Research Article

An Orthogonal Learning Differential Evolution Algorithm for Remote Sensing Image Registration

Wenping Ma, Xiafei Fan, Yue Wu, and Licheng Jiao

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi 710071, China

Correspondence should be addressed to Wenping Ma; wpma@mail.xidian.edu.cn

We introduce an area-based method for remote sensing image registration. We use orthogonal learning differential evolution algorithm to optimize the similarity metric between the reference image and the target image. Many local and global methods have been used to achieve the optimal similarity metric in the last few years. Because remote sensing images are usually influenced by large distortions and high noise, local methods will fail in some cases. For this reason, global methods are often required. The orthogonal learning (OL) strategy is efficient when searching in complex problem spaces. In addition, it can discover more useful information via orthogonal experimental design (OED). Differential evolution (DE) is a heuristic algorithm. It has shown to be efficient in solving the remote sensing image registration problem. So orthogonal learning differential evolution algorithm (OLDE) is efficient for many optimization problems. The OLDE method uses the OL strategy to guide the DE algorithm to discover more useful information. Experiments show that the OLDE method is more robust and efficient for registering remote sensing images.

1. Introduction

Image registration is an important step for many fields [1], such as change detection, image fusion, and object recognition. In order to provide complete information about the image, it is necessary to register the images taken from different sensors or from the same sensor at different times. The result of image registration will greatly influence the performance of the follow-up procedure. So remote sensing image registration methods should be efficient, robust, and accurate.

Image registration methods are usually divided into two categories: feature-based and intensity-based methods [2, 3]. Many feature-based methods have been proposed [4, 5]. These methods usually need to initially extract salient features, such as point, edge, contour, and region. Those features are matched using similarity measures to establish the geometric correspondence between two images. One of the main advantages of these approaches is that they are efficient and robust to noise, complex geometric distortions, and significant radiometric differences. However, they will only perform well on the condition that suitable features

are extracted and reliable algorithms are used [3]. For some images, where features are not obvious, intensity-based methods perform better than feature-based methods.

The key procedure of intensity-based method is to find the optimal similarity metric. The similarity metric is to measure how closely the gray values of two images are matched. The similarity metric for remote sensing image registration must be robust. There are many commonly used similarity metrics [2, 6–11]. Mutual information which is based on the Shannon definition of entropy [9, 10] was widely used in a lot of current work. Mutual information has been shown to be robust and does not depend on the intensity scaling or specific dynamic range of the images [12].

The searching strategy optimizes the similarity metric. Both local and global search strategies are commonly used. Many local methods have been used in image registration [13, 14]. These local methods yield the best registration when the initial orientation is very close to the true transformation. Besides, they are still easy to be trapped in a local optimum [14, 15]. So the global optimization is often required and global methods have been successfully applied to image registration [8, 12, 16–22].

The differential evolution (DE) algorithm was proposed by Storn and Price for global optimization over continuous search space [23]. DE is a version of evolutionary algorithm (EA) that has proven to be fast and reliable in many applications [23–29]. The DE algorithm has shown to be efficient in remote sensing image registration [30]. The seminal idea of DE is to generate a new vector by adding the weighted difference between two trial vectors to a third vector. The new vector is defined as $V = X_1 + F(X_2 - X_3)$, where X_1, X_2, and X_3 are three randomly selected trial vectors from the population and F is a multiplier, which is the main parameter of the DE algorithm [25].

The orthogonal experimental design (OED) offers an ability to discover the best combination levels for different factors with a reasonably small number of experimental samples [31]. Owing to the OEDs orthogonal prediction ability and test ability, the orthogonal learning (OL) strategy can construct a guidance exemplar with the ability to predict promising search directions toward the global optimum [31]. The OL strategy has been successfully applied to many areas [32–38]. In this paper, we will apply the OL strategy to image registration problem. Considering the effectiveness of the DE algorithm in image registration [30], we combine the OL strategy with the DE. The main idea of our method is based on the observation that the major step in the DE can be considered to be an "experiment." Based on this "experiment," the OL strategy will construct a guidance exemplar with an ability to predict promising search directions toward the global optimum. In our method, the OED is used to discover the best combination of three trial vectors.

The rest of this paper is organized as follows. In Section 2, image transformation, similarity metrics, and optimization techniques are discussed. In Section 3, the orthogonal learning strategy is formulated. In Section 4, the proposed OLDE for image registration is presented. Experimental results are described in Section 5. Finally, conclusions are drawn in Section 6.

2. Image Registration

2.1. Image Transformation. The image registration process is actually a process to seek the one-to-one mapping between two images. The process links the points in two images corresponding to the same spatial position. The mapping is commonly referred to as a transformation. It is a two-dimensional transformation in a two-dimensional space. The proposed approach in this paper is to be used in image registration in two-dimensional space. We are using a widely applied affine transformation model to transform the target image. This will allow us to demonstrate the efficacy of the OLDE method in image registration.

2.2. Similarity Metrics. At the correct registration, similarity metrics must be robust. They should attain a global or a very distinct local maximum. Most of the current work on remote sensing image registration utilizes mutual information which has been shown to be robust for remote sensing image registration. Mutual information represents the relative entropy of

two images [6]. The larger the value of mutual information, the better the registration of the two images. In general, given two images U and V, their mutual information is

$$I(U, V) = H(U) + H(V) - H(U, V), \qquad (1)$$

where $H(U, V)$ is the joint entropy and the $H(U)$ and $H(V)$ are the entropies of U and V. Respectively, $H(U)$, $H(V)$, and $H(U, V)$ are given as

$$H(U) = -\sum_u P_U(u) \log P_U(u),$$

$$H(V) = -\sum_v P_V(v) \log P_V(v), \qquad (2)$$

$$H(U, V) = -\sum_{u,v} P_{U,V}(u, v) \log P_{U,V}(u, v),$$

where $P_U(u)$ is the marginal probability density function of U, $P_V(v)$ is the marginal probability density function of V, and $P_{U,V}(u, v)$ is the joint probability density function of U and V. The $P_U(u)$, $P_V(v)$, and $P_{U,V}(u, v)$ can be estimated with the Parzen windows [10]. Normalized mutual information I_N is given as

$$I_N(U, V) = \frac{H(U) + H(V)}{H(U, V)}. \qquad (3)$$

Normalized mutual information is less sensitive to the size of overlap area. It has been proved that normalized mutual information is an accurate and robust image registration similarity metric in previous studies [6, 8, 39]. Therefore, in current study, normalized mutual information was selected as the similarity measure. In image registration, the greater the value of I_N, the better the match between the two images. So this is a maximization problem. Many optimization problems are formulated as minimization problems, where the objective function is denoted as $f(x)$, and, thus, it is understood that for image registration, the goal is to minimize $-f(x)$, without loss of generality.

