Multi-objective Discrete Particle Swarm Optimization for Multi-sensor Image Alignment

J. Senthilnath, Student Member, IEEE, S.N. Omkar, V. Mani, and T. Karthikeyan

Abstract—A new technique is proposed for multi-sensor image registration by matching the features using Discrete Particle Swarm Optimization (DPSO). The feature points are first extracted from the reference and sensed image using improved Harris corner detector available in the literature. From the extracted corner points, DPSO finds the three corresponding points in the sensed and reference images using multi-objective optimization of distance and angle conditions through objective switching technique. By this, the global best matched points are obtained which are used to evaluate the affine transformation for the sensed image. The performance of the image registration is evaluated and concluded that the proposed approach is efficient.

Index Terms— Multi-sensor image registration, Multi-objective optimization, Discrete particle swarm optimization

I. INTRODUCTION

IMAGE REGISTRATION [1] is the process of geometrically aligning an image obtained from different sensor, different times or from different viewpoints with a reference image. This image which is aligned is called the sensed image and the image with respect to which alignment is carried out is called the reference image. The transformed sensed image which aligns with the reference image is called registered image.

The process of image registration is classified into two types: area based registration [1] and feature based registration [2]. Area based methods [3] are correlation or probabilistic based method. Here the common overlapping region is obtained by the traversal of window of predefined size across the sensed and reference image. The limitation of this approach is that it compares the pixel-intensities rather than analyzing the shape structure within the window [1].

Feature based methods [1, 4] requires essential features such as contours or corner points to be first extracted and mapped using suitable matching techniques. Earlier Random Sample Consensus (RANSAC) was used as a matching technique. Fischer and Bolles [5] introduced RANSAC to the location determination problem for scene analysis. However, it was found that RANSAC maps the correspondence from the outliers. Tackling this, various variants of RANSAC have been proposed. Torr et. al, used MLESAC [6] for estimating the image geometry by sampling points based on the likelihood of their belonging in the solution set. The same technique has been extended to MAPSAC [7] which takes posterior probability into account. However, this fails when the prior information about the sampling points is insignificant. Further, Genetic Algorithm (GA) has been used along with RANSAC namely GASAC [8] for parameter estimation and found to obtain better results in comparison to RANSAC. Likewise, GA [9] and particle swarm optimization [10] have been used for image matching problems.

In addition to the above problem for efficient point matching, several challenges exist for multi-sensor images. For instance, finding common region is difficult especially when the pixel intensity distribution is not same as the different sensor share different pixel information. Because of this, the feature matching is also difficult [1]. Therefore, corner points should be extracted efficiently [11] and along with this, common corner points picked between the multi-sensor images should be more so that matching corner points becomes efficient. With this, the registration process becomes more complex and less robust.

In literature RANSAC [5] and particle swarm optimization [11, 12] has been successfully applied for matching correspondence between same sensor images. In these studies, distance condition was sole fitness function for matching correspondence, but this fails incase images are of different resolution. Hence to improve this, a new approach is proposed for matching the feature points between the multi-sensor images using Discrete Particle Swarm Optimization (DPSO) based on multi-objective function. This objective function combines angle and distance condition to obtain the sets of global best matched points. The method is evaluated using quality measures and is found to be an efficient approach for multi-sensor image matching.

II. PROBLEM FORMULATION

For multi-sensor images, distance condition [12] cannot be used as the only fitness function because of its ineffective nature for different scale. Hence, additional angle condition [13] is taken into account. The combination of the two fitness functions will give better matching of the corner points in the
reference and the sensed image. Multi-sensor satellite images can be registered using the affine transformation which takes into account scaling, rotation and translation. A set of three matched points is necessary to obtain the registered image.

Let Fig. 1(a) and 1(b) be the reference and sensed image with a cluster of three corner points \( \{I, J, K\} \) and \( \{i, j, k\} \) chosen for affine transformation among the pool of corner points extracted in the reference and sensed image respectively.

\[ l_{\mu} = |x_{i} - x_{j}| \]
\[ l_{\nu} = |x_{J} - x_{K}| \]
\[ l_{\kappa} = |x_{i} - x_{k}| \]

The distance condition aims to minimize the difference of the norm of the corresponding distances in the reference and sensed image. Therefore, for a good match \( R_{y} \) must be equal to \( R_{b} \). However in practical aspects, if the deviation in the distance ratio \( \delta_{D} \) as represented in Eq. 4, lies within the threshold \( t_{l} \), then the distance condition is satisfied.

\[ \delta_{D} = |R_{y} - R_{b}| \]

\[ 0 < \delta_{D} < t_{l} \]  

The constraints involved in this case are: only three sets of points are considered at a time for the fitness function evaluation and the threshold \( t_{l} \) is user defined.

