A Modified PSO Algorithm for Remote Sensing Image Template Matching

Ru An, Peng Gong, Huilin Wang, Xuezhi Feng, Pengfeng Xiao, Qi Chen, Qing Zhang, Chunye Chen, and Peng Yan

Abstract
Image template matching is essential in image analysis and computer vision tasks. Cross-correlation algorithms are often used in practice, but they are sensitive to nonlinear changes in image intensity and random noise, and are computationally expensive. In this paper, we propose a template-matching algorithm based on a modified particle swarm optimization (PSO) procedure with a mutual information (MI) similarity measure. The influence of MI on the performance of template matching, calculated by different histogram bins, is analyzed first. A modified PSO method (C-MPSO) is then presented. The proposed algorithm is tested with remote sensing imagery from different sensors and for different seasons. Our experimental results indicate that the proposed approach is robust in practical scenarios and outperforms the standard PSO, multi-start PSO, and cross-correlation algorithms in accuracy and efficiency with our test data. The proposed method can be used for position estimation of aircraft, object recognition, and image retrieval.

Introduction
Image template matching is a process of determining the presence and location of an input image or an object inside a reference image (Choi et al., 2002). The input image and the reference image may be dissimilar because they may exhibit relative translation, rotation, and different scales or contain additive random noise due to different sensors and image captured at different dates. All these factors make image matching a challenging task (Brown, 1992; Zitova et al., 2003). Image template matching has been widely used for object recognition, image retrieval, motion tracking, aircraft position estimation, change detection, and multi-image registration (An et al., 2005; Gong et al., 1992; Liu et al., 2006; Oh et al., 2006; Sim et al., 1999a and 1999b).

Cross-correlation-based algorithms such as the normalized cross-correlation (NCC), the mean absolute error (MAE), and the mean squares error (MSE) methods are often used in image template matching. These methods do not generally require extensive preprocessing, such as segmentation or feature extraction, but they often have a lower probability of correct matching when the image intensity has nonlinear changes. They are also computationally expensive. Other commonly used methods in template matching are feature-based techniques. These methods often fail when the images to be matched contain few salient features such as points, lines, and regions (Brown, 1992; Zitova et al., 2003).

Position estimation of aircraft through image matching is an important method for autonomous navigation. Because of its benefits of higher self-determination and accuracy, the development of image-based navigation systems has been a hot research topic (Sim et al., 1999a and 1999b; Oh et al., 2006). Various approaches, such as terrain contour matching (TERCOM), inertial navigation systems (INS), and global positioning systems (GPS), have been used for navigation, but they all have some drawbacks. For example, the estimation error by the INS tends to increase as an aircraft goes on flying, the GPS can be disturbed by external signals, and terrain contour matching encounters difficulties in estimating its own position in the plain regions where change in the elevation is small and can be out of control by external signals (Oh et al., 2006). In practice, a hybrid system, such as, INS and image-based navigation system is commonly used. INS usually supplies orientation and altitude information of the aircraft and image-based navigation is used to rectify cumulative error and increase accuracy of INS systems.

In general, navigation systems have been developed under the assumption that an aircraft flies along a predetermined (planned) trajectory. Therefore, it is assumed that the system contains reference images, and it is thus possible to measure the performance of various matching methods (Oh et al., 2006). A matching algorithm suitable for position estimation of aircraft has to meet the requirements of reliability (higher success rate for matching), accuracy
(precise location), and real-time implementation. Currently, few approaches can meet all these requirements. It is especially difficult to achieve a high success rate and real-time efficiency at the same time.

A template matching algorithm for absolute position estimation (APE) was introduced based on the Hausdorff distance (HD) (Sihn et al., 1999a and 1999b), in which aerial images were used as import images, and SPOT images were used as reference images. An accumulated buffer (AB) matching method was suggested by Oh et al. (2006), in which some accumulated buffers, such as those employed in the Hough transform, are used as a set of accumulator cells. AB matching is faster than HD matching, and has performance comparable to that of NCC matching.

From the studies mentioned above, we find that the performance of commonly used matching methods does not satisfy practical applications. The NCC method often exhibits slow matching speed and a low matching rate when the image intensity is nonlinear, while AB matching exhibits a fast speed but a low matching rate. Therefore, a more robust similarity measure and a faster matching strategy are required.

Mutual information (MI) has been considered a suitable similarity measure for template matching in which the images to be matched are dissimilar. It is more flexible and independent of image content (Chen et al., 2003). Since its introduction (Viola and Wells, 1995; Maes et al., 1997), MI has been used widely in many multi-modal medical image registration problems because of its generality and high accuracy (Pluim et al., 2001; Lo, 2003; Xu et al., 2007). The registration of remote sensing images using mutual information has also been studied by Ingleda (2002), Johnson et al. (2001), Cole-Rhodes et al. (2003), Chen et al. (2003), and Kern (2007). They found that MI produced consistently sharper peaks at the correct registration position than cross correlation, which is important for accurate registration. They also found that mutual information was more robust against Gaussian noise when working at low signal to noise ratio. MI is also robust against nonlinear intensity relationships between images, such as contrast reversals.

The search strategy optimizes the similarity metric. A fast search strategy is very important in practical applications in order to improve efficiency. Examples include local or global searches, multi-resolution approaches, or other optimization techniques. Particle swarm optimization (PSO) is an evolutionary computation technique. Since its introduction (Kennedy and Eberhart, 1995), it has attracted considerable attention from researchers because of its simplicity in concept, requirement of fewer parameters, and ease of implementation (Eberhart et al., 2004). It has found applications in many areas, such as evolving neural networks, tracking dynamic systems, tackling multi-objective optimization, constraint optimization problems, reactive power, and voltage control and ingredient mix optimization (Eberhart et al., 2004; Hu et al., 2004; Alrashidi et al., 2006). In image processing, Wachowiak et al. (2004) adapted particle swarm optimization for 3D-to-3D biomedical image registration. The images to be matched were obtained from different modalities. PSO was used as a search strategy to maximize the similarity metric for registering single slice biomedical images to 3D volumes. Knowledge can be incorporated in the initialization of the PSO. Du et al. (2005) applied PSO in search of the maximum threshold for infrared image segmentation. Zahara et al. (2005) also used a PSO based hybrid optimization approach to optimize multi-threshold image segmentation. Maitra et al. (2007) used a PSO based hybrid cooperative-comprehensive learning algorithm for optimizing multilevel threshold for image segmentation. Yin (2006) applied PSO for point pattern matching, and the experimental results indicate that the PSO-based method is superior to a genetic algorithm and a simulated annealing algorithm in terms of both effectiveness and efficiency.

The main objective of this research is to improve both the accuracy and efficiency in image matching. The focus is on the search strategy for maximizing the similarity metric (MI) in matching. We propose a modified PSO method (CRI-PSO) that improves the swarm diversity and enlarges the search space. It does not destroy the structure of the current particles, and might have high search accuracy. We empirically show that the CRI-PSO achieves better performance than standard PSO and some PSO variants in template matching. A novel image matching method based on MI and CRI-PSO is suggested for aircraft navigation. Generally, INS can ensure only a small flying parametric error, such as 0.5 degree rotation and less than four percent scale variance between the input image and the reference image, so, in the paper, attention is mainly paid on translation transformation.

The paper is organized as follows. In the next Section, a brief discussion of the similarity metric and the influence of different histogram bin sizes on the matching performance are introduced followed by a description of the adaptation of PSO to template matching, and an improved PSO for template matching. An experiment to test the modified algorithm and compare its performance with several other methods is then presented, and finally, some concluding remarks regarding the proposed method.

Similarity Metric

Similarity metrics are indicators of how closely the features or intensity values of two images match. They must be robust; that is, they should attain a global maximum at the correct matching position.

Mutual Information Similarity Metric

Mutual information is a concept developed from information theory. It indicates how much information one random variable tells about another. The MI registration criterion can be thought of as a measure of how well one image explains the other. It is applied to measure the statistical dependence between image intensities of corresponding pixels in both images, which is assumed to be a maximum if the images are correctly aligned geometrically.

Mutual information, $I(U,V)$, of two random variables $U$ and $V$ can be obtained from Equation 1 (Cover et al., 1991):

$$I(U,V) = H(U) + H(V) - H(U,V)$$

where $H(U)$ and $H(V)$ are the entropies of $U$ and $V$, and $H(U,V)$ is their joint entropy. Considering $U$ and $V$ as two images, the MI-based matching criterion states that the images shall be registered when $I(U,V)$ is maximal. The entropies and joint entropy can be computed from:

$$H(U) = -\sum_{u} p_u(u) \log p_u(u)$$

$$H(V) = -\sum_{v} p_v(v) \log p_v(v)$$

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

where $p_u(u)$ and $p_v(v)$ are the marginal probability mass functions, and $p_{uv}(u,v)$ is the joint probability mass function. These probability mass functions can be obtained from their histograms and joint histogram, respectively (Viola and Wells, 1995; Maes et al., 1997; Chen et al., 2003).
The size of the overlapping part of the images influences the mutual information measure. A normalized measure of mutual information has been proposed (Studholme et al., 1999), which is less sensitive to changes in overlap:

\[ NMI(U, V) = \frac{H(U) + H(V)}{H(U, V)}. \] (5)

In all the following experiments (except specific indication is given), Equation 5 is used to calculate similarity measure value for matching.