2.3. Optimization of Similarity Metrics. The goal of image registration is to find the best transformation parameters to maximize the object function (similarity metric). Both image interpolation and joint probability density estimation are involved in image registration. Thus, computing input parameters for (3) is computationally expensive. Therefore, it is necessary to get an effective method to reduce this cost. The OLDE algorithm, which this paper proposes, is an efficient global method for image registration.

3. Motivation

3.1. Motivation of Orthogonal Learning Strategy. The OL strategy can discover more useful information toward the global optimum [31]. Here is a simple case to illustrate the importance of the OL strategy. Given a 3-dimension sphere function, $f(x) = x_1^2 + x_2^2 + x_3^2$, whose global optimum value

TABLE 1: Factors and levels of the chemical experiment example.

| Level | Factors | | |
	A Temperature (°C)	B Time (min)	C Alkali (%)
L1	80	90	5
L2	85	120	6
L3	90	150	7

TABLE 2: Orthogonal array of $L_9(3^3)$.

Combination	Factor 1	Factor 2	Factor 3
1st	1	1	1
2nd	1	2	2
3rd	1	3	3
4th	2	1	2
5th	2	2	3
6th	2	3	1
7th	3	1	3
8th	3	2	1
9th	3	3	2

TABLE 3: Orthogonal array $L_9(3^3)$ of the chemical experiment.

| Combination | Factors | | |
	A: temperature (°C)	B: time (Min)	C: alkali (%)
1st	80	90	5
2nd	80	120	6
3rd	80	150	7
4th	85	90	6
5th	85	120	7
6th	85	150	5
7th	90	90	7
8th	90	120	5
9th	90	150	6

is 0 and the minimum point is $[0, 0, 0]$. Suppose that the current trial vectors are $X_1 = [3, 0, 1]$, $X_2 = [2, 1, 4]$, and $X_3 = [0, 1, 2]$. Furthermore, assume that the value of F is 0.5, and then the new vector is $V = X_1 + F(X_2 - X_3) = [4, 0, 2]$. This results in the new vector with a cost value of 20, which is worse than both X_1 and X_3. However, if we can discover good dimensions of the three vectors, we can then combine them to form three new trial vectors $X_{1new} = [0, 0, 2]$, $X_{2new} = [0, 1, 2]$, and $X_{3new} = [0, 1, 1]$. Then, the updated vector is $V = [0, 0, 1.5]$, resulting in the updated vector with a cost value of 2.25. Thus, the object function is moving toward the global optimum value of 0.

The simple case above has illustrated the importance of designing three new trial vectors to generate the updated vector V. In order to find the best combination of trial vectors, we use orthogonal experimental design (OED) [40], which can get a relatively good vector through only a few experimental tests [31].

3.2. Orthogonal Experimental Design. We use a simple example to explain OED. In this example, to yield maximum chemical product, we should find the best level combination of the three factors: temperature, time, and alkali. These three factors which will affect experimental results are shown in Table 1 and we denoted them as factors A, B, and C. The temperature has three levels: 80°C, 85°C, or 90°C. The time can be 90 min, 120 min, and 150 min. And the alkali can be 5%, 6%, and 7%. Therefore, there are 27 ($3^3 = 27$) combinations of experimental designs totally. It is desirable to obtain or predict the best combination by sampling only a few representative experimental cases.

Let $L_k(n^m)$ denote an orthogonal array, where m is the number of factors and each factor has n levels, k is the number of the combinations of levels, and "L" denotes a Latin square. Table 2 shows an orthogonal array $L_9(3^3)$. Each row in this table shows one combination of levels. The orthogonal array has three properties. First, for the factor in any column, every level occurs in equal times. Secondly, for the two factors in any two columns, every combination of two levels occurs in equal times. Thirdly, the selected combinations are uniformly distributed over the whole space of all the possible combinations [32]. We apply $L_9(3^3)$ to the example of chemical experiment. An orthogonal array with three factors is shown in Table 3.

In this paper, we use OL strategy to guide the DE algorithm to select promising search directions towards the global optimum, which will enable us to achieve the best

image registration results in terms of the normalized mutual information similarity metric.

4. OLDE Algorithm for Image Registration

4.1. Encoding. We make use of the aforementioned affine transformation model:

$$T = \begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ 0 & 0 & 1 \end{bmatrix},\qquad(4)$$

where a_{11}, a_{12}, a_{21}, a_{22}, b_1, and b_2 are six transformation parameters. Then, the transformation formulas of image are represented as

$$x' = a_{11}x + a_{12}y + b_1,$$
$$y' = a_{21}x + a_{22}y + b_2,\qquad(5)$$

where (x, y) is the coordinate of the target image and (x', y') is the coordinate of the transformed target image. Given 6 real parameters, a_{11}, a_{12}, a_{21}, a_{22}, b_1, and b_2, trial vectors are formed based on those values. Therefore, each trial vector in the initial population of n trial vectors is an array with 6 positions, with the parameter vector denoted by $T = (a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2)$. The initial population is randomly initialized so that each parameter can uniformly vary within a range of its own.

4.2. Population Evolution. To create a new population, both orthogonal crossover and the DE algorithm are executed consecutively. We refer to this combination as the OLDE algorithm.

4.2.1. Orthogonal Crossover Based on Orthogonal Array. Here, we introduce the process of n-to-g orthogonal crossover. For example, we have a population with N vectors. We select n of them to be the parent vectors and each parent vector with l real values. The details of the orthogonal crossover that these n parents are based on $L_k(n^m)$ to produce g vectors are as follows.

n-to-g Orthogonal Crossover

Step 1. Input n parent vectors $X_i = (x_{i1}, x_{i2}, \ldots, x_{il})$, where $1 \le i \le n$ and X_i has l real values.

Step 2. For all $1 \le i \le l$, randomly and independently generate $t(i) \in \{1, 2, \ldots, m\}$.

Step 3. For all $1 \le i \le k$, produce the vectors $p_i = (x_{c_{t(1)}(i),1}, x_{c_{t(2)}(i),2}, \ldots, x_{c_{t(l)}(i),l})$ based on the ith combination of factor levels $(c_1(i), c_2(i), \ldots, c_m(i))$ in the orthogonal array $L_k(n^m)$.

Step 4. Evaluate the fitness of vectors $p_1, p_2, \ldots, p_k$, and then select g of them to be the offspring (these offspring are denoted as $p'_1, p'_2, \ldots, p'_g$).

Step 5. Output g offspring vectors $p'_i = (p'_{i1}, p'_{i2}, \ldots, p'_{il})$, for all $1 \le i \le g$.

Example. We choose three parent vectors X_1, X_2, and X_3, and each vector has 6 real values. Consider three-to-one orthogonal crossover. Let the parent vectors be

$$
\begin{aligned}
X_1 &= (x_{1,1}, x_{1,2}, x_{1,3}, x_{1,4}, x_{1,5}, x_{1,6}) \\
X_2 &= (x_{2,1}, x_{2,2}, x_{2,3}, x_{2,4}, x_{2,5}, x_{2,6}) \quad (6) \\
X_3 &= (x_{3,1}, x_{3,2}, x_{3,3}, x_{3,4}, x_{3,5}, x_{3,6}).
\end{aligned}
$$

In Step 2, we can choose $t(1) = 2$, $t(2) = 3$, $t(3) = 1$, $t(4) = 3$, $t(5) = 1$, and $t(6) = 2$. After that, using Step 3, we get 9 vectors $p_1, p_2, \ldots, p_9$ based on the 9 combinations of the orthogonal array in Table 2 as follows:

$$
\begin{aligned}
p_1 &= (x_{1,1}, x_{1,2}, x_{1,3}, x_{1,4}, x_{1,5}, x_{1,6}) \\
p_2 &= (x_{2,1}, x_{2,2}, x_{1,3}, x_{2,4}, x_{1,5}, x_{2,6}) \\
p_3 &= (x_{3,1}, x_{3,2}, x_{1,3}, x_{3,4}, x_{1,5}, x_{3,6}) \\
p_4 &= (x_{1,1}, x_{2,2}, x_{2,3}, x_{2,4}, x_{2,5}, x_{1,6}) \\
p_5 &= (x_{2,1}, x_{3,2}, x_{2,3}, x_{3,4}, x_{2,5}, x_{2,6}) \\
p_6 &= (x_{3,1}, x_{1,2}, x_{2,3}, x_{1,4}, x_{2,5}, x_{3,6}) \\
p_7 &= (x_{1,1}, x_{3,2}, x_{3,3}, x_{3,4}, x_{3,5}, x_{1,6})
\end{aligned}
$$

$$
\begin{aligned}
p_8 &= (x_{2,1}, x_{1,2}, x_{3,3}, x_{1,4}, x_{3,5}, x_{2,6}) \\
p_9 &= (x_{3,1}, x_{2,2}, x_{3,3}, x_{2,4}, x_{3,5}, x_{3,6}).
\end{aligned}
$$

$$(7)$$

Evaluate the fitness of vectors $p_1, p_2, \ldots, p_9$, and then select the best one of them to be the offspring.

4.2.2. Differential Evolution Algorithm. Then we execute DE N times to generate N new vectors. Each new vector V is created by combining three randomly selected trial vectors from the population. This combination process is defined as follows:

$$V = X_1 + F(X_2 - X_3), \quad (8)$$

where X_1, X_2, and X_3 are three randomly selected trial vectors from the population and F is a multiplier, which is the main parameter of the DE.

4.3. Fitness of Image Registration Using the OLDE. We take (3) as a fitness function. Given two images A and B, the aim of the problem becomes finding the best affine transformation T for A so that the normalized mutual information of B and $T(A)$ is maximized.

4.4. The Procedure of Image Registration Using the OLDE. The procedure of the OLDE is as follows.

Step 1 (input). Input target image and reference image.

Step 2 (initialization). Within the range of 6 parameters, randomly initialize population X_0 consisting of N trial vectors, denoted as $X_0 = \{p_1, p_2, \ldots, p_N\}$. Each trial vector is made up of 6 real values, represented as $p_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6})$, where $1 \le i \le N$ and we initialize the generation number g to 0.

Step 3 (population evolution)

Step 3.1. DE, as (8), is executed N times using multiplier F. A new population $X_{gm} = \{p_{1m}, p_{2m}, \ldots, p_{Nm}\}$ is generated with $p_{im} = (x_{im1}, x_{im2}, x_{im3}, x_{im4}, x_{im5}, x_{im6})$, $1 \le i \le N$.

Step 3.2. Randomly choose three vectors in X_{gm} to undergo three-to-three orthogonal crossover. After evaluating the fitness of vectors $p_{1m}, p_{2m}, \ldots, p_{Nm}$ as (3), the worst three vectors of them are recorded. The worst three vectors are replaced with three new vectors generated by the orthogonal crossover. This results in a new population X_{gc}.

Step 3.3. Evaluate the fitness of vectors in X_{gc}, and choose the best one p_g and record its fitness value $f(p_g)$. After that increment the generation number g by 1.

Step 3.4. If stopping criterion is satisfied, then the algorithm terminates, and generate the best vector $p_{g\text{best}} = (x_{\text{best1}}, x_{\text{best2}}, x_{\text{best3}}, x_{\text{best4}}, x_{\text{best5}}, x_{\text{best6}})$; otherwise, return to Step 3.1.
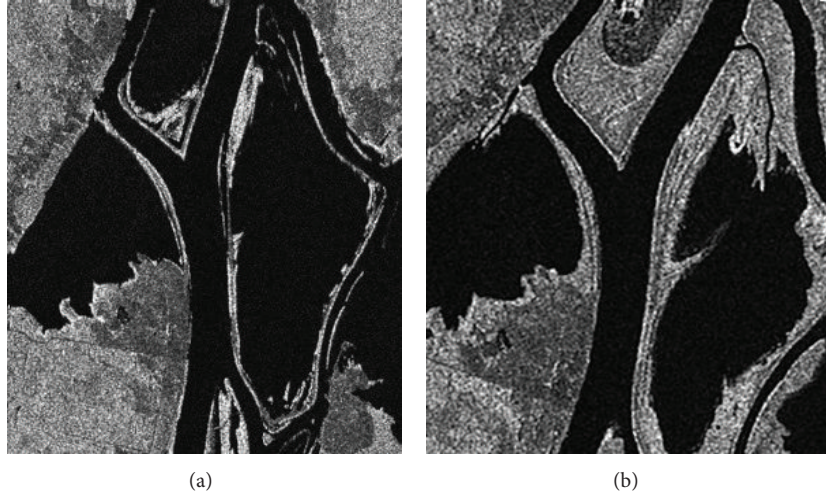
FIGURE 1: Ottawa dataset: (a) target image acquired in May 1997. (b) Reference image acquired in August 1997.

Step 4 (transform the target image). Affine transform the target image as (5), where $a_{11} = x_{\text{best1}}$, $a_{12} = x_{\text{best2}}$, $a_{21} = x_{\text{best3}}$, $a_{22} = x_{\text{best4}}$, $b_1 = x_{\text{best5}}$, and $b_2 = x_{\text{best6}}$, and get T (target) image.

Step 5 (fusion). Fuse the T (target) image and reference image to generate the result image.

In Step 3, the population is evolved and improved iteratively until halting criterion is satisfied. One possible halting criterion is to stop when the number of generation g is equal to a given maximum value. In Step 3.1, we must check that if x_{imj} has been out of the range of jth parameters, we should replace x_{imj} with the initialized value x_{ij}, where $1 \leq i \leq N$, $1 \leq j \leq 6$. In Step 3.2, execute three-to-three orthogonal crossover. An orthogonal array is generated as in Table 2. Carry out the procedure of n-to-g orthogonal crossover. Here, we set $n = 3$ and $g = 3$. In this step, three worst vectors are eliminated. Experiments proved that the speed of convergence can be improved and the diversity of the population can be kept by replacing the three worst vectors with three vectors generated by orthogonal crossover.

5. Experiments and Discussion

To investigate the performance of our method, we have compared it against three image registration methods: genetic algorithm (GA), particle swarm optimization (PSO) [12], and the differential evolution (DE) [30] algorithm.

In both experiments, we set the size of the population $n = 30$ and the maximum allowed number of generations $g = 200$. In the GA algorithm, we set $M = 0.05$, $C = 0.8$, where M is the probability of mutation and C is the probability of crossover of GA algorithm. In the PSO algorithm, the weight ω is declining linearly from 0.9 to 0.4 and $c_1 = c_2 = 2.0$. In the DE algorithm, we set the multiplier FD $= 0.7$ and the probability of crossover CR $= 0.5$. In the OLDE algorithm, we set initial value $F = 0.7$, where F is the multiplier of DE

TABLE 4: Variation ranges of 6 affine transformation parameters.

	a_{11}	a_{12}	a_{21}	a_{22}	b_1	b_2
Min	0.50	−0.50	−0.50	0.50	−200	−200
Max	1.50	0.50	0.50	1.50	200	200

procedure in the OLDE. The allowed variation ranges of 6 affine transformation parameters used in our experiments are shown in Table 4. To test the performance of all methods, we have run each method for 20 times. We record the value of the normalized mutual information (NMI) for the best solution NMI_{best}, the worst solution $\text{NMI}_{\text{worst}}$, and the mean value $\overline{\text{NMI}}$ over the 20 times and ω is the standard deviation (see Tables 5 and 6).

5.1. Ottawa Dataset Task. In the first experiment, we select two images from Ottawa dataset as shown in Figure 1 as the experimental images. These two images are the portions of the city of Ottawa acquired by RADARSAT SAR sensor in May 1997 and August 1997, respectively. They were provided by Defense Research and Development Canada (DRDC), Ottawa. Figure 1(a) shows the image acquired in May 1997 during the summer flooding and Figure 1(b) shows the image acquired in August 1997 after the summer flooding. The resolution of both images is 290×350 with 8 bits per pixels.

In Figure 2, image registration results are shown only for the best experiment out of 20 experiments; that is, an image for each search method corresponding to the highest NMI value is shown in the figure. The matching image is created by one superimposed on the other. Table 5 records the statistical results for each method. The table also contains optimal transformation parameter values obtained by each optimization method. Our method can get the best result compared with the other three methods. In addition, our method can receive much smaller ω value than the other methods. In fact, the standard deviation of 0.0002 for the OLDE is very small. Therefore, the OLDE method is robust with respect to the initial parameter values.
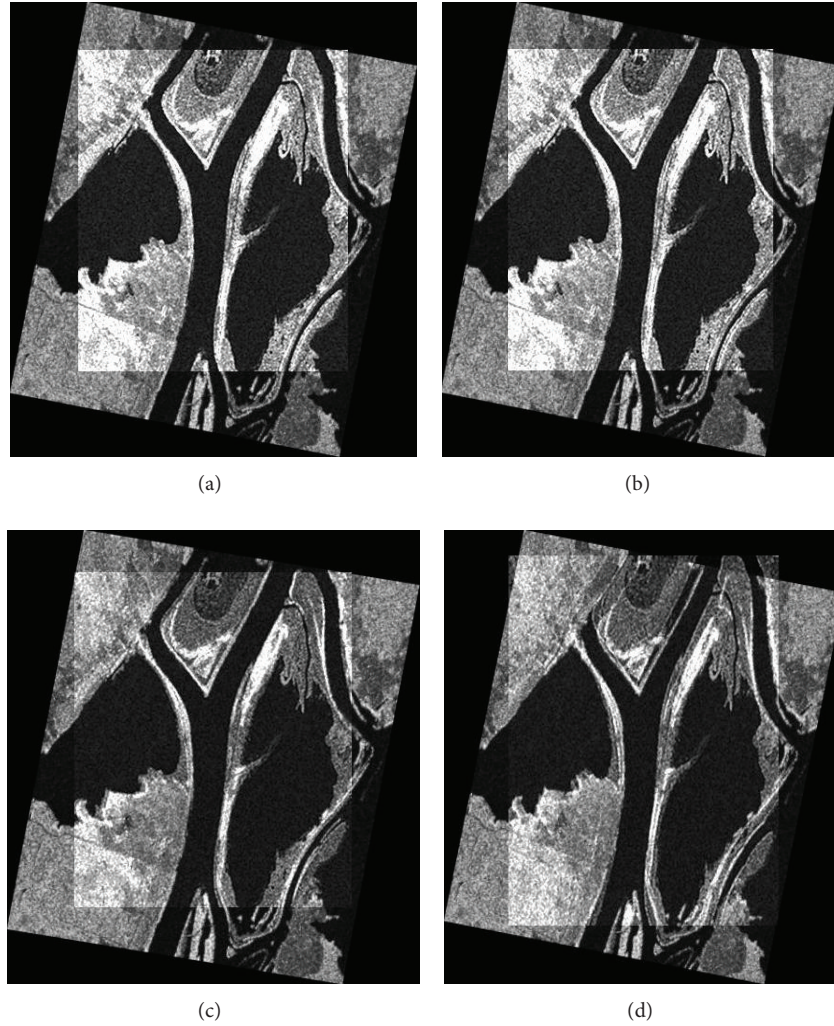
(a)

(b)

(c)

(d)

Figure 2: Result of Ottawa dataset: image (a) is the result of image registration using OLDE. Image (b) is the result of image registration using DE. Image (c) is the result of image registration using PSO. Image (d) is the result of image registration using GA.

Table 5: Comparison of the image registration results on the Ottawa dataset obtained by different methods.

Method	NMI_{best}	NMI_{worst}	$\overline{NMI}$	ω	a_{11}	a_{12}	a_{21}	a_{22}	b_1	b_2
OLDE	0.5588	0.5573	0.5581	0.0002	1.2180	0.2381	−0.2371	1.2135	8.4485	−52.7456
DE	0.5563	0.4847	0.5019	0.0005	1.2213	0.2390	−0.2371	1.2138	9.7243	−50.9686
PSO	0.5509	0.3732	0.4457	0.0204	1.2255	0.2179	−0.2034	1.1520	2.7101	−40.4702
GA	0.5274	0.3633	0.4503	0.0135	1.3817	0.0082	−0.1569	1.0422	−31.1995	−1.6593

Table 6: Comparison of the image registration results on the Yellow River dataset obtained by different methods.

Method	NMI_{best}	NMI_{worst}	$\overline{NMI}$	ω	a_{11}	a_{12}	a_{21}	a_{22}	b_1	b_2
OLDE	0.2187	0.2066	0.2126	0.0002	0.9841	0.3140	−0.3217	0.9494	45.7927	−111.1126
DE	0.2135	0.1865	0.1976	0.0018	0.9482	0.3151	−0.3217	0.9482	45.6404	−111.3498
PSO	0.1976	0.1420	0.1765	0.0147	0.9260	0.4094	−0.4227	0.9283	39.3913	−102.9269
GA	0.1618	0.1256	0.1469	0.0130	1.1479	0.3741	−0.2629	0.9641	−41.8955	−136.3805

TABLE 7: Comparison of the image registration results on the Yellow River dataset obtained by different methods.

Method	NMI$_{best}$	NMI$_{worst}$	$\overline{\text{NMI}}$	ω	a_{11}	a_{12}	a_{21}	a_{22}	b_1	b_2
OLDE	0.6974	0.6950	0.6959	0.0001	0.9492	0.2523	−0.3966	0.9524	22.1166	−44.9263
DE	0.6317	0.5985	0.6027	0.0014	0.9680	0.2982	−0.3572	0.9333	25.5351	−45.7148
PSO	0.6292	0.3691	0.5527	0.0250	1.0163	0.2645	−0.4910	1.0429	58.4900	−11.6157
GA	0.1918	0.1479	0.4974	0.0130	1.1054	0.1851	−0.2129	0.9952	30.0790	−113.4680
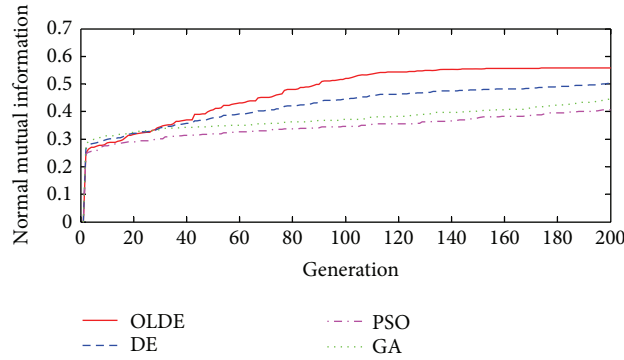
FIGURE 3: Result of Ottawa dataset: behavior of GA, PSO, DE, and OLDE for optimizing the normalized mutual information.

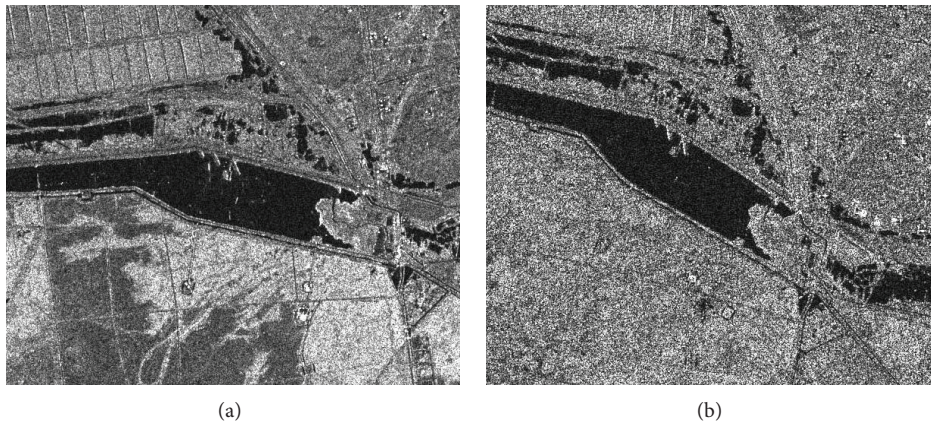

(a)   (b)

FIGURE 4: Yellow River dataset: (a) target image was acquired in June 2008. (b) Reference image was acquired in June 2009.

In Figure 3, the NMI values are shown as a function of the generation number. As shown in the figure, the OLDE has a much higher NMI value than the other methods. Thus, our method outperforms all the comparison methods.

5.2. Yellow River Dataset Task. In the second experiment, we use two images acquired by RADARSAT-2 at the Yellow River Estuary region in China in June 2008 and June 2009 as the experimental images. The two images are shown in Figure 4. The resolution of both Figures 4(a) and 4(b) is 600 × 500 pixels with an 8-bit dynamic range.

The numerical results of independently running the optimization algorithms 20 times are shown in Table 6. Also in this case, the OLDE method is the best method in terms of the NMI. Our method can receive much smaller ω value than DE, PSO, and GA. Therefore, the OLDE method is robust with respect to the initial parameter values.

In Figure 5, the NMI values are shown as a function of a generation number. As can be observed from the results, with

the increase of generation, the performance of the other two methods, PSO and GA, is far worse than the OLDE method and the DE method. In more detail, after the 60th generation, the OLDE method emerges as a better method than DE method. Therefore, using OL to guide DE can improve the performance of DE method.

In Figure 6, image registration results are shown only for the best experiment out of 20 experiments; that is, an image for each search method corresponding to the highest NMI value is shown in the figure. The best results for each one of the four optimization methods are included in the figure.

In the third experiment, we also select two images acquired by RADARSAT-2 at the Yellow River Estuary region in June 2008 and June 2009. The two images are shown in Figure 7. The resolution of both images is 400 × 400 pixels with an 8-bit dynamic range.

The numerical results of independently running the optimization algorithms 20 times are shown in Table 7. Also in this case, the OLDE method can receive the largest value
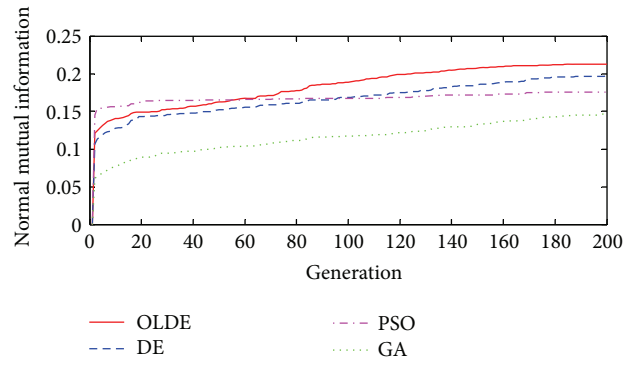
FIGURE 5: Result of Yellow River dataset: behavior of GA, PSO, DE, and the OLDE for optimizing the normalized mutual information.
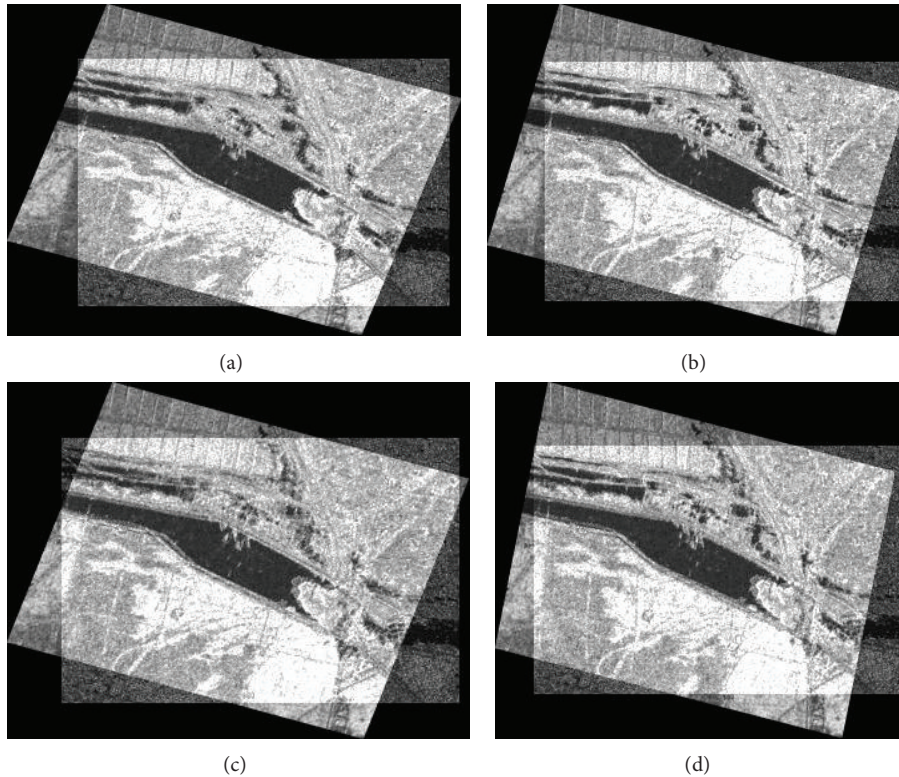

(a)

(b)

(c)

(d)

FIGURE 6: Result of Yellow River dataset: image (a) is the result of image registration using the OLDE method. Image (b) is the result of image registration using DE. Image (c) is the result of image registration using PSO. Image (d) is the result of image registration using GA.
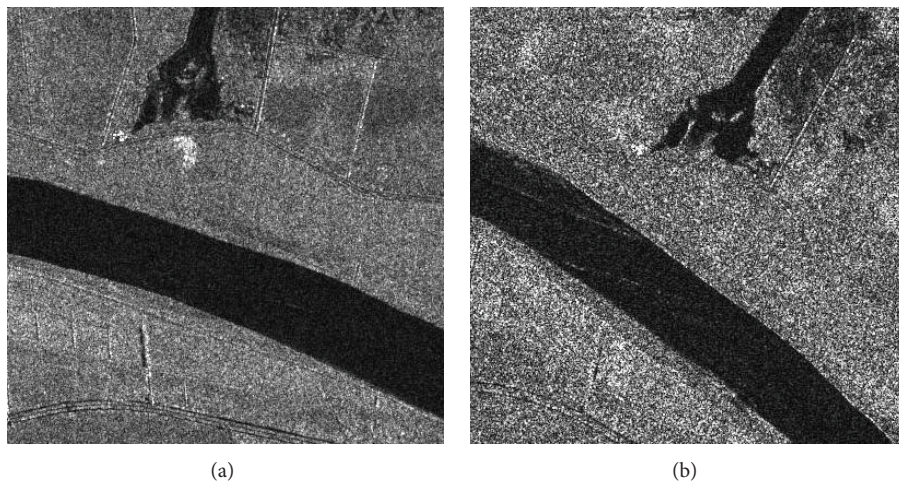

(a)

(b)

FIGURE 7: Yellow River dataset: (a) target image was acquired in June 2008. (b) Reference image was acquired in June 2009.
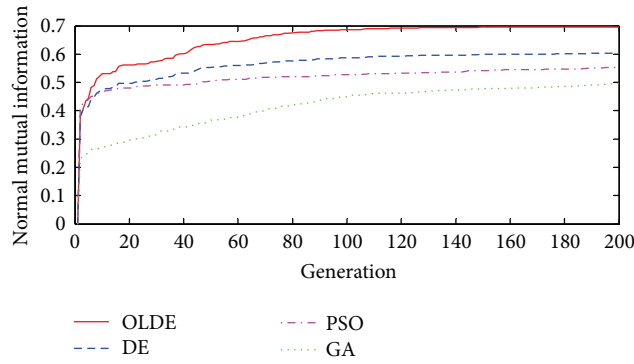
FIGURE 8: Result of Yellow River dataset: behavior of GA, PSO, DE, and the OLDE for optimizing the normalized mutual information.
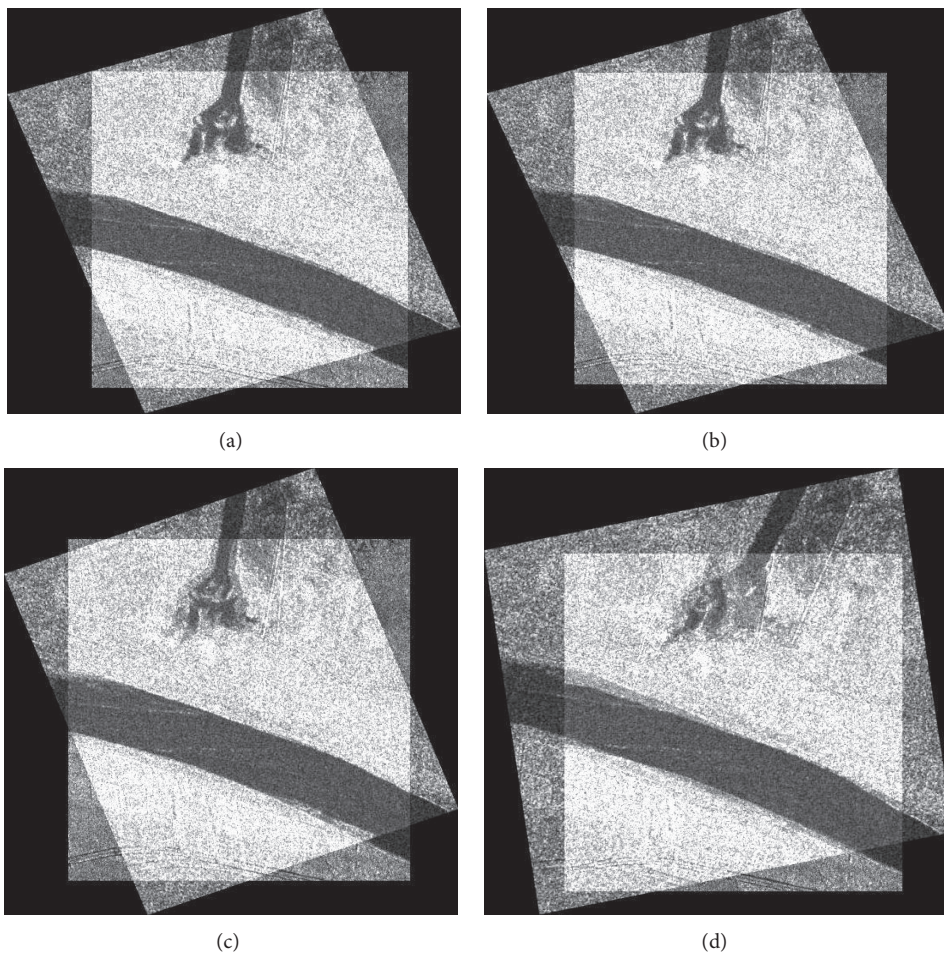

(a)

(b)

(c)

(d)

FIGURE 9: Result of Yellow River dataset: image (a) is the result of image registration using the OLDE method. Image (b) is the result of image registration using DE. Image (c) is the result of image registration using PSO. Image (d) is the result of image registration using GA.

of the NMI and much smaller ω value. In Figure 8, the NMI values are shown as a function of a generation number. In Figure 9, image registration results are shown only for the best experiment out of 20 experiments; that is, an image for each search method corresponding to the highest NMI value is shown in the figure. The best results for each one of the four optimization methods are included in the figure.

6. Conclusion

This paper proposed a method for remote sensing image registration using the orthogonal learning differential evolution (OLDE). The orthogonal learning (OL) strategy can construct a guidance exemplar with an ability to predict promising search directions toward the global optimum. Differential

evolution (DE) is a version of evolutionary algorithm (EA) that has proven to be fast and reliable in many applications. The OLDE method uses the OL strategy to guide the DE algorithm to select promising search directions towards the global optimum.

To investigate the performance of our method, we have compared it against three image registration methods: genetic algorithm (GA), particle swarm optimization (PSO), and the differential evolution (DE) algorithm. The OLDE method was shown to be able to achieve the best image registration results in terms of the normalized mutual information similarity metric. Furthermore, experiments showed that the OLDE method was robust and efficient with respect to the initial parameter values.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] Z. Z. Shi, S. Vadera, A. Aamodt, and D. Leake, "Pattern recognition," in *Intelligent Information Processing V*, pp. 326–351, 2010.

[2] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[3] M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4328–4338, 2014.

[4] Y. Wu, W. P. Ma, M. G. Gong, L. Z. Su, and L. C. Jiao, "A novel point-matching algorithm based on fast sample consensus for image registration," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 1, pp. 43–47, 2015.

[5] H. Goncalves, J. A. Goncalves, and L. Corte-Real, "HAIRIS: a method for automatic image registration through histogram-based image segmentation," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 776–789, 2011.

[6] D. L. G. Hill and P. Batchelor, "Registration methodology: concepts and algorithms," in *Medical Image Registration*, pp. 39–67, 2001.

[7] A. Roche, A. Pennec, G. Malandain, and N. Ayache, "Rigid registration of 3-d ultrasound with mr images: a new approach combining intensity and gradient information," *IEEE Transactions on Medical Imaging*, vol. 20, no. 10, pp. 1038–1049, 2001.

[8] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Physics in Medicine and Biology*, vol. 46, no. 3, pp. R1–R45, 2001.

[9] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997.

[10] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[11] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.

[12] M. Wachowiak, R. Smolíková, Y. Zheng, J. M. Zurada, and A. S. Elmaghraby, "An approach to multimodal biomedical image registration utilizing particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 289–301, 2004.

[13] F. Maes, D. Vandermeulen, and P. Suetens, "Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information," *Medical Image Analysis*, vol. 3, no. 4, pp. 373–386, 1999.

[14] J. L. Bernon, V. Boudousq, J. F. Rohmer et al., "A comparative study of Powell's and downhill simplex algorithms for a fast multimodal surface matching in brain imaging," *Computerized Medical Imaging and Graphics*, vol. 25, no. 4, pp. 287–297, 2001.

[15] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual information matching in multiresolution contexts," *Image and Vision Computing*, vol. 19, no. 1-2, pp. 45–52, 2001.

[16] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical Image Analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[17] G. K. Matsopoulos, N. A. Mouravliansky, K. K. Delibasis, and K. S. Nikita, "Automatic retinal image registration scheme using global optimization techniques," *IEEE Transactions on Information Technology in Biomedicine*, vol. 3, no. 1, pp. 47–60, 1999.

[18] J.-M. Rouet, J.-J. Jacq, and C. Roux, "Genetic algorithms for a robust 3-D MR-CT registration," *IEEE Transactions on Information Technology in Biomedicine*, vol. 4, no. 2, pp. 126–136, 2000.

[19] R. He and P. A. Narayana, "Global optimization of mutual information: application to three-dimensional retrospective registration of magnetic resonance images," *Computerized Medical Imaging and Graphics*, vol. 26, no. 4, pp. 277–292, 2002.

[20] M. P. Wachowiak and A. S. Elmaghraby, "The continuous tabu search as an optimizer for 2D-to-3D biomedical image registration," in *Proceedings of Medical Image Computing and Computer Assisted Intervention*, vol. 2208, pp. 1273–1274, Springer, Berlin, Germany, 2001.

[21] H. V. Phan, M. Lech, and T. D. Nguyen, "Registration of 3D range images using particle swarm optimization," in *Advances in Computer Science—ASIAN 2004. Higher-Level Decision Making*, vol. 3321 of *Lecture Notes in Computer Science*, pp. 223–235, Springer, Berlin, Germany, 2005.

[22] X. Yang, J. Pei, and W. Xie, "Maximization of feature potential mutual information in multimodality image registration using particle swarm optimization," in *Medical Imaging—Image Processing*, Proceedings of the SPIE, pp. 1300–1309, San Diego, Calif, USA, February 2005.

[23] K. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, Natural Computing Series, Springer, 2005.

[24] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, 1995.

[25] K. Kachitvichyanukul, "Comparison of three evolutionary algorithms: GA, PSO, and DE," *Industrial Engineering & Management Systems*, vol. 11, no. 3, pp. 215–223, 2012.

[26] K. Chandrasekar and N. V. Ramana, "Performance comparison of GA, DE, PSO and SA approaches in enhancement of total transfer capability using facts devices," *Journal of Electrical Engineering & Technology*, vol. 7, no. 4, pp. 493–500, 2012.

[27] R. Storn and K. Price, "Differential evolution-a simple and efficient heuristic for global optimization over continuous

spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[28] M. Omran, A. P. Engelbrecht, and A. Salman, "Differential evolution methods for unsupervised image classification," in *Proceedings of the IEEE Congress on Evolutionary Computation*, vol. 2, pp. 966–973, 2005.

[29] I. D. Falco, A. D. Cioppa, and E. Tarantino, "Automatic classification of handsegmented image parts with differential evolution," *Computer Science*, vol. 3907, pp. 403–414, 2006.

[30] I. de Falco, A. Della Cioppa, D. Maisto, and E. Tarantino, "Differential Evolution as a viable tool for satellite image registration," *Applied Soft Computing Journal*, vol. 8, no. 4, pp. 1453–1462, 2008.

[31] Z.-H. Zhan, J. Zhang, Y. Li, and Y.-H. Shi, "Orthogonal learning particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 6, pp. 832–847, 2011.

[32] Q. Zhang and Y.-W. Leung, "An orthogonal genetic algorithm for multimedia multicast routing," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 1, pp. 53–62, 1999.

[33] Y.-W. Leung and Y. Wang, "An orthogonal genetic algorithm with quantization for global numerical optimization," *IEEE Transactions on Evolutionary Computation*, vol. 5, no. 1, pp. 41–53, 2001.

[34] X. M. Hu, J. Zhang, and J. H. Zhong, "An enhanced genetic algorithm with orthogonal design," in *Proceedings of the IEEE Transactions on Evolutionary Computation*, pp. 3174–3181, 2006.

[35] S.-Y. Ho, S.-J. Ho, Y.-K. Lin, and W. C.-C. Chu, "An orthogonal simulated annealing algorithm for large floorplanning problems," *IEEE Transactions on Very Large Scale Integration (VLSI) System*, vol. 12, no. 8, pp. 874–876, 2004.

[36] S.-J. Ho, S.-Y. Ho, and L.-S. Shu, "OSA: orthogonal simulated annealing algorithm and its application to designing mixed H_2/H_∞ optimal controllers," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans.*, vol. 34, no. 5, pp. 588–600, 2004.

[37] X.-M. Hu, J. Zhang, and Y. Li, "Orthogonal methods based ant colony search for solving continuous optimization problems," *Journal of Computer Science & Technology*, vol. 23, no. 1, pp. 2–18, 2008.

[38] S.-Y. Ho, H.-S. Lin, W.-H. Liauh, and S.-J. Ho, "OPSO: orthogonal particle swarm optimization and its application to task assignment problems," *IEEE Transactions on Systems, Man, and Cybernetics A: Systems and Humans*, vol. 38, no. 2, pp. 288–298, 2008.

[39] B. Likar and F. Permiš, "A hierarchical approach to elastic registration based on mutual information," *Image and Vision Computing*, vol. 19, no. 1-2, pp. 33–44, 2001.

[40] D. C. Montgomery, *Design and Analysis of Experiments*, John Wiley & Sons, New York, NY, USA, 3rd edition, 1991.

Submit your manuscripts at
http://www.hindawi.com