The fitness functions considered are as follows:

### A. Distance Condition

The distance between the corner points \( I, J \) and \( K \) in Fig. 1(a) is given by the norm of the difference between the two points as shown in Eq. 1.

\[ l_{\mu} = |x_{i} - x_{j}| \]
\[ l_{\kappa} = |x_{J} - x_{K}| \]
\[ l_{\nu} = |x_{i} - x_{k}| \]  

Similarly, the distances evaluated between the corner points \( i, j \) and \( k \) in Fig. 1(b) are as shown in Eq. 2.

\[ l_{\mu} = |x_{j} - y_{j}| \]
\[ l_{\kappa} = |x_{k} - y_{k}| \]
\[ l_{\nu} = |x_{i} - y_{i}| \]  

The ratio between the distances of the corresponding corner points in the reference and sensed image are obtained using Eq. 3.

\[ R_{y} = \frac{l_{\mu}}{l_{\nu}} \]
\[ R_{b} = \frac{l_{\kappa}}{l_{\nu}} \]  

\[ 0 < \delta_{D} < t_{l} \]  

The constraints involved in this case are: only three sets of points are considered at a time for the fitness function evaluation and the threshold \( t_{l} \) is user defined.

### B. Angle Condition

The angle condition aims to minimize the angle enclosed by segments of the similar points. From the Fig. 1(a), let slope of the line segment between the points \( i, j \) and \( k \) are as shown in Eq. 6.

\[ m_{ij} = \frac{y_{j} - y_{i}}{x_{j} - x_{i}} \]
\[ m_{jk} = \frac{y_{k} - y_{j}}{x_{k} - x_{j}} \]  

From co-ordinate geometry, the angle enclosed between the slopes \( m_{ij} \) and \( m_{jk} \) in the reference image is given by Eq. 7.

\[ \phi_{ref} = \tan^{-1} \left( \frac{m_{ij} - m_{jk}}{1 + m_{ij}m_{jk}} \right) \]  

Similarly, the angle enclosed in the sensed image is given by Eq. 8.

\[ \phi_{sensed} = \tan^{-1} \left( \frac{m_{ij} - m_{jk}}{1 + m_{ij}m_{jk}} \right) \]  

If the points are correspondingly matched then the difference between the angles should be less than the threshold ‘\( t_{\theta} \)’. This is illustrated in Eq. 9.

\[ |\phi_{ref} - \phi_{sensed}| < t_{\theta} \]  

This condition also has similar constraints as distance condition.

Let \( f_{1}(x) \) corresponds to the distance condition and \( f_{2}(x) \) corresponds to the angle condition. Mathematically, this optimization procedure can be formulated as

Minimize distance: \( \min \{d = f_{1}(x)\} \) subject to the constraints involved in Eq. 4

Minimize angle: \( \min \{a = f_{2}(x)\} \) subject to the constraints involved in Eq. 9

The objective functions \( f_{1}(x) \), may be conflicting with other objective function (i.e. \( f_{2}(x) \)), thus, most of the time it is impossible to obtain for all objectives the global minimum at the same point. Instead there exists a set of optimal trade-offs which forms the solution set – the Pareto Set.

### III. METHODOLOGY

This section describes the methodology involved in feature based image registration using DPSO.

Here, an improved version [11] of Harris Corner Detector (HCD) [14] is used in extracting the corner points in the reference and the sensed image.

Let \( U \) represents the set of \( m \) points detected for the reference image and \( V \) represents the set of \( n \) points for the sensed image by the feature extractor [15] as represented in Eq. 10 and Eq. 11 below.

\[ U = \{a_{1}, a_{2}, a_{3}, \ldots, a_{n}\} \]
\[ 1 \leq i \leq m \]  

\[ V = \{b_{1}, b_{2}, b_{3}, \ldots, b_{n}\} \]
\[ 1 \leq j \leq n \]  

where the co-ordinate of point \( a_{i} \) and \( b_{j} \) is given by \( (x_{ai}, y_{ai}) \) and \( (x_{bj}, y_{bj}) \) in their respective co-ordinate system.

