function is very important. In order to solve such optimization problems speedily, various meta-heuristics optimization algorithms have been proposed. Particle swarm optimization (PSO), which was originally proposed by J. Kennedy and R. Eberhart[1][2], is one of such meta-heuristics algorithms. The PSO algorithm is useful tool for optimization problems[3]-[5].

The original PSO is described as

$$\boldsymbol{v}_i^{t+1} = w\boldsymbol{v}_i^t + c_1\boldsymbol{r}_1(\boldsymbol{pbest}_i^t - \boldsymbol{x}_i^t) + c_2\boldsymbol{r}_2(\boldsymbol{gbest}^t - \boldsymbol{x}_i^t) \quad (1)$$

$$\boldsymbol{x}_i^{t+1} = \boldsymbol{x}_i^t + \boldsymbol{v}_i^{t+1} \quad (2)$$

where $w \geq 0$ is an inertia weight coefficient, $c_1 \geq 0$ and $c_2 \geq 0$ are acceleration coefficients, and $\boldsymbol{r}_1 \in [0, 1]^N$ and $\boldsymbol{r}_2 \in [0, 1]^N$ are two separately generated uniformly distributed random number vectors. $\boldsymbol{x}_i^t \in \mathbb{R}^N$ denotes the position vector of the i-th particle on the t-th iteration in the N-dimensional parameter space, and $\boldsymbol{v}_i^t \in \mathbb{R}^N$ denotes the velocity of the i-th particle on the t-th iteration. $\boldsymbol{pbest}_i^t$ represents the position that gives the best value of the evaluation function of the i-th particle until the t-th iteration. $\boldsymbol{gbest}$ is the position that gives the best value of the evaluation function until the t-th iteration in the swarm.

The conventional PSO has quick convergence property. This characteristic serves as hindrance to search the solution in multi-modal function with local optima. Paying attention to convergence of particles, some techniques to improve performance of searching the solution are also proposed. The reason why the convergence speed is quick is that all particles to refer to $\boldsymbol{gbest}$. Then, particles with good evaluation value are chosen stochastically, and the method to search the best solution by using the position of the particle is proposed.

2. PSO using Gradient

We define a pseudo gradient to update the particle position as the following.

$$\boldsymbol{g}_{ik}^t = \frac{f(\boldsymbol{x}_k^t) - f(\boldsymbol{x}_i^t)}{\boldsymbol{x}_k^t - \boldsymbol{x}_i^t} \quad (3)$$

$$\boldsymbol{a}_{ik}^t = \begin{cases} 1 & (\boldsymbol{g}_{ik}^t > 0) \\ 0 & (\boldsymbol{g}_{ik}^t \leq 0) \end{cases} \quad (4)$$

$f(x)$ expresses the evaluation value of each particle. The probability which chooses a certain particle i as an element of a next searching vector is determined by the following equation.

$$\boldsymbol{p}_i^t = \frac{\sum_{k=0}^{K} \boldsymbol{a}_{ik}^t}{(K^2 - K)/2} \quad (5)$$

K expresses the population size of particles. The position of the particles is selected by the probability p is set to $\boldsymbol{x}_p$. $\boldsymbol{x}_p$ is added to $\boldsymbol{gbest}$ of Eq. (1).

$$\begin{aligned}\boldsymbol{v}_i^{t+1} = {} & w\boldsymbol{v}_i^t + c_1\boldsymbol{r}_1(\boldsymbol{pbest}_i^t - \boldsymbol{x}_i^t) + \alpha c_2\boldsymbol{r}_2(\boldsymbol{gbest}^t - \boldsymbol{x}_i^t) \\ & + (1-\alpha)c_2\boldsymbol{r}_3(\boldsymbol{x}_p^t - \boldsymbol{x}_i^t)\end{aligned} \quad (6)$$

The dynamics of the proposed method is described in Eq. (6). The parameter α means the mixture ratio between the global best position $gbest$ and the stochastic calculated x_p position. When the α is 1, the dynamics of the system is equivalent to the conventional PSO.