The Number of Histogram Bins and Mutual Information-based Template Matching

In template matching, the size of the input image is often small (e.g., 65 \( \times \) 65), so determining how many histogram bins should be used to estimate probability distributions is a problem. If the number of bins is too large (e.g., 256 bins), the statistical power of the probability distribution estimation by mutual information will be reduced due to the small number of data points (e.g., pixels for the image) in a bin (Knops et al., 2006). If the number of bins is too small (e.g., 4 bins), important information might be lost. Moreover, in order to reduce the computational complexity, the number of bins used should be as small as possible. Therefore, it is important to select a reasonable number of histogram bins for MI-based template matching. We achieved this aim with empirical experiment. A histogram with 16 bins produced a more highly successful matching rate than a 256 bins histogram. Here, the success rate of matching refers to the ratio between the numbers of successful trials and total trials. For each test image cropped from the input source image, we assume that its theoretically correct matching position in the reference image (e.g., upper left corner) is known in advance. Here, a successful match means that the matching error is within three pixels. The matching error denotes the difference between the theoretically correct matching position and the actual matching position along the x-axis and/or y-axis, respectively. In Figure 1, the correct matching position of an input image is (10, 99), whereas its actual matching position is (13, 98). Since its location error is in the allowable range (less than three pixels along the x-axis and/or y-axis), it is accepted as a successful match. Computation time refers to the time used in a matching between the reference image and an input image. Moreover, the reduced number of bins dramatically reduces the runtime for MI-based template matching. This image area in Figure 1 is the Farmland.

In Figure 2, the surface of MI similarity measure using different numbers of histogram bins is shown. We can see in Figure 2a that the MI similarity surface is non-convex and irregular, because the size of the images to be matched is only 65 \( \times \) 65 pixels. This leads the distribution of each pixel sample to spread when the number of bins is 256. We can also observe that the peak of the MI similarity surface at the correct matching position, which is pre-defined (162, 56), is not salient, and it is not the global optimum. The reduced number of histogram bins makes the peak of the histogram much sharper and more distinctive (Figure 2b), and it appears at the correct matching position when the 16-bin histogram is used. However, when the number of bins is too small, it results in multiple peaks in the similarity surface (shown in Figure 2c). This leads to an increased number of incorrect matches. In Figure 2c, we find that the peak of the MI similarity surface at the correct match position is not salient, and it is not the global optimum. Therefore, it is necessary to select a reasonable number of bins for successful matching.

The success rates of MI-based image matching for different numbers of bins are summarized in Table 1. The test data are shown in Figures 1 and 4. We can see that the success rate is greatest when 16 histogram bins are used, and the number of input images for the trials is 100, respectively.

The computation time using MI-based template matching with 16 histogram bins is relatively low (Table 2). The test data of Table 2 is a Stream image pair. Considering the trade-off between the success rate and computation time, we choose 16 histogram bins as a desirable number of bins for relatively high match rates and low computation cost. For a joint histogram, the number of bins is 16 \( \times \) 16.

Figure 1. The MI-based template-matching concept: (a) IRS-C source of the template image, (b) a template image cropped from (a), and (c) match result for (b) in (c).
good solutions much more quickly than other evolutionary algorithms (Eberhart et al., 1998). It is conceptually simple, requires fewer parameters, and is easy to implement (Eberhart et al., 2004).

The Particle Swarm Optimization (PSO) Algorithm

The PSO is an evolutionary computation technique. It is a population-based search algorithm and is initialized with a population of random solutions, called a particle swarm. Here, one particle indicates a possible matching position in the input image inside the reference image, and is generally randomly produced. A particle swarm consists of a number of particles or “possible solutions” that proceed (fly) through the feasible solution space to explore optimal solutions.

In PSO, the coordinates of each particle represent a possible solution associated with two vectors, the position $X_i(t)$ and velocity $V_i(t)$ vectors. $V_i(t)$ denotes a change of a particle’s position. The size of vectors $X_i(t)$ and $V_i(t)$ is equal to the problem space dimension $D$. Here, $D$ equals two, since only the translations along the $x$-axis and $y$-axis are considered. Assume a swarm consisting of $M$ particles. The position of the $i^{th}$ particle is, in effect, a $D$-dimensional vector:

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})^T.$$  

The velocity of this particle is also a $D$-dimensional vector:

$$V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})^T$$

where $i = 1, 2, \ldots, M$ is a particle index. $V_i$ is dynamically adjusted in terms of a particle’s own previous best solution, $pbest_i(t)$, and the previous best solution of the entire swarm, $gbest(t)$. The velocity updates are calculated as a linear combination of the position and velocity vectors according to the following equations (Shi et al., 1998):

$$V_i(t + 1) = \omega V_i(t) + c_1 \times \text{rand()}(pbest_i(t) - X_i(t)) + c_2 \times \text{rand()}(gbest(t) - X_i(t))$$  

(6)

$$X_i(t + 1) = X_i(t) + V_i(t + 1)$$  

(7)

where $t$ is the iteration number, $X_i(t)$ is the current position of particle $i$, $\text{rand()}$ is the random number between zero and one, $c_1$ and $c_2$ are learning factors, $pbest_i(t)$ is the best position of particle $i$ achieved based on its own experience, $gbest(t)$ is the best particle position based on the overall swarm experience, and $pbest(t)$ and $gbest(t)$ are given by the following equations, respectively:

![Figure 2. Surface of the MI similarity measure using different numbers of bins: (a) 256 bins, (b) 16 bins, and (c) 4 bins.](image)

**Table 1. The Success Rate of Template Matching Based on Different Numbers of Bins**

<table>
<thead>
<tr>
<th></th>
<th>Farmland image</th>
<th>Stream image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of bins</td>
<td>Success rate (%)</td>
<td>Number of bins</td>
</tr>
<tr>
<td>Normalized mutual information</td>
<td>256</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td>128</td>
<td>25.5</td>
</tr>
<tr>
<td></td>
<td>64</td>
<td>47.7</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>74.4</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>80.0</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>72.2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>64.3</td>
</tr>
</tbody>
</table>

**Table 2. Computation Time Required with Different Numbers of Bins in the Normalized Mutual Information-Based Template Matching with Exhaustive Search**

<table>
<thead>
<tr>
<th>Number of bins</th>
<th>Time cost (S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>256</td>
<td>20.8</td>
</tr>
<tr>
<td>128</td>
<td>11.4</td>
</tr>
<tr>
<td>64</td>
<td>7.5</td>
</tr>
<tr>
<td>32</td>
<td>5.9</td>
</tr>
<tr>
<td>16</td>
<td>4.5</td>
</tr>
<tr>
<td>8</td>
<td>4.4</td>
</tr>
<tr>
<td>4</td>
<td>4.3</td>
</tr>
</tbody>
</table>
the MI function, and is an inertia weight that is employed
in PSO2. A Modified PSO: CRI-PSO Algorithm

A Modified PSO: CRI-PSO Algorithm

One of the methods to satisfy various application purposes,
and to prevent premature convergence of the PSO, is to
enhance the particle diversity. Some researchers have
merged or combined PSO with other evolutionary computa-
tion techniques (Lvbjerg et al., 2001; Higashi et al., 2003)
to enhance the particle diversity. Van den Bergh (2001)
proposed a multi-start PSO, which maintains the historic
optimal position of the particle swarm and reinitializes all
the particles after having iterated for a certain amount of
time in order to improve the swarm diversity and enlarge the
searching space. This algorithm can escape from the
local minimum and ensure that the search will converge to
the global optimization, but it destroys the structure of the
current particles with reinitialization, which results in a
slowing down of the convergence speed and a great decrease
of search accuracy (Li et al., 2004). It is found that reinitial-
ization of particle velocity is a powerful method to avoid the
premature convergence of the swarm (Pasupuleti et al.,
2006), but how and when to implement reinitialization is
important to overcome the drawback of the multi-start PSO.
In the following, we introduce an improved condition for
PSO reinitialization (CRI-PSO).
CRI-PSO accepts that each individual is not simply
influenced by the best performer among its neighbors. If
each particle has avoided premature convergence in the
earby stages of the search, the capacity for convergence to
the global optimum could be enhanced during subsequent
stages of the search. Thus, in each iteration in CRI-PSO, if one
particle’s f(pbest) (i.e., its historical best fitness function
value, i.e., the MI similarity measure value at position pbest)
does not vary after two continuous iterations, this particle’s
velocity will be reinitialized. Additionally, the total number of
particles to be reinitialized in each iteration is less than a
threshold Npredefined to prevent much more particles from
being reinitialized at the later stages of the search. The
velocity vector V(t + 1) is updated as follows:

\[ V(t + 1) = \begin{cases} 
    \chi(t) V(t) + c_1 \cdot \text{rand} \cdot (pbest(t) - X(t)) & \text{if } f(pbest(t)) = f(pbest(t - 1)) \\
    \text{else } V(t + 1) = \text{rand} \cdot (V_{\text{max}} - V_{\text{min}}) + c_2 \cdot \text{rand} \cdot (gbest(t) - X(t)) 
\end{cases} \]

In this manner, whether a particle is reinitialized or not is
determined by its solution quality, i.e., whether its f(pbest)
value is improved after a certain number of iterations. If the
particle is trapped into a local optimum, it would be
reinitialized at once. For each time of iteration, some
particles with the worst solution quality will be reinitial-
ized. Therefore, they can avoid premature convergence.
Particles with better solution quality will update their
velocity by the constriction factor. This avoids blind reini-
itialization for all particles. Thus, with CRI-PSO, the swarm’s
diversity can be improved and the searching space can be
enlarged. In addition, it does not destroy the structure of the
current particles. High search accuracy is empirically
validated (see Table 5). From Table 5, we can see that the
success rate for matching is improved, from 59.8 percent to
76.1 percent for PSO1 and CRI-PSO, respectively.