The corner points of the sensed image are matched with the corner points of the reference image using DPSO.

Instead of using the brute force method [15] of finding the corresponding corner points in the sensed and reference image, DPSO optimizes the search for the sets of global best matched points based on the distance and angle conditions utilizing the objective switching technique.

It treats the points as discrete entities in a vector which are labeled with an index using Eq. 10 and 11. The position vector of a particle is represented by 6 strings. The first three elements represent the index of the corner points chosen from the reference image while the latter 3 represent the index of...
the corner points chosen from the sensed image. A typical position vector is shown in Eq. 12 below:

\[ P = \{a_i, a_j, a_k, b_i, b_j, b_k\} \]  

(12)

The aim is to find three pairs of reference and its corresponding sensed points to obtain the best affine transformation for image registration by optimizing the fitness functions - distance and angle condition.

1) Initialization

The position vector of the \(i^{th}\) particle \(X(i)\) is represented using Eq. 12. The fitness value of this particle is evaluated using Eq. 4 and 9 by utilizing the objective switching technique. Further, the previous best vector \(B(i)\) which represents the best position vector of the \(i^{th}\) particles of the current iteration, is also initialized same as \(X(i)\). The fitness value corresponding to the previous best vector \(f(B(i))\) is also initialized same as \(f(X(i))\). This process is carried out for all the other particles in the population. The global best vector, representing the best fit particle is initialized as null.

2) Particle Update

The position vector of the particle \(X(i)\) is updated using velocity vector. Velocity vector is of dimension 2 \(X N\). The first row contains the weight associated with a particular attribute in the same column. The second row contains the indices of the corner points. Initially all the weights are defined as 1. If the total number of points is \(N\), then the velocity vector for \(i^{th}\) particle is defined as:

\[ v(i) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & N \end{bmatrix} \]  

(13)

The weights are changed if the attributes present in the same column as that of the weight is also present in the global best position vector \((G)\) or particle’s position vector \((X(i))\) or the particle’s previous best position vector \((B(i))\). Three positive constants \(\alpha\), \(\beta\) and \(\gamma\) are used to designate the contribution of the \(X(i)\), \(B(i)\) and \(G\) which are used for updating the weights of the velocity vector. For all those attributes present in velocity vector which are also present in \(X(i)\), the weight is increased by \(\alpha\), likewise \(\beta\) for \(B(i)\) and \(\gamma\) for \(G\). The velocity vector is updated using Eq. 14.

\[ v(i) = \begin{bmatrix} w_1 & w_2 & \cdots & w_N \\ 1 & 2 & \cdots & N \end{bmatrix} \]  

(14)

Where

\[ w_j = 1 + \alpha \delta(\alpha) + \beta \delta(\beta) + \gamma \delta(\gamma) \]  

(15)

\[ \delta(\alpha) = \begin{cases} 1 & \text{if } v(2,j) \in X(i) \\ 0 & \text{otherwise} \end{cases} \]

\[ \delta(\beta) = \begin{cases} 1 & \text{if } v(2,j) \in B(i) \\ 0 & \text{otherwise} \end{cases} \]

\[ \delta(\gamma) = \begin{cases} 1 & \text{if } v(2,j) \in G \\ 0 & \text{otherwise} \end{cases} \]

After this, each weight is randomly multiplied by a number between 0 and 1. The vector \(v(i)\) is rearranged column-wise in decreasing order of the cumulative weights. The first three indices in the second row will give the cluster of 3 reference points to be considered for the \(i^{th}\) particle. After the corner points in the reference image are found, the corresponding corner points in the sensed image are found using switching technique in multi-objective optimization of the distance and angle condition as discussed in previous section. The corner points in the sensed image which satisfy the fitness functions within the threshold limits as represented in Eq. 4 and Eq. 9 for the particular corner points chosen in the reference image using Eq. 13 - 15 are chosen as the best candidate for the position vector of the particle as shown in Eq. 12. The fitness value of the new position vector of the \(i^{th}\) particle is compared with its previous best and global best and is updated if it is found to be better than the latter.

3) Termination

The process terminates at the end of \(M^{th}\) iteration and the global best vector, \(G\) which contains the indices of the corner points in the reference image and its corresponding points in the sensed image which satisfies the fitness functions the most, is used in evaluating the affine transformation for the sensed image for it to align with the reference image.