3. Numerical Simulations

In order to confirm the performance of the proposed method, we compare with the conventional PSO. We carry out 50 times trials. The parameter of PSO are set to $w = 0.729$ and $c_1 = c_2 = 1.494$.

3.1. Benchmark Function

The numerical simulations are carried out by using four standard benchmark functions as shown in TABLE 1. Sphere and Rosenbrock functions are uni-modal functions. Rastrigin and Griewank functions are multi-modal functions. Excepting Rosenbrock function, the optimum value of each function is 0. The optimum value of Rosenbrock function is 0 and the corresponding optimum solution is $x_d = 1$.

Table 1: Benchmark Function

Function		Optimum value
Sphere function	$f_1(\boldsymbol{x}) = \sum_{d=1}^{N} x_d^2$	$f_1(0, 0, 0, \ldots, 0) = 0$
Rosenbrock function	$f_2(\boldsymbol{x}) = \sum_{d=1}^{N-1}(100(x_{d+1} - x_d^2)^2 + (x_d - 1)^2)$	$f_3(1, 1, 1, \ldots, 1) = 0$
Rastrigin function	$f_3(\boldsymbol{x}) = 10N + \sum_{d=1}^{N}((x_d)^2 - 10\cos(2\pi x_d))$	$f_3(0, 0, 0, \ldots, 0) = 0$
Griewank function	$f_4(\boldsymbol{x}) = 1 + \frac{1}{4000}\sum_{d=1}^{N} x_d^2 - \prod_{d=1}^{N}\cos\left(\frac{x_d}{\sqrt{d}}\right)$	$f_4(0, 0, 0, \ldots, 0) = 0$

Table 2: Function Parameter for the Benchmarks

Function	Search Range	$v_{\max}$	Initial Range
$f_1(x)$	$(-100,\ 100)$	100	$(50,\ 100)^N$
$f_2(x)$	$(-100,\ 100)$	100	$(50,\ 100)^N$
$f_3(x)$	$(-10,\ 10)$	10	$(2.56,\ 5.12)^N$
$f_4(x)$	$(-600,\ 600)$	600	$(300,\ 600)^N$

The parameters; searching range, initializing range, and $v_{\max}$ are set as shown in TABLE 2 for each benchmark function. $v_{\max}$ is a divergent control parameter. The upper bound is given at each particle velocity which is calculated in Eqs. (1) and (6). Initializing range is determined as an asymmetric range in the searching range. This operation provides the biased initial values.

3.2. Influence of Parameter α

We confirm an influence to solution search performance of the mixture parameter α. Figures 1 and 2 show the characteristic of the search performance. The horizontal axis denotes the parameter α, and the vertical axis denotes the average evaluation value. These results indicate that the parameter α improves the performance around $\alpha = 0.5$.

3.3. Results

The numerical simulation results are shown in Fig. 3. The horizontal axis denotes iteration and the vertical axis denotes average evaluation value.

In the cases of Rastrigin function and Griewank function, the solution search performance is improved. Namely, the proposed procedure improves the search performance of multi-modal functions. On the other hand, the solution search performance dose not improve in Sphere function and Rosenbrock function. Moreover, also in which benchmark functions, the convergence is slow as compared with the conventional PSO.

Figures 4 and 5 show the examples of the decay of the particle velocity. In these cases, we apply Rastrigin function as the benchmark function. These figures indicate that the proposed procedure keeps the particle velocity comparing with the conventional PSO procedure. From these results, when each particle refers to each target position, the convergence can be delayed.

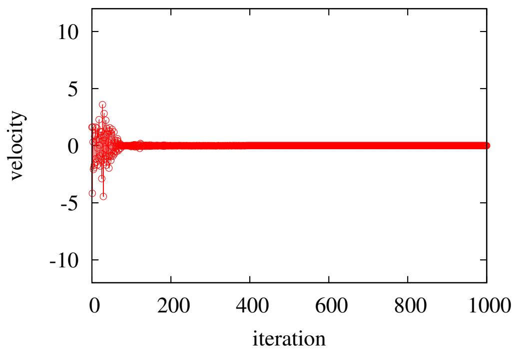


Figure 4: Velocity of a particle in a certain dimension for the conventional PSO (Rastrigin function)

4. Conclusions

In order to prevent too early converging property of the particle of the PSO, we proposed the novel procedure which refers the pseudo gradient of the evaluation value. We confirmed the performance is depended on the parameter α by numerical simulations. Also, we confirmed the proposed procedure keeps the particle velocity comparing with the conventional PSO procedure. Based on these properties, the proposed procedure exhibits an excellent performance for multi-modal function. However the system dose not search local. To overcome these points is our future problems.

References

[1] James Kennedy and Rusell Eberhart, “Particle Swarm Optimization”, in *Proc. IEEE Int. Conf. Neural Net-*

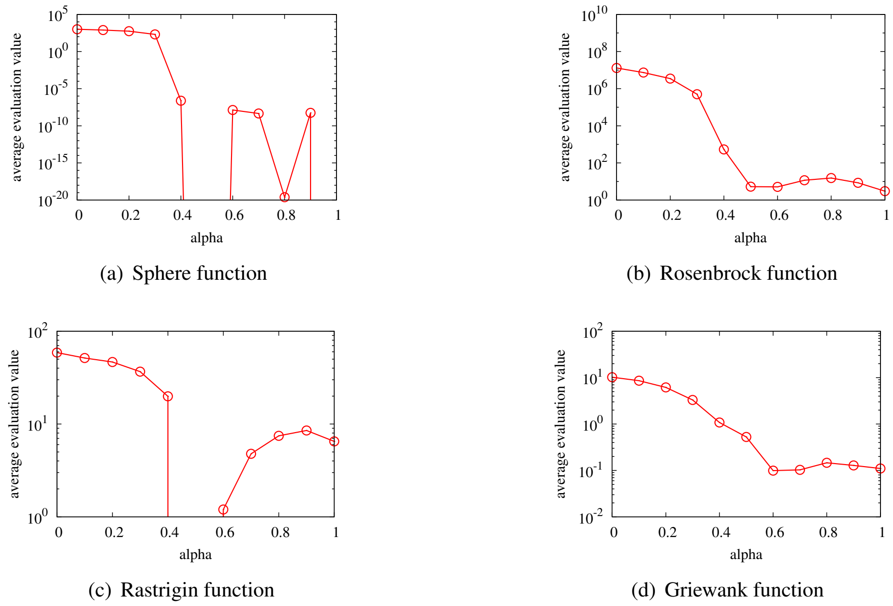


Figure 1: Influence of parameter α to solution search performance

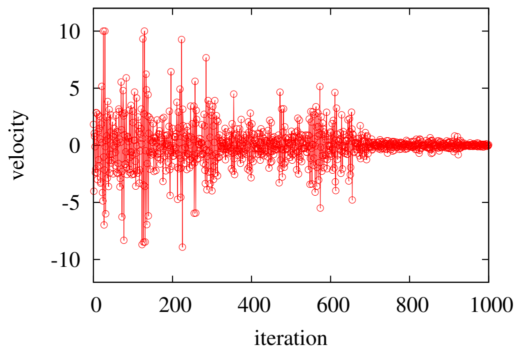


Figure 5: Velocity of a particle in proposed method (Rastrigin function)

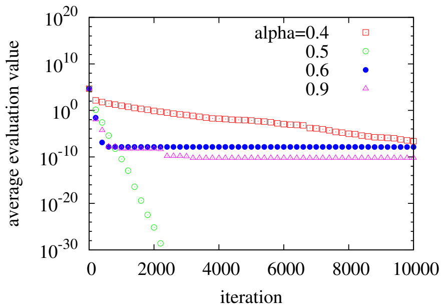
works, pp. 1942-1948, 1995.

[2] James Kennedy, "The particle swarm: Social adaptation of knowledge," in *Proc. IEEE Int. Conf. Evolutionary Computation,* pp. 303-308, 1997.

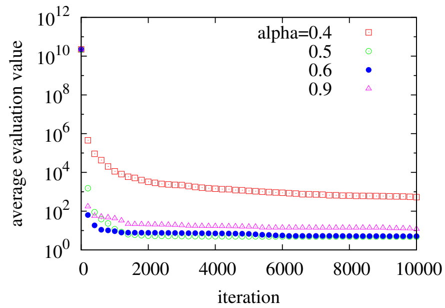
[3] Y. Shi and R. Eberhart, "Empirical study of particle swarm optimization," in *Proc. ICEC 1999*, pp. 1945-1950, 1999.

[4] M. Clerc and James Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Comput.,* vol. 6, no. 1, pp. 58-73, 2002.

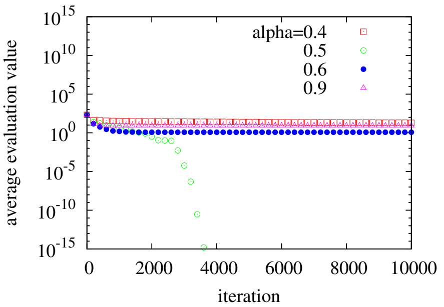
[5] F. van den Bergh and A. P. Engelbrecht, "A cooperative approach to particle swarm optimization," *IEEE Trans. Evol. Comput.,* vol. 8, no. 3, pp. 225-239, 2004.



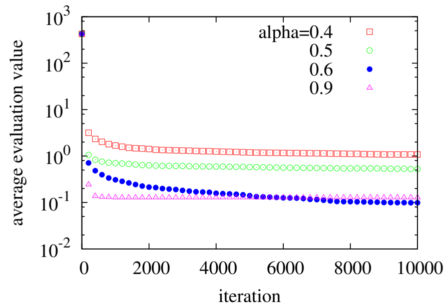
(a) Sphere



(b) Rosenbrock

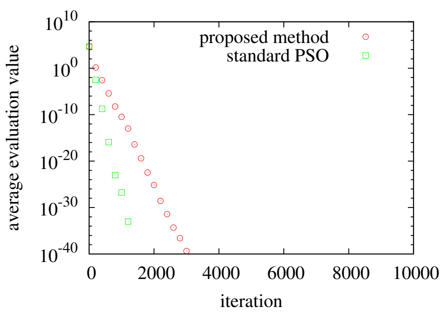


(c) Rastrigin

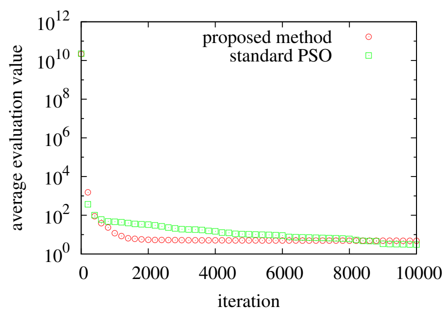


(d) Griewank

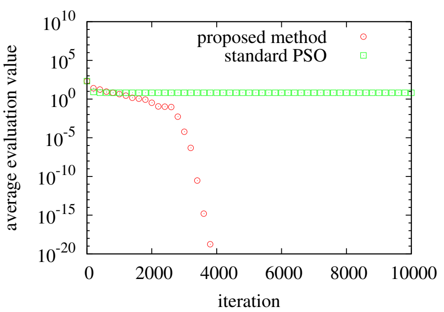
Figure 2: Influence of parameter α



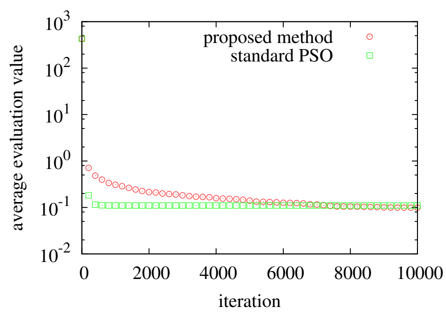
(a) Sphere (α = 0.5)



(b) Rosenbrock (α = 0.5)



(c) Rastrigin (α = 0.5)



(d) Griewank (α = 0.6)

Figure 3: Simulation results