AUTOMATED IMAGE REGISTRATION USING GEOMETRICALLY INVARIANT PARAMETER SPACE CLUSTERING (GIPSC)

Gamal Seedahmed and Lou Martucci

Remote Sensing and Electro-Optics Group
Engineering Physics Division
Pacific Northwest National Laboratory*
902 Battelle Blvd. Richland, WA, 99352 USA
Gamal.Seedahmed@pnl.gov
Lou.Martucci@pnl.gov

Commission III, WG III/I

KEY WORDS: Automation, Image Registration, Hough Transform, Geometric Invariance, Clustering

ABSTRACT:

Accurate, robust, and automatic image registration is a critical task in many typical applications that employ multi-sensor and/or multi-date imagery information. In this paper we present a new approach to automatic image registration, which obviates the need for feature matching and solves for the registration parameters in a Hough-like approach. The basic idea underpinning GIPSC methodology is to pair each data element belonging to two overlapping images, with all other data in each image, through a mathematical transformation. The results of pairing are encoded and exploited in histogram-like arrays as clusters of votes. Geometrically invariant features are adopted in this approach to reduce the computational complexity generated by the high dimensionality of the mathematical transformation. In this way, the problem of image registration is characterized, not by spatial or radiometric properties, but by the mathematical transformation that describes the geometrical relationship between the two images or more. While this approach does not require feature matching, it does permit recovery of matched features (e.g., points) as a useful by-product. The developed methodology incorporates uncertainty modeling using a least squares solution. Successful and promising experimental results of multi-date automatic image registration are reported in this paper.

1. INTRODUCTION

The goal of image registration is to geometrically align two or more images so that respective pixels or their derivatives (edges, corner points, etc) representing the same underlying structure (object space) may be integrated or fused. In some applications image registration is the final goal (interactive remote sensing, medical imaging, etc) and in others it is a required link to accomplish high-level tasks (multi-sensors fusion, surface reconstruction, etc). In a multi-sensor context, registration is a critical starting point to combine multiple attributes and evidence from multiple sensors. In turn, multi-sensors registration or fusion can be used to assess the meaning of the entire scene at the highest level of abstraction and/or to characterize individual items, events (e.g. motion), and other types of data.

The sequential steps of feature extraction, feature matching, and geometric transformation have evolved into a general paradigm for automatic image registration, (see Brown, 1992). Many algorithms have been invented around this paradigm to handle the automatic image registration with a major focus on solving the matching (correspondence) problem. The basic idea behind most of these algorithms is to match image features according to their radiometric or geometric properties using a pre-specified cost function to assess the quality of the match; (see Dare and Dowman, 2001; Thepaut et al., 2000; Hsieh et al., 1997; Li et al., 1995; Wolfson, 1990). While these methods have certain advantages in computing the transformation parameters in a single step and in retaining the traditional way of thinking about registration in the sense of identifying similar features first and then computing the parameters of the geometric transformation, they have considerable drawbacks in meeting the current challenges of image registration. First of all, they require feature matching, which is difficult to achieve in a multi-sensor context since the common information, which is the basis of registration, may manifest itself in a very different way in each image. This is because different sensors record different phenomena in the object scene. For instance, take a radar image vs. optical image. Second, the feature extraction algorithms are far inferior in the sense of detecting complete image features. For instance, missing information such as edge gaps, and occlusion are two famous examples that could lead to incorrect matching

In the late nineties and through 2001, Hough Transform (HT)-like approaches emerged as a powerful class of registration methods for image and non-image data. This new class of methods provides a remedy to the above-mentioned problems and considers different strategies to reduce the computational complexity that hampered the wide use of the original HT (Hough, 1962). In comparison to the previous approaches, this new class is a correspondence-less strategy since it does not use feature correspondence to recover the transformation. Instead, a search is conducted in the space of possible transformations. The Modified Iterative Hough Transform (MIHT) is a representative method that belongs to Hough-like approaches. MIHT is developed to solve automatically for different tasks such as single photo-resection, relative orientation, and surface matching, (see Habib and Schenk, 1999; Habib et al., 2000; Habib and Kelly, 2001[a,b]; Habib et al., 2001). In MIHT, an ordered sequential recovery of the registration parameters is adopted as a strategy to reduce the computational complexity. This ordered sequential solution considers quasi- invariant parameters to reduce the computational complexity. These parameters are associated either with specific locations that de-correlate them or with the selection of data elements that contribute to specific parameter(s). The basic idea behind the Hough-like approaches, such as MIHT, is the exploitation of the duality between the observation space and the

* Operated by Battelle for the US Dept. of Energy (DOE)

parameter space. Similar registration problems can be formulated as Hough-like approaches, but the computational complexity will remain an overwhelming challenge.

In this paper we present a new approach to automatic image registration, which is compatible with the general ideas of Hough-like approaches, but differs in the way of how the computational complexity is handled and the application task. This approach exploits the duality between the observation space and the parameter space in the sense of HT. In this approach, as well as the Hough-like ones, the problem of image registration is characterized, not by the geometric or radiometric properties, but by the mathematical transformation that describes the geometrical relationship between two images. The proposed approach considers different strategy to reduce the computational complexity, and is tailored to handle 2-D registration, which is a typical case in most of remote sensing imagery. The basic idea underpinning the proposed approach is to pair each data element belonging to two sets of imagery, with all other data in the set, through a mathematical transformation that describes the geometrical relationship between them. The results of pairing are encoded and exploited in histogram-like arrays (parameter space) as clusters of votes. Binning in the specified range of the registration parameters generates these clusters. The process of using geometrically invariant features is considered as a strategy to reduce the computational complexity generated by the high dimensionality of the mathematical transformation. This approach does not require feature matching. Matched features will be recovered as a by-product of this approach. The developed approach is accommodated with full uncertainty modeling and analysis using a least squares solution.

This paper is organized as follows. Section 2 presents the proposed methodology, section 3 presents the experimental results, section 4 discusses the obtained results, and finally section 5 concludes the paper.

2. METHODOLOGY

The basic idea underpinning the proposed approach is to compare common data elements of two images with all other data contained in those images through a mathematical transformation that describes the geometrical relationship between them. This approach considers two basic assumptions. First, the characteristics of the object space give rise to detectable features such as points and lines in both images, and at least part of these features are common to both images. Second, the two images can be aligned by a 2-D transformation. The basic process starts with feature extraction, followed by geometric invariant features construction, and then parameter space clustering. For the interest of developing an intuitive understanding of the basic process, each step is highlighted briefly, while a through discussion is deferred to the subsections below. First, in the presented study point features are dealt with. Second, in order to construct geometric invariant features, each point in the first image is related to a collection of other points in the same image defining a geometric arrangement whose properties remain invariant under a chosen transformation. The same process is applied to the second image. By constructing geometric invariant features, we did not impose any geometric constraint on the original image features such as straightness. However, the geometric properties of the mathematical transformation are considered. Invariant features, constructed from the two images to be registered, characterize these properties. Invariant features gave rise to a set of mathematical transformations with a reduced dimensionality. This set of mathematical transformations was used as voting (clustering) functions in the parameter space. Third, the basic idea of parameter space clustering is to compare the data element gathered from two sets according to a pre-specified observation equation (voting function). The results of comparison will point to different locations in the parameter space. The pointing is achieved by incrementing each admissible location by one during the voting process. A coexisting location in the parameter space, defined by the data elements that satisfy the observation equation, will be incremented several times forming a global maximum in the parameter space. This maximum will be evaluated as a consistency measure between the two data sets.

In the sequel of the three subsections below, we present a detailed derivation of geometric invariant features, the principle of parameter space clustering and least squares solution.

2.1 Construction of Geometric Invariant Features

In general, geometric invariants can be defined as properties (functions) of geometric configurations that do not change under a certain class of transformations (Mundy and Zisserman, 1992). For instance, the length of a line does not change under rigid motion such as translations and rotation. In this subsection geometric invariance will be developed for point sets under similarity transformation. Assume that we have two point sets, *P* and *Q*, extracted from two images, where $P = \{(x_i, y_i)^T \mid i = 1,...,m\}$ and $Q = \{(x_j, y_j)^T \mid j = 1,...,n\}$. A registration is to find a correspondence between a point p_i in *P* and a certain point q_j in *Q*; that makes this corresponding pair consistent under a selected mathematical transformation. The similarity transformation, $f(T_x, T_y, s, \theta)$, is used as registration function between the two sets, where T_x, T_y are the translation along the *x* and *y*-axes, *s* is the scale factor, and θ is the rotation angle between the two images. Let (p_{i1}, p_{i2}) and (q_{j1}, q_{j2}) be two corresponding pairs in *P* and *Q* respectively. Geometric invariant quantities under the similarity transformation can be derived as follows:

$$\begin{bmatrix} x_{j1} \\ y_{j1} \end{bmatrix} = \begin{bmatrix} T_x \\ T_y \end{bmatrix} + s \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_{i1} \\ y_{i1} \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} x_{j2} \\ y_{j2} \end{bmatrix} = \begin{bmatrix} T_x \\ T_y \end{bmatrix} + s \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_{i2} \\ y_{i2} \end{bmatrix} \tag{2}$$

By deriving the vector quantities between (p_{i1}, p_{i2}) and (q_{j1}, q_{j2}) we will end up with:

$$\begin{bmatrix} x_{j2} - x_{j1} \\ y_{j2} - y_{j1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + s \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{R} \begin{bmatrix} x_{i2} - x_{i1} \\ y_{i2} - y_{i1} \end{bmatrix} \tag{3}$$

where R is the rotation matrix. By computing the vector quantities we derived translation invariant geometric primitives (vectors).

Let

$v_1 = \begin{bmatrix} x_{j2} - x_{j1} \\ y_{j2} - y_{j1} \end{bmatrix}$ and $v_2 = \begin{bmatrix} x_{i2} - x_{i1} \\ y_{i2} - y_{i1} \end{bmatrix}$. By computing the squared norm of equation (3) we will end up with the following equation:

$$\|v_1\|^2 = s^2 \underbrace{(R^T R)}_{Identity} \|v_2\|^2 \tag{4}$$

$R^T R = I$, since R is an orthogonal matrix. By computing vector norms, one can derive rotation invariant geometric primitives. One can conclude from equation (4) that the scale factor can be determined by the following equation:

$$s = \frac{\|v_1\|}{\|v_2\|} \tag{5}$$

The rotation angle (θ) between the two images will be recovered as a difference between the directions of the normal to each vector. The directions of the normals can be obtained by the following equations:

$$\psi_1 = \cos^{-1}\left(\frac{x_{i2} - x_{i1}}{\|v_1\|}\right) \tag{6}$$

$$\psi_2 = \cos^{-1}\left(\frac{x_{j2} - x_{j1}}{\|v_2\|}\right) \tag{7}$$

$$\theta = \psi_2 - \psi_1 , \tag{8}$$

where ψ_1 and ψ_2 are the directions of the normals to the two vectors. The direction of the normal to the vector is preferred over the direction of the vector to avoid the difficulties in vertical vectors. After solving for the scale factor and the rotation angle using equations (5) and (8) the translations along the x- and-y axes can be solved for by:

$$T_x = x_{j1} - (s(\cos\theta)x_{i1} - s(\sin\theta)y_{i1}) \tag{9}$$

$$T_y = y_{j1} - (s(\sin\theta)x_{i1} + s(\cos\theta)y_{i1}) \tag{10}$$

2.2 Parameter Space Clustering

As was mentioned above, the basic idea behind this approach is a pairing process between two data sets. In statistical sense, the pairing process is nothing more than a determination of a parameter distribution function of the specified unknown parameter(s). Equations (5), (8), (9), and (10) were used to pair the extracted features from the first and the second image. For instance, equation (5) was used to pair the derived norms from the first and the second data sets to recover the parameter distribution function of the scale factor. By the same token, equations (8), (9), and (10) were used to recover the parameter distribution functions of the rotation and translations parameters respectively. The results of pairing can be encoded in the parameter space, and is implemented by an accumulator array. The correct pairs will generate a peak in the parameter space. This peak will be evaluated as a consistency measure between the two images to be registered. Incorrect pairing will give rise to non-peaked clusters in the parameter space. In this method and other ones, which adopt similar approaches to match data sets, the admissible range of the transformation parameters, encoded in the parameter space, defines a probability distribution function, as indicated previously. Then, the best transformation parameters are estimated by the mode, that is, by the maximum value (the peak). It is well known that the mode is a robust estimator (Rosseeuw and Leroy, 1987) since it is not biased to outliers. In automatic image registration, outliers correspond to transformation parameters originate by matching some image features to noise or to some features that do not exist in the other image. Hence, we can conclude that parameter space clustering should be capable of handling incorrect matches in a way that do not affect the expected solution.

2.3 Least Squares Solution

In order to propagate the accuracy of the extracted feature (points) into the registration parameters in an optimal way, a least squares solution was used. Equations (11) and (12) below describe the similarity transformation with the uncertainty associated with extracted points.

$$x_{j1} - e_{xj1} = T_x + (s(\cos\theta) - s(\sin\theta))\begin{bmatrix} x_{i1} - e_{yi1} \\ y_{i1} - e_{yi1} \end{bmatrix} \tag{11}$$

$$y_{j1} - e_{yj1} = T_y + (s(\sin\theta) + s(\cos\theta))\begin{bmatrix} x_{i1} - e_{xi1} \\ y_{i1} - e_{yi1} \end{bmatrix} \tag{12}$$

$$\begin{bmatrix} e_{xi1} \\ e_{yi1} \\ e_{xj1} \\ e_{yj1} \end{bmatrix} \sim (0, \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix})$$

e: is the true error associated with each coordinate, ~: stands for the normal distribution and Σ_1, Σ_2 are the variance-covariance matrices associated with each data set. We assume that the two data sets are stochastically independent.

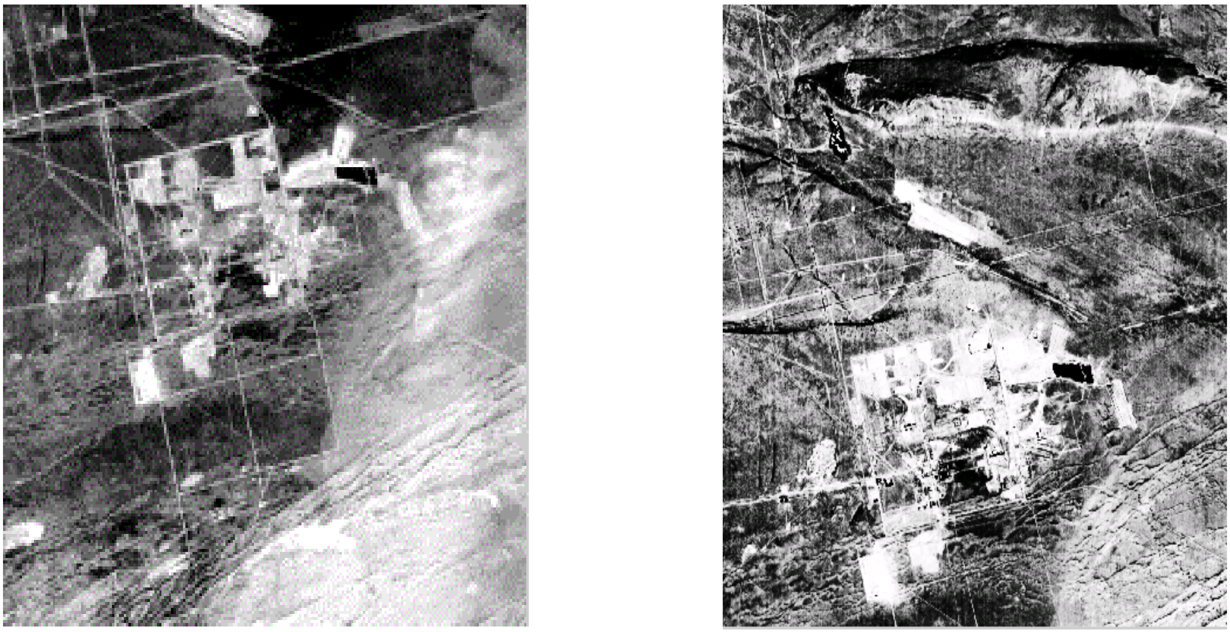
The proper stochastic model of equations (11) and (12) is the condition equations with parameters (see Schaffrin, 1997), which can be stated as follows:

$$bY = A\Xi + be \tag{13}$$

b: is the partial derivatives with respect to the observation (extracted features), A: is the partial derivatives with respect to the registration parameters, Ξ : is the correction values to the registration parameters, and e: is the true error.

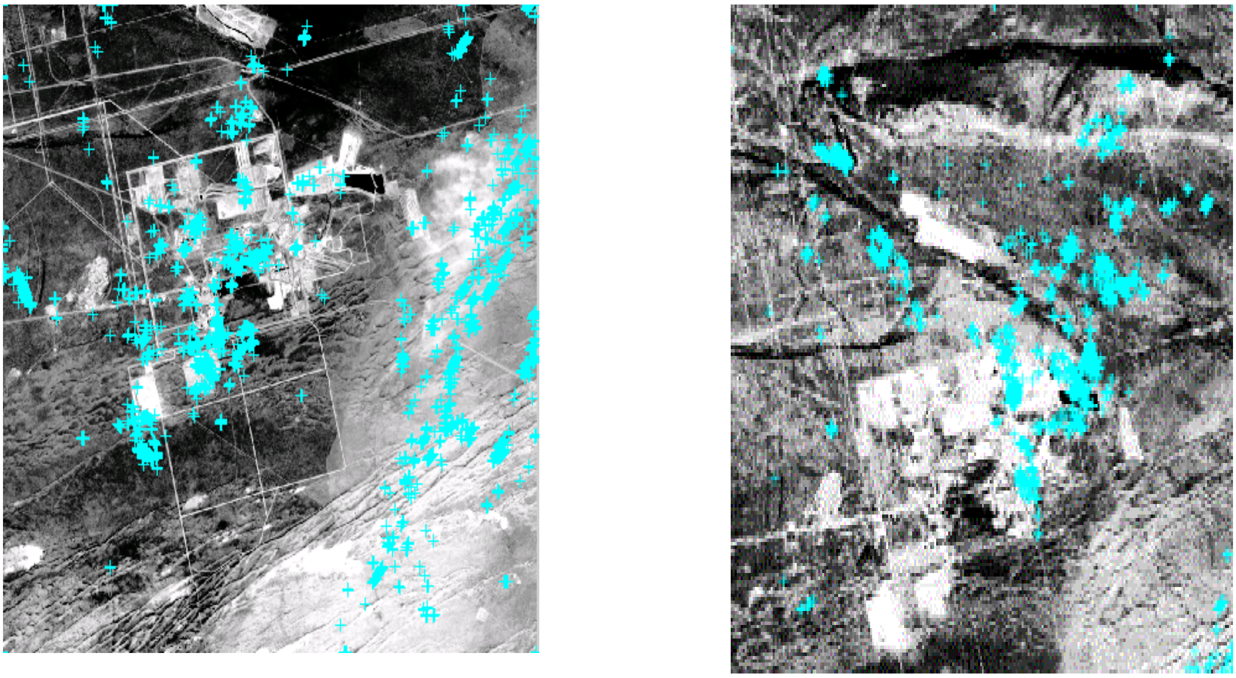
3. EXPERIMENTAL RESULTS

This section presents a complete experiment of a typical example of workflow of automatic image registration using GIPSC. Two subimages of SPOT scenes of size 1024×1024 , were used in this experiment, see Fig. (1). These subimages shared a common overlap area and separated by a time difference of four years. The two images were corrected up level 1A. In order to remove the random noise, the two subimages were convolved by a moving average filter. The process started by point features extraction using Moravec operator (Moravec, 1977), see Fig. (2), and then followed by constructing geometric invariant features as described in section 2. The two image features were paired according to equations (5), (8), (9), and (10). The results of pairing were encoded in the relevant parameter space as depicted in Fig. (3). The expected registration parameters were recovered by searching for the peak value in the parameter space. The locus of the peak indicates the values of the registration parameters and its peak height indicates the number of matched points. Matched points were recovered by backtracking the process, as show in Fig. (4). Table (1) shows the number of detected and matched points between the two images. The matched points are combined in a single least squares adjustment, and Table (2) shows the results. The adjusted parameters were used to resample the second image (SPOT 1991) to the space of the first image (SPOT 1987) and Fig. (5) shows the results of resampling as image mosaic. Bilinear transformation is used as an interpolation method in the resampling process.



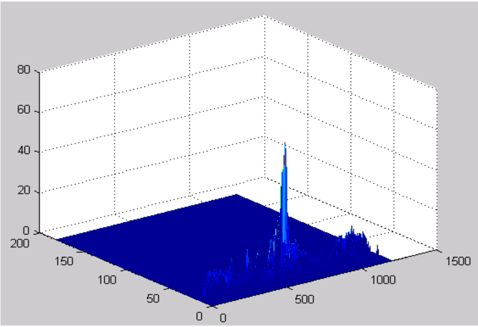
SPOT 1987 SPOT 1991

Figure 1: Two SPOT subimages, taken at different time (1987 and 1991), over the Hanford Reservation in Washington State, USA.

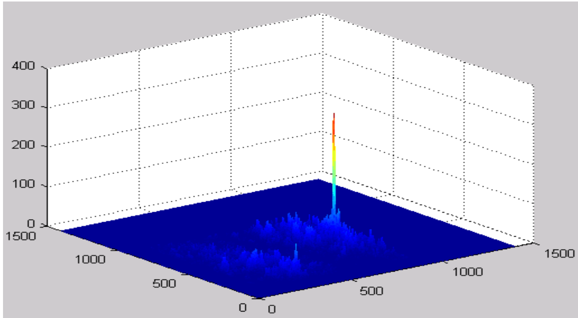


SPOT 1987 SPOT 1991

Figure 2: Shows the results of point features extraction using Moravec operator.



3.a



3.b

Figure 3.a: The results of the parameter space clustering with respect to the scale and rotation angle. Figure 3.b: shows the results of the parameter space clustering with respect to the translations along the x and y-axes. The emerged peaks in both figures point to the locus of the expect solution that can align the two images. The grids of the two plots are based on the matrix indices.

Figure 4: The automatically matched points overlaid over their original subimages. The automatically matched points are used as a basis for precise registration using least squares solution.

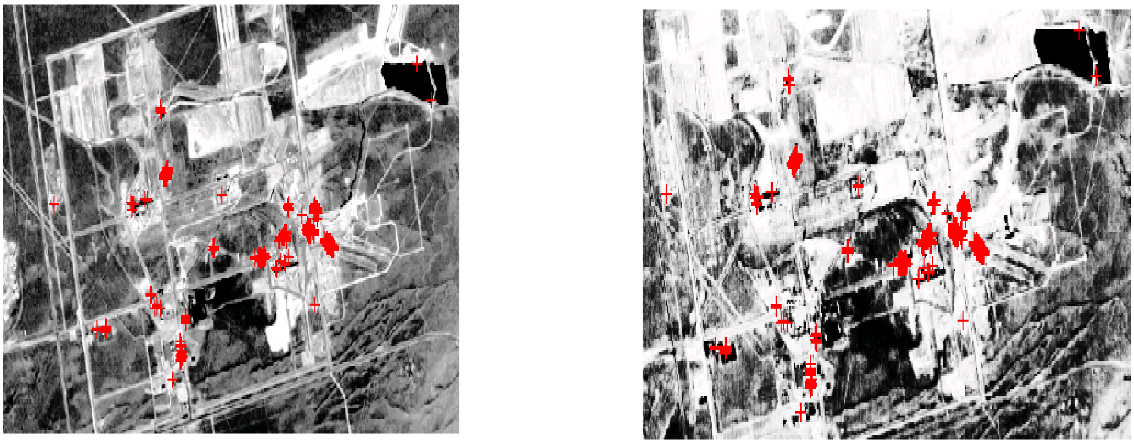


Figure 4: shows the automatically matched points overlaid over their original subimages. The automatically matched points are used as a basis for precise registration using least squares solution.

Point description	Number of points
Detected points in image (1987)	1962
Detected points in image (1991)	1932
Matched Points	328

Table 1: The number of the detected and matched points.

Parameter	Value	Std Dev.
X-translation	35.33 pixels	± 0.0917 Pixel
Y-translation	330.5 pixels	± 0.0917 Pixel
Scale	0.9768	$\pm 10^{-4}$
Rotation	-0.0023deg	$\pm 1.74 \times 10^{-6}$ deg.

Table 2: The registration parameters and their standard deviations.

4. DISCUSSION

The developed approach, detailed in this paper, successfully registers the two images, as shown in Fig. (5) . The correct matches define a peak in the parameter space, (see Fig. (3)). Incorrect matches define non-peaked clusters. It is evident from table (1) that this approach is highly robust , since the percentage of the matched point compared to the number of the detected points in each image is very small (<16%). In other words, this approach is able to handle more than 84% of incorrect match (outliers). The results of the least squares solution, presented in table (2), give important information about the final accuracy of the registration, which is about $1/10^{th}$ of the pixel size in the x and y directions. It is interesting to note that the accuracy of feature extraction is around ± 1 pixel. This excellent subpixel registration accuracy, in the final localization, is obtained because all of the points that have been identified as a corresponding pairs (328 points) are used in the final adjustment.

5. CONCLUSIONS

In this paper we have developed a formulation and methodology to handle the problem of image registration in an autonomous, robust and accurate manner. In this method the problem of image registration is characterized, not by the geometric or radiometric properties, but by the mathematical transformation that describes the geometrical relationship between two images. In other words, this approach does not require feature matching. Instead, a search is conducted in the space of admissible transformation. Geometrically invariant features are adopted to decompose the computational complexity of the transformation. This approach solves simultaneously for the registration parameters and the matched features.

This approach is highly robust as compared to the traditional M-estimators (Rosseeuw and Leroy, 1987), which tolerates only up to 50% of outliers. Combining the developed approach with the least squares estimator facilitate the achievement of subpixel accuracy in the final registration parameters. Research effort is underway to characterize performance metrics and pathological cases, in order to extend this approach in its methodology and applications.

1987-1991

Figure 5: Shows the results of resampling using the registration parameters.

6. ACKNOWLEDGEMENTS

We would like to express our great appreciation to the Pacific Northwest National Laboratory for the full financial support under project No. 43429.

7. REFERENCES

Brown, L., (1992). A Survey of Image Registration Techniques. *ACM Comput. Surv.* 24, No. 4, pp. 325-376.

Dare, P., I., Dowman, (2001). An Improved Model for Automatic Feature-Based Registration of SAR and SPOT Images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 56 (2001) pp. 13-28.

Habib, A., A., Asmamaw, D., Kelly, (2000). New Approach to Solve the Matching Problem in Photogrammetry. XIXth ISPRS Congress, Amsterdam, The Netherlands (16-23 July 2000).

Habib, A., D., Kelley. 2001^{a}. Single Photo Resection Using Modified Iterative Hough Transform. *Journal of Photogrammetric Engineering and Remote Sensing*, Vol. 67, August 2001, pp. 909-914.

Habib, A., D., Kelley. 2001^{b}. Automatic Relative orientation of Large Scale Imagery over Urban Areas Using Modified Iterative Hough Transform. *International Journal of Photogrammetry and Remote Sensing*, 56 (2001), pp. 29-41.

Habib, A., T., Schenk, (1999). New Approach for Matching Surfaces from Laser Scanners and Optical Sensors. Joint Workshop of ISPRS III/5 and III/2 on Mapping Surface Structure and Topography by Air-borne and Space-borne Lasers, La Jolla, San Diego, CA (November 9-11 1999).

Habib, A., Y., Lee, M. Morgan, (2001). Surface Matching and Change Detection Using Modified Iterative Hough Transform for Robust Parameter Estimation. *Photogrammetric Record*, 17 (98), October 2001, pp. 303-315.

Hough, P., (1962). Methods and Means for Recognizing Complex Pattern, US Patent 3,069,654.

Hsieh J., H. Liao, K. Fan, M. Ko, Y. Hung, (1997). Image Registration Using a New Edge-Based Approach, *Computer Vision and Image Understanding,* Vol. 67,No. 2, August, pp. 112-130.

Li, H., B., Manjunath, S., Mitra, (1995). A Contour-Based Approach to Multi-sensor Image Registration. *IEEE Transaction on Image Processing*, Vol. 4, No. 3, March 1995, pp. 320-334.

Moravec, H., (1977). Towards Automatic Visual Obstacle Avoidance. Proc. 5th Int. Joint Conference on Artificial Intelligence, MIT, Cambridