A Novel Intelligent Particle Optimizer for Global Optimization of Multimodal Functions

Zhen JI, Member, IEEE, Huilian LIAO, Yiwei WANG and Q. H. WU, Senior Member, IEEE

Abstract—A novel intelligent particle optimizer based on subvectors (IPO) is proposed in this paper, which is inspired by conventional particle swarm optimization (PSO). IPO uses only one particle instead of a particle swarm. The position vector of this particle is partitioned into a certain number of subvectors, and the updating process is based on subvectors and evolved to subvectors updating process, in which the particle adjusts the velocity intelligently by introducing a new learning factor. This learning factor utilizes the information contained in the previous updating process. The particle is capable of increasing its velocity towards the global optimum in lower dimensional subspaces and not being trapped in local optima. Experimental results have demonstrated that IPO has impressive ability to find global optimum. IPO performs better than recently developed PSO-based algorithms in solving some complicated multimodal functions.

I. INTRODUCTION

Optimization has been a hotspot of research for several decades. Particle swarm optimization (PSO), which has been used to successfully optimize a wide range of problems, was proposed by Kennedy and Eberhart[1] in 1995 based on an analogy with models of the social behavior of groups of simple individuals. Later, Shi et al.[2] improved PSO algorithm by introducing an inertia weight \( w \) into the velocity updating equation of the original PSO algorithm (PSO-\( w \)). However, PSO usually suffers from premature convergence when strongly multimodal problems are being optimized. To overcome the drawback of premature convergence, various improved algorithms based on conventional PSO were developed mainly by increasing the diversity of solutions in the swarm[3][4][5][6]. In 2003, Peram et al.[7] proposed Fitness-distance-ratio based particle swarm optimization (FDR-PSO), which moves particles towards nearby particles of higher fitness, instead of attracting each particle towards just the best position discovered so far by any particle. It is accomplished by using the ratio of the relative fitness and the distance of other particles to determine the direction in which each group of the particle position needs to be changed. In 2006, Liang et al.[8] proposed a comprehensive learning particle swarm optimizer (CLPSO), which uses a novel learning strategy whereby all other particles’ historical best information is used to update a particle’s velocity.

In this paper, a novel intelligent particle optimizer based on subvectors (IPO) is proposed with the inspiration of conventional PSO. This algorithm uses only one particle instead of a particle swarm, and the complete position vector is partitioned into a certain number of subvectors, which are updated in sequence repeatedly. IPO uses a novel way to update the particle’s velocity and position. During the process of updating each subvector, the algorithm adjusts the velocity intelligently by introducing a learning factor, which utilizes the information in the updating process. For instance, the velocity of the particle will be increased if the fitness value is improved; the velocity will be slowed down when the particle skips over the optimum; when the fitness value is not improved after several iterations in the subvector updating process, the particle will increase the diversity of velocity in order to escape from the local optimum. Experiments were performed on complicated multimodal functions, such as the Rosenbrock’s, Rastrigin’s and Ackley’s function. The experimental results have demonstrated that the new optimizer performs much better in solving some complicated multimodal problems in comparison with PSO-\( w \), FDR-PSO and CLPSO algorithms.

II. PARTICLE SWARM OPTIMIZATION

Particle swarm optimization is a population-based optimization algorithm. It is popular in the research field of optimization in the last decade for its simplicity of implementation, few parameters and high convergence rate. PSO algorithm works on the social behavior of particles in the swarm by remembering the best location of itself and the best experience of other individuals in the swarm. The particles alter their velocities according to their records at each iteration. The position vector and velocity vector of the particle \( i \) (\( i=1,2,\ldots,n \)) in the \( D \)-dimensional search space are denoted by \( P_i=(z_{id}^{1},z_{id}^{2},\ldots,z_{id}^{D}) \) and \( \dot{V}_i=(v_{id}^{1},v_{id}^{2},\ldots,v_{id}^{D}) \) respectively. The best position of the particle \( i \) and the fittest particle found so far in the swarm are represented by \( P_i=(p_{id}^{1},p_{id}^{2},\ldots,p_{id}^{D}) \) and \( P_f=(p_{f1},p_{f2},\ldots,p_{fD}) \) respectively. In PSO, the particle’s velocity and position are updated according to the following equations at each iteration:

\[
v_{id}^{k+1} = w v_{id}^{k} + c_1 r_1 (p_{id} - z_{id}^{k}) + c_2 r_2 (p_f - z_{id}^{k}) \quad (1)
\]

\[
z_{id}^{k+1} = z_{id}^{k} + v_{id}^{k+1} \quad (2)
\]

where \( k \) is the iteration number; \( w \) is the inertia weight; \( r_1 \) and \( r_2 \) are random numbers in the range [0.1]. Constants \( c_1 \) and \( c_2 \), called acceleration factors, make the particle adjust its trajectory according to its own previous best position and excellent individuals in the group, and to some extent move...
III. INTELLIGENT PARTICLE OPTIMIZER BASED ON SUBVECTORS

In the updating process of IPO, velocity and position vector are updated based on subvectors, which are obtained through splitting the solution space into lower dimensional subspaces, instead of the complete position vector. In the subvector updating process, to adjust the velocity subvector dynamically, learning factor is introduced to utilize the information in the updating process, and determine the new velocity.

A. Subvectors

For most stochastic optimization algorithms, including PSO, their performance deteriorates as the dimensionality of the search space increases. At each iteration of the conventional PSO algorithm, it updates all dimensions of the complete position vector at a time. The fitness value obtained represents the quality of the complete position vector, however, it can not tell whether some dimensions move towards global optima or not. For instance, in solving a 3-dimensional function whose global optimum is [0.0,0], the initial position vector is set to [1,1,1]. A perturbation [0.2,-0.5,0.3] is added into the initial position vector, therefore the updated position vector is [1.2,0.5,1.3]. At this moment, although the second dimension moves towards the global optimum, other dimensions move away from the global optima. The conventional PSO algorithm does not take some dimensions’ optimum directions into consideration. To overcome the drawback, the search space is partitioned into lower dimensional subspaces as reference[6], as long as the algorithm can guarantee that it is capable of searching every possible region of the search space.

![Fig. 1. The illustration of m subvectors.](image)

In the proposed IPO algorithm, the D-dimensional search space is partitioned into m parts. The particle represents a complete position vector, and is the potential solution of the problem. The position vector is split into m subvectors. Each position subvector and the corresponding velocity subvector of the jth part are denoted by \( \tilde{Z}_j \) and \( \tilde{V}_j \) (\( j = 1, \ldots, m \)) respectively. For simplicity, D dimensions can be divided exactly into m subvectors, therefore each subvector contains \( l \) (\( l = D/m \)) dimensions as shown in Figure 1. Arbitrary partition may make the new algorithm become trapped in pseudo-minima created by the partitioning process. Therefore, it is hoped that some strongly correlated dimensions will end up in the same subvector (refer to reference[6] for more details).

B. Subvectors updating process

The updating process in IPO is based on subvectors, which are updated iteratively, instead of the complete position vector. Therefore, the updating process is evolved to subvectors updating process. The main difference between the optimizer in reference[6] and IPO is that IPO uses a novel method to update subvectors instead of the updating rules of conventional particle swarm optimization. In the updating process of subvector \( j \), the position and velocity subvectors are updated \( N \) times iteratively according to the following equations:

\[
\tilde{V}_j^{k+1} = (a/k^s) \times r + b \times L_j^{k+1} \tag{3}
\]

\[
\tilde{Z}_j^{k+1} = \begin{cases} 
\tilde{Z}_j^k + \tilde{V}_j^k & \text{if } f(X_1^k) \leq f(X_2^k) \\
\tilde{Z}_j^k & \text{otherwise}
\end{cases} \tag{4}
\]

\[
L_j^{k+1} = \begin{cases} 
L_j^{k+1} & \text{if } f(X_1^k) > f(X_2^k) \\
0 & \text{if } f(X_1^k) \leq f(X_2^k)
\end{cases} \tag{5}
\]

where \( X_1^k = [Z_1, Z_2, \ldots, Z_j^{k-1}, Z_j^k, \ldots, Z_m] \) and \( X_2^k = [Z_1, Z_2, \ldots, Z_j^{k-1}, Z_j^k, \ldots, Z_m] \), \( k = 1, \ldots, N \). The vector \( L \), called learning factor, is used to help update the velocity of the jth subvector, and it is set to zero if less than \( \varepsilon \), which is a very small value. The constant \( s \) is a shrink factor. The vector \( r \) consists of random values in the range \([-0.5,0.5]\), which increases the diversity of the velocity. The parameters \( a \) and \( b (b \geq 1) \) are acceleration constants. The calculation of fitness function \( f() \) still requires a complete D-dimensional vector. The fitness value is updated after each iteration in the subvector updating process.