IV. PERFORMANCE MEASURE

The performance of the proposed technique is evaluated using two measures namely feature correspondence and RMSE [16].

A. Feature correspondence accuracy

The feature matching accuracy is measured by checking how many times the DPSO matches the features correctly in the reference and the sensed image, out of the total number of matched points satisfying the multi-objective function. To test the novelty of the algorithm, it is run for several times. The accuracy of matching is as shown in Eq. 16.

\[ A = \frac{N_c}{N_s \times N_r} \]  

(16)

where, \(N_c\) is the total number of correct feature matches in all runs, \(N_s\) is the number of matches satisfying the multi-objective fitness conditions in each run and \(N_r\) is the number of runs.

B. Root Mean Square Error

Root Mean Square Error (RMSE) will be closer to zero when the ground truth and automatic aligned image are similar and will increase when the dissimilarity increases. RMSE is defined as:

\[ RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} E_k^2} \]  

(13)

where \(E_k\) is the difference between pixel values of the ground truth image and the automatic aligned image. \(N\) signifies the total number of pixels in the image.

This error value was computed to assess the image alignment of multi-sensor satellite images. The RMSE was calculated with different sets of manually aligned images and automatically registered image and the average RMSE is calculated. This is done in order to minimize the error in manually aligned image which is used as ground truth image.

V. RESULTS AND DISCUSSION

In this section, results from two multi-sensor satellite images for image registration are presented using DPSO based multi-
objective optimization. Further, using the performance evaluations, this method is compared with RANSAC incorporating only distance condition as its objective function.

**Image set 1**: location, Ulsoor Lake, Bangalore, India. The reference image is a QuickBird multi-spectral (MS) of resolution 2.4 m and the sensed image is QuickBird panchromatic (Pan), resolution 0.61 m.

**Image set 2**: The reference image is same as considered in Image set 1 while the sensed image is Linear Imaging Self-Scanning Sensor – 4 (LISS-4), resolution 5.8 m.

### A. Image set 1
The corner points are extracted using modified HCD. The corner points detected in the reference and sensed image were to be 40 and 62 respectively. The population size for the DPSO was fixed to 10 particles each containing cluster of 3 points from the reference and sensed image as shown in Eq. 12. Table I illustrates the number of matched points obtained by the selection of the different DPSO parameter values. In this table we can observe that the number of matched points varies for different parameter value. The best number of matched points is obtained by setting the parameter value of $\alpha$, $\beta$, and $\gamma$ to 0.6, 0.8 and 1 respectively.

### TABLE I. DPSO parameter analysis

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>Matched points</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0.8</td>
<td>0.6</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>2</td>
</tr>
</tbody>
</table>

### B. Image set 2
For this image set, a total of 40 corner points in the reference image and 25 points in the sensed image were extracted. The same analysis is done for the selection of parameters, thresholds and termination conditions as for the Image set 1. Ten trials ($N_r = 10$) were carried out to find the set of 3 matched points in the sensed and reference image ($N_s = 3$). Out of these, a set of 15 matches ($N_c$) were found to be correct. Hence, the feature correspondence was found to be 0.5 where as RANSAC fails to pick 3 matched points. The RMSE value...
for the registered image using RANSAC was found to be 0.81 while the value obtained using DPSO was found to be equal to 0.29. Hence the multi objective optimization using DPSO picks better matched points between the sensed and the reference image. The matched points and the final registered image are shown in Fig. 4 and Fig. 5 respectively.

![Fig. 5: Registered Image](image)

### VI. CONCLUSION

In this study, a new multi-objective optimization of the angle and distance conditions using DPSO based on switching technique is proposed for feature matching. It is found that the proposed technique is able to register the sensed image with the reference image by matching the corner points obtained from the modified Harris Corner Detector. There are few restrictions in DPSO. The initial population is random and the parameters are set empirically for better match. From the result obtained using both the case study, DPSO is found to be better than RANSAC.

In the past, RANSAC and its variants were applied to match the points. Based on the results obtained, it clearly indicates that the proposed method is more efficient than RANSAC for multi-sensor image registration as it incorporates angle condition besides the distance condition for feature matching.

### ACKNOWLEDGMENT

The authors would like to thank the associate editor and reviewers for their comments that helped in improving this paper.

### REFERENCES