This approach is different from that proposed by Li
et al. (2004) at least in two respects. The first is the condi-
tion under which reinitialization should be implemented.
The condition of reinitialization in Li et al. (2004) is that the
global historical optimal position gbest remains continuously
unchanged or has changed only a little. In other words,
the condition for reinitialization is whether the particle swarm
is aggregated heavily. The maximum Euclidean distance
between all particles and gbest were used to express the
degree of aggregation of the particle swarm. The reinitial-
ization condition needed in CRI-PSO is based on individual
particles. The second difference is who should be reinitial-
ized. In the approach proposed by Li et al. (2004), once the
condition under which reinitialization occurs is satisfied,
gbest should be kept and some dimensions of all particles’
position and velocity vectors in the swarm should be
reinitialized according to the mutation probability p\text{m.} In CRI-
PSO, once one particle’s f(pbest) does not vary after two
continuous iterations, only this particle’s velocity is reini-
itialized randomly. Its position vector X\text{t should not be
reinitialized randomly. On the contrary, X should still be
adjusted according to Equation 7.

Figure 3 shows the curves of the similarity measure
value obtained at gbest(f(gbest)) for the CRI-PSO and the
standard PSO along with the times of iteration, respectively.
For the standard PSO, it can be seen that a large variation
occurred for the similarity measure value before 20 itera-
tions, and an M of 0.064 was found quickly but no further
variation took place in subsequent iterations. This indicates

\[
\omega(t + 1) = \omega(t) + do, \quad do = \frac{(\omega_{\text{min}} - \omega_{\text{max}})}{T}
\]

We refer to Equations 6, 7, 8, and 9 as the Standard PSO.
Like other global optimization algorithms, PSO has the
property of being trapped into premature convergence for
introduced a constriction factor into PSO to control the
convergence properties of the particles. If this constriction
factor is included in Equation 6, then:

\[
V(t + 1) = \chi V(t) + c_1 \cdot \text{rand} \cdot (pbest(t) - X(t)) \\
+ c_2 \cdot \text{rand} \cdot (g\text{best}(t) - X(t))
\]

with

\[
\chi = \frac{2k}{1 - c - \sqrt{c^2 - 4c} l}
\]

\[
c = c_1 + c_2, \quad c > 4, \quad k \in [0, 1].
\]

In our experiment, we set k = 1.0, c_1 = 2.8, c_2 = 1.3
(Wachowiak et al., 2004; Carlisle and Dozier, 2001). This
variant is commonly used in many fields and promising in
reaching the optimal performance. Here, this approach is
named PSO2.
that the standard PSO method has a strong exploring ability in the earlier iterations, but once it is trapped into a local optimal position, it is difficult to find the global optimal solution. The CRI-PSO found the global best solution, \( M_I = 0.091 \), after 54 iterations. From Figure 3, we can find that the modified CRI-PSO method has a stronger ability to find global optimal solutions than the standard PSO. The following additional experiment also verified this conclusion.

**Experiment and Discussions**

**Experimental Data**

Parts of our test images captured by different sensors are shown in Figure 1 and Figure 4. Test images from several sensors are summarized in Table 3. We have three sets of images to be matched from different sensors: SPOT4 panchromatic (pan) and IRS-C (pan) (Dataset (A)), RadarSAT-1 C and SPOT5 multispectral (Dataset (B)), and SPOT5 (pan) at different dates (Dataset (C)). The RadarSAT-1 C image has a wavelength of 5.6 cm, HH polarization, and was acquired at an angle of incidence of 23.989°.

![Figure 3. Curves of similarity values at gbest(\(f(\text{gbest})\)).](image)

![Figure 4. Experimental datasets: Dataset (A): IRS-C pan (above) and SPOT4 pan (below); Dataset (B): RaderSAT-1 C band (above) and SPOT5 multi-spectral (below); and Dataset (C): SPOT5 pan at different dates.](image)
TABLE 3. DIFFERENT SATELLITE SENSORS AND THEIR PROPERTIES

<table>
<thead>
<tr>
<th>Sensors</th>
<th>SPOT4</th>
<th>IRS-C</th>
<th>SPOT5</th>
<th>RadarSAT-1</th>
<th>SPOT5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imaging site</td>
<td>Shaoxing, Zhejiang, China, Zhejiang, China</td>
<td>Shaoxing, Zhejiang, China, Zhejiang, China</td>
<td>Nanjing, Jiangsu, China, Nanjing, Jiangsu, China</td>
<td>Nanjing, Jiangsu, China, Nanjing, Jiangsu, China</td>
<td>Shaoxing, Zhejiang, China, Nanjing, Jiangsu, China</td>
</tr>
<tr>
<td>Band Spatial resolution (meters)</td>
<td>Pan</td>
<td>Pan</td>
<td>NIR, R, G</td>
<td>C</td>
<td>Pan</td>
</tr>
</tbody>
</table>

All image pairs are 256 × 256 pixels in dimension. In each pair, one image is used as the reference image and the other as the source image for selecting templates for matching. Several template image samples of 65 × 65 pixels were obtained by cropping the source image with an interval of 20 pixels apart from the upper left corner of each template image sample. Thus, 100 template image samples should be derived from a source image. The test areas include flat terrain with farmlands, roads, villages, towns, water areas, or city zones. The reference images and source images are approximately aligned and the pixel size of the image with the lower resolution in a pair is resampled to that of the higher resolution in that pair.