In the subvector updating process, the new position subvector will be determined by the velocity subvector, which consists of two parts: the diversity part \( (a/k^s) \times r \) and the learning part \( b \times L_j^{k+1} \). The term \( (a/k^s) \), as a decreasing function of time, makes the particle possess more exploitation ability at the beginning, then have more exploration ability to fine search the area around. The particle adjusts the velocity subvector dynamically according to the current updating status through the learning part, which can be seen from the following:

- At each iteration of updating subvector, the velocity subvector will be multiplied by \( b \) times if a position subvector with better fitness value is found, according to Eqs.(3) and (5).
- If the fitness value is improved in the first iteration, and declines in the second iteration, the reason may be that the position subvector skips over the optimum, since the velocity subvector of the second iteration is \( b \) times of that in the first iteration. In this case, the velocity needs to be slowed down. Therefore, the velocity subvector of
the first iteration will be divided by \((s/b)\) as the learning part in the third iteration.

- If the fitness value is not improved continuously by several iterations during the subvector updating process, the learning factor \(L\) will be decreased to a very small number due to the shrink factor, and it will be set to zero if less than \(\varepsilon\). At the moment, only the term \((a/k')\times\pi\) determines the velocity subvector until the fitness value is improved again. In this case, the diversity of the velocity subvector is increased and the particle could jump out of local optima easily.

What mentioned above is the reason why the particle is called intelligent particle. The pseudocode of IPO is given as follows. Here, only one particle is applied in this algorithm.

\[
\text{begin} \\
\text{initialize } Z^0 = (Z_1, Z_2, \ldots, Z_m) \text{ and calculate } f(Z^0) \\
\text{for } (k = 1 \text{ to iteration number}) \\
\quad \text{for } (j = 1 \text{ to } m) \\
\quad \quad L_j^k = 0 \\
\quad \quad \text{for } (k' = 1 \text{ to } N) \\
\quad \quad \quad \text{evaluate } V_j^{k'} \text{ according to Eq.3} \\
\quad \quad \quad \text{evaluate } Z_j^{k'} \text{ according to Eq.4} \\
\quad \quad \quad \text{evaluate } L_j^{k'} \text{ according to Eq.5} \\
\quad \quad \text{end for} \\
\quad \text{end for} \\
\text{end for} \\
\text{end for} \\
\text{end}
\]

IV. EXPERIMENTS AND PERFORMANCE ASSESSMENT

Comparison was performed through three complicated multimodal functions: Rosenbrock, Rastrigin and Ackley. Rosenbrock’s function can be treated as a multimodal problem, and it is very difficult to minimize for it has a narrow valley from the perceived optima to the global optimum. Rastrigin’s function is a complicated multimodal problem with a large number of local optima, and most algorithms are easily trapped into local optima in solving this kind of problem. They are defined as follows:

- Rosenbrock: \(f_1(x) = \sum_{d=1}^{D-1} (100(x_d^2 - x_{d+1}) + (x_d - 1)^2)\);
- Rastrigin: \(f_2(x) = \sum_{d=1}^{D} (x_d^2 - 10 \cos(2\pi x_d) + 10)\);
- Ackley: \(f_3(x) = -20 \exp \left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^{D} x_i^2}\right) - \exp \left(\frac{1}{D} \sum_{i=1}^{D} \cos(2\pi x_i)\right) + 20 + e\);

where \(D=30\). \(f_1(x), f_2(x)\) and \(f_3(x)\) are treated as fitness functions respectively. In Table I, the global optimum \(x^*\), the corresponding fitness value \(f(x^*)\), the search range, and the initialization range for IPO and other algorithms, similar to reference [8]. The performance of IPO was compared with algorithms PSO-\(w\), FDR-PSO and CLPSO. To the authors’ knowledge, FDR-PSO is the best algorithm in solving Rosenbrock’s function among existing algorithms, and CLPSO is a good algorithm in solving multimodal functions which was proposed recently. The parameter settings for PSO-\(w\) are \(w = 0.8\) and \(c_1 = c_2 = 1.8\). The parameter settings for IPO are \(b = 2; s = 4; \varepsilon = 10^{-15}; a = 30\) for Rosenbrock while \(a = 3\) for other two functions. IPO employs only one particle, which is partitioned into 15 subvectors for Rosenbrock, while 30 for other two functions.

The experimental results of FDR-PSO and CLPSO are from reference [8]. PSO-\(w\) and IPO were performed in Matlab 7.0 for 20 trials. The averages and standard deviations, to assess the robustness of the algorithm, were calculated as listed in Table II. IPO achieves significant improvements respectively in solving Rosenbrock’s function in comparison with PSO-\(w\), FDR-PSO and CLPSO. As for the Rastrigin’s function, it also performs better than other algorithms. IPO performs better than PSO-\(w\) on Ackley’s function, however, not so well as FDR-PSO and CLPSO.

The convergence characteristics of IPO and PSO-\(w\) are shown in Figure 2, where the fitness values are shown in log scale. The smaller the fitness value is, the better the solution is. The IPO is capable of improving the fitness value continuously throughout the simulation, especially at the early stage of the optimization process. For some iterations, the fitness value is decreased dramatically, since the particle escapes from local optima. In contrast, the performance of PSO-\(w\) was found to be significantly poor, especially in solving the Rastrigin’s and Ackley’s function.

To verify the search ability of the IPO algorithm, the search range and initialization range of IPO and PSO-\(w\) are extended to a wider range \([-100, 100]\) because of a great number of traps existing beyond the ranges given by Table I. The results are listed in Table III. Although the initialization and search ranges of IPO become larger, the results obtained by IPO on function Rosenbrock and Rastrigin are similar with those in Table II. However, PSO-\(w\) performs much poorer in comparison with the results obtained when smaller initialization and search range are used. Both of PSO-\(w\) and IPO perform poor on function Ackley when then initialization and search ranges are expanded.

V. CONCLUSION

Many existing optimization algorithms are easily trapped in local optima in solving multimodal functions. Experimental results have demonstrated that the proposed IPO optimizer achieves a significant improvement in comparison with the conventional PSO-\(w\), FDR-PSO and CLPSO algorithms when optimizing some complicated multimodal functions, which have a narrow valley from the perceived optima to the global optimum, or have a large number of local optima.

ACKNOWLEDGMENT

This work was partially supported by the National Natural Science Foundation of China under grant no. 60572100, Royal Society (U.K.) International Joint Projects 2006/R3 - Cost Share with NSFC, Foundation of State Key Laboratory
TABLE I
THE GLOBAL OPTIMUM $x^*$, THE CORRESPONDING FITNESS VALUE $f(x^*)$, THE SEARCH RANGE, AND THE INITIALIZATION RANGE FOR ALGORITHMS PSO-w, FDR-PSO, CLPSO AND IPO

<table>
<thead>
<tr>
<th>Functions</th>
<th>$x^*$</th>
<th>$f(x^*)$</th>
<th>Initialization Range</th>
<th>Search Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosenbrock</td>
<td>[1, 1, ... , 1]</td>
<td>0</td>
<td>[-2.048, 2.048]</td>
<td>[-2.048, 2.048]</td>
</tr>
<tr>
<td>Rastrigin</td>
<td>[0, 0, ... , 0]</td>
<td>0</td>
<td>[-5.12, 5.12]</td>
<td>[-5.12, 5.12]</td>
</tr>
<tr>
<td>Ackley</td>
<td>[0, 0, ... , 0]</td>
<td>0</td>
<td>[-32, 32]</td>
<td>[-32, 32]</td>
</tr>
</tbody>
</table>

TABLE II
PERFORMANCE COMPARISON OF FITNESS VALUE $f(x^*)$ FOR MULTIMODAL FUNCTIONS ($D=30$)

<table>
<thead>
<tr>
<th>Functions</th>
<th>PSO-w</th>
<th>FDR-PSO</th>
<th>CLPSO</th>
<th>IPO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosenbrock</td>
<td>2.32e2 ±1.95e2</td>
<td>5.39 ±1.76</td>
<td>21±2.98</td>
<td>0.3±1.00</td>
</tr>
<tr>
<td>Rastrigin</td>
<td>1.56e2 ±3.31e2</td>
<td>28.4 ±8.71</td>
<td>4.85e-10 ±3.63e-10</td>
<td>1.8e-10 ±1.75e-10</td>
</tr>
<tr>
<td>Ackley</td>
<td>1.37e1 ±3.56</td>
<td>2.84e-14 ±4.10e-15</td>
<td>0 ±0</td>
<td>7.67e-7 ±3.80e-7</td>
</tr>
</tbody>
</table>

TABLE III
PERFORMANCE COMPARISON OF FITNESS VALUE $f(x^*)$ FOR MULTIMODAL FUNCTIONS ($D=30$) WITH [-100,100] AS THE INITIALIZATION AND SEARCH RANGES

<table>
<thead>
<tr>
<th>Functions</th>
<th>PSO-w</th>
<th>IPO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosenbrock</td>
<td>3.01e8 ±3.42e8</td>
<td>0.04±0.04</td>
</tr>
<tr>
<td>Rastrigin</td>
<td>8.22e3 ±5.52e3</td>
<td>9.36e-11 ±6.65e-11</td>
</tr>
<tr>
<td>Ackley</td>
<td>20 ±0</td>
<td>20 ±0</td>
</tr>
</tbody>
</table>

Fig. 2. The convergence characteristics of 30-dimensional problems for PSO-w and IPO.

REFERENCES