The computer configuration is as follows: CPU: Mobile Intel Pentium M 705; Main Frequency: 1.5GHz; EMS memory: 512M; Programming language: VC+ +6.0.

Optimal Choice of Parameters for PSO

In PSO, problem-based tuning of parameters is also a key factor for finding the optimal solution accurately and efficiently. Many empirical values of the parameters for PSO are investigated for different problems. Here, the parameters for PSO are set as follows:

1. \( V\text{max} \): This determines the largest distance the particle can move in each iteration. In our experiments, the range of a particle’s movement is set to the range of the template image moving in the reference image. \( V\text{max} = 256 - 65 = 191 \).
2. Learning factors \( c_1 \) and \( c_2 \): These are set to \( c_1 = c_2 = 2 \) for the standard PSO, according to Eberhart and Shi (2001) and Clerc (1999).
3. The range of the position of the particle \( X\text{min} \) and \( X\text{max} \): These refer to the range of the solution determined by the optimal problem. \( X\text{min} = 0 \) and \( X\text{max} = 191 \).
4. Dimensions of the particle: They are determined by the optimal problem. The dimensions of the particle are set to 2 because the matching is only done in 2D.
5. Stopping criterion: The stopping criterion used in the experiment is the iteration that is regarded as convergence when \( f\text{(gbest)} \) has not been improved after iteration for a given number of times. The stopping criterion is set to 30 based on the test shown in Figure 5, in which we can see that the success rate increases as the increase of the given number of consecutive iterations, and the time cost is the same. For example, when the stopping criterion is set to 30, the success rate is 93 percent, and time cost is 0.64(s). Whereas when the stopping criterion is set to 100, the success rate is 95 percent, and time cost is 1.32 [s]. We find that the success rate is increased slightly as a whole, but the time cost is increased greatly. So in the paper’s experiments, the stopping criterion is set to 30 except the experiment in Table 4.

The test data used in Figure 5 are those Stream image pairs, and 100 template image samples are used. The CRI-PSO algorithm is used and the size of swarm population is 90.

6. Size of swarm population: The population size selected is problem-dependent. Sizes of 20 to 50 are most common. In some situations, large population sizes may be used to

Table 4. The Success Rate and Time Cost for Matching with Different Size of Swarm Population

<table>
<thead>
<tr>
<th>Size of swarm population</th>
<th>Time cost (S)</th>
<th>Total iteration times</th>
<th>Success rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1.13</td>
<td>116</td>
<td>77.8</td>
</tr>
<tr>
<td>90</td>
<td>0.89</td>
<td>113</td>
<td>72.4</td>
</tr>
<tr>
<td>80</td>
<td>0.71</td>
<td>115</td>
<td>71.1</td>
</tr>
<tr>
<td>70</td>
<td>0.71</td>
<td>115</td>
<td>66.8</td>
</tr>
<tr>
<td>60</td>
<td>0.55</td>
<td>118</td>
<td>67.7</td>
</tr>
<tr>
<td>50</td>
<td>0.54</td>
<td>117</td>
<td>69.0</td>
</tr>
<tr>
<td>40</td>
<td>0.39</td>
<td>116</td>
<td>59.3</td>
</tr>
<tr>
<td>30</td>
<td>0.28</td>
<td>120</td>
<td>53.3</td>
</tr>
<tr>
<td>20</td>
<td>0.18</td>
<td>107</td>
<td>43.2</td>
</tr>
</tbody>
</table>
adapt to different requirements. The experiment in the paper also illustrates this. The relationship between the size of the swarm population and the success rate and computation is shown in Table 4. The test data used is the Stream image pair. The standard PSO is used. Here, the stopping criterion is set to 80. From it, we can see that the success rate increases as the increase of the size of swarm population as a whole, and the time cost is the same. The size of the swarm population from 50 to 100 is all acceptable. The larger the size of swarm population is, the more reliable result can be obtained. Here, it is set to 90 in order to adapt the matches for different types of images processed.

Experimental Results
Matching Results of Different PSO Variants
The PSO variants used for comparison include the following methods:

1. PSO1: Standard PSO (Shi et al., 1998). Equations 6, 7, 8, and 9 are referred to as the Standard PSO. Suppose \( c_1 = c_2 = 2 \) (Eberhart and Shi, 2001; Clerc, 1999). Linearly decreasing inertia weight \( \omega \) is set from 0.9 to 0.2, according to Shi et al. (1998 and 1999) and Wachowiak et al. (2004). Suppose \( T = 120, \) according to empirical experiment.

2. PSO2: PSO with constriction factors. A constriction factor is included in the PSO to control the convergence properties of the particles. The initial particle position is produced randomly. Suppose \( k = 1.0, c_1 = 2.8, c_2 = 1.3 \) (Wachowiak et al., 2004; Carlisle and Dozier, 2001).

3. PSO3: PSO with constriction factors and initial particle position distributed uniformly. The parameters are the same as in PSO2.

4. PSO4: Multi start PSO. The parameters are the same as in PSO1.

5. PSO5: CRI-PSO proposed in the paper. The CRI-PSO, \( c_1 \) and \( c_2 \) are determined by following equations in terms of Ratnaweera et al. (2004):

\[
\begin{align*}
    c_1 &= (c_{1f} - c_{1i}) \frac{\text{iter}}{T} + c_{1i} \\
    c_2 &= (c_{2f} - c_{2i}) \frac{\text{iter}}{T} + c_{2i}
\end{align*}
\]

where \( c_{1f} \) and \( c_{2f} \) are acceleration coefficients in the final iteration; \( c_{1i} \) and \( c_{2i} \) are acceleration coefficients in the initial iteration, \( T \) is the maximum number of iterations; iter is the current times of iterations. Generally, \( c_{1f} = 2.5, c_{1i} = 0.5, c_{2i} = 0.5, \) and \( c_{2f} = 2.5 \). The value of acceleration coefficient \( c_1 \) decreases from 2.5 to 0.5 and \( c_2 \) increases from 0.5 to 2.5. Set \( N = 45 \) validated by empirical test. Other parameters are the same as in PSO1.

Comparison results are shown in Figure 6 and Figure 7. The test data we used is the Stream image pair. The number of template images is 100. The success rate of the method, in which MI calculated with 16 bins and an exhaustive searching strategy are used, is 100 percent (exhaustive searching means all possible matching positions are searched out and compared with each other, and the most optimal solution is found; this search strategy is very slow in practice, although it produces the highest match rate). It is found that PSO1, PSO2, and PSO3 have lower success rates in our applications because the similarity measure distribution is non-convex and irregular, and it has many local maxima (shown in Figure 2). PSO4 has better performance than the three previously mentioned PSO variants. Clearly, reinitialization is effective for particles to escape from local minima. PSO5 (i.e., CRI-PSO), the method suggested here, outperforms all the other methods. In Figure 7, it can be seen that the computation cost of PSO5 is slower slightly than PSO3, but it is faster than other three methods.

Matching Experiment using Different Scenarios in SPOT4 Pan and IRS-C Pan Datasets
Different image pairs (shown in Figure 1 and Figure 4) were used in this experiment in order to test the stability and reliability of the proposed method. Performances were compared among the MI-based methods and NCC (Table 5). It is found that image matching with 16 bins for the MI is very robust, because it achieved a high success rate even though the matching images contain contrast reversals and more self-similar patterns, such as the Village6 image pair and the Farmland pair. The success rate of NCC is much lower for the Village6 and Farmland image pairs.

The CRI-PSO outperformed the Standard PSO and NCC in terms of the success rate of matching. The computation time required by CRI-PSO is similar to that of the Standard PSO, and about six times faster than that of the NCC. Compared with the MI-based method with 16 bins and the exhaustive searching strategy, the CRI-PSO has a relatively lower success rate (about 10 percent lower than that of the exhaustive search method).

Matching Experiments using Multi-sensor Dataset
In addition to the SPOT4 and IRS-C dataset, datasets using RadarSAT-1 C band and SPOT5, and SPOT5 on different dates are also tested in our experiment. The matching performance is given in Table 6.
TABLE 5. COMPARISONS OF MATCHING RESULTS OF FOUR METHODS

<table>
<thead>
<tr>
<th>Images used to trail</th>
<th>Number of template images</th>
<th>16-bin MI (Exhaustive searching)</th>
<th>16-bin MI Standard PSO (PSO1)</th>
<th>16-bin MI CRI-PSO (PSO5)</th>
<th>NCC (Exhaustive searching)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Success rate (%)</td>
<td>Time cost (s)</td>
<td>Success rate (%)</td>
<td>Time cost (s)</td>
</tr>
<tr>
<td>Town</td>
<td>100</td>
<td>95.0</td>
<td>4.52</td>
<td>55.0</td>
<td>0.64</td>
</tr>
<tr>
<td>Farmland</td>
<td>100</td>
<td>80.0</td>
<td>4.52</td>
<td>58.0</td>
<td>0.68</td>
</tr>
<tr>
<td>River1</td>
<td>100</td>
<td>89.0</td>
<td>4.52</td>
<td>73.0</td>
<td>0.62</td>
</tr>
<tr>
<td>Village</td>
<td>100</td>
<td>86.0</td>
<td>4.51</td>
<td>56.0</td>
<td>0.65</td>
</tr>
<tr>
<td>City</td>
<td>100</td>
<td>92.0</td>
<td>4.50</td>
<td>65.0</td>
<td>0.81</td>
</tr>
<tr>
<td>Village6</td>
<td>100</td>
<td>66.0</td>
<td>4.51</td>
<td>39.0</td>
<td>0.73</td>
</tr>
<tr>
<td>Stream</td>
<td>100</td>
<td>100.0</td>
<td>4.51</td>
<td>73.0</td>
<td>0.60</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>86.7</td>
<td>4.51</td>
<td>59.8</td>
<td>0.67</td>
</tr>
</tbody>
</table>

TABLE 6. COMPARISONS OF MATCHING RESULTS FOR DIFFERENT DATASETS

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Number of template images</th>
<th>16-bin MI (Exhaustive searching)</th>
<th>16-bin MI Standard PSO (PSO1)</th>
<th>16-bin MI CRI-PSO (PSO5)</th>
<th>NCC (Exhaustive searching)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Success rate (%)</td>
<td>Time cost (s)</td>
<td>Success rate (%)</td>
<td>Time cost (s)</td>
</tr>
<tr>
<td>SPOT4 pan and IRS-C pan</td>
<td>700</td>
<td>86.7</td>
<td>4.51</td>
<td>59.8</td>
<td>0.67</td>
</tr>
<tr>
<td>RadarSAT-1 C and SPOT5 multi-band</td>
<td>600</td>
<td>62.1</td>
<td>4.27</td>
<td>41.0</td>
<td>0.63</td>
</tr>
<tr>
<td>SPOT5 pan at different dates</td>
<td>400</td>
<td>100.0</td>
<td>4.31</td>
<td>78.0</td>
<td>0.62</td>
</tr>
</tbody>
</table>

From the experimental result, it is observed that the success rate of the MI with the exhaustive searching method is much higher than that of NCC, whether the matching images are captured by radar or the optical sensors. However, the time consumed by the MI methods is appreciably longer than that of NCC. For the test images from radar and optical sensors, e.g., the RadarSAT-1 and SPOT datasets, the success rate of MI is much higher than that of NCC. For images acquired from optical sensors, e.g., the SPOT and IRS-C dataset, or the SPOT5 on different dates, the success rate of MI is also higher than that of NCC. Because the time and season of the images captured by these two sensors are different, and there is a change of the landscapes, the performance of the NCC is poor, and the MI measure still maintains good performance. This illustrates that the MI is robust, reliable, and suitable for matching satellite images acquired by multiple sensors with similar resolution, such as SPOT4 pan (10 meter) and IRS-C pan (5.8 meter), and RadarSAT-1 C (12.5 meter) and SPOT5 multi-spectral (10 meter).

For the CRI-PSO proposed in the paper, its time cost for matching is improved greatly compared to that of MI with exhaustive searching and NCC. Its success rate is much better than that of NCC, but the success rate is slightly reduced compared to that of MI with exhaustive searching. From the trade-off between success rate and time cost, we can see that CRI-PSO is a promising optimizing method for fast image matching.

SIFT method (Lowe, 2004) is also tested using RadarSAT-1 C and SPOT5 images. The matching is almost all failure. The matching results of different scene for RadarSAT-1 C and SPOT5 images by CRI-PSO are shown in Figure 8 and Figure 9.

From the test results, we can see that the match rates between the optical sensors are greater than those between optical and radar sensors. Therefore, to make the mutual information criterion more useful for the application of SAR image registration, some efforts should be made to further improve the robustness of MI, for example, reducing speckle in radar images and improving interpolation algorithms for suppressing interpolation artifacts.

Conclusions

Cross-correlation based algorithms are often used in template matching in practice, but they are sensitive to nonlinear changes of the image intensity and random noise. Here, we
introduce the normalized mutual information (MI) to overcome the drawback of the cross-correlation algorithm. The validity of MI is checked by experiments using multisensor images. Because the MI-based matching algorithms are time consuming, a global optimal method is needed to accelerate the match process. Particle swarm optimization (PSO) methods can be used to accelerate the matching process. We proposed a modified PSO, named CRI-PSO, to improve convergence. We carried out experiments with various PSO in MI-based template matching. The experimental results show that the CRI-PSO can not only speed up the matching but also improve the match rates when compared to the standard PSO and NCC methods. Our results indicate that the proposed method is promising and viable for template matching. How to enhance the calculation efficiency for the MI-based template matching needs further research. There is still room to improve the match rates while overcoming the premature convergence problem of the PSO methods. For radar image matching, means to reduce speckle noise and suppress interpolation artifacts also need further study. Finally, it is desirable to develop additional measures to determine genuinely correct matching, i.e., how to discard mismatches from those matches with high similarity.

Acknowledgments
This research is supported by the State Key Laboratory of Remote Sensing Science, a grant from National Nature Science Foundation of China (No. 40771137), a grant from the National 863 Program (2008AA12Z106), partially supported by 2006103269 NNSFC, and a one hundredth talent grant from the Chinese Academy of Sciences.

The authors wish to thank three anonymous referees and editors for providing many constructive suggestions that greatly improved the quality on an earlier version of this paper.

References

Figure 9. Time cost for various methods among different scenes of RaderSAT-1 C and SPOT5.


(Received 27 January 2008; accepted 03 March 2009; final version: 15 July 2009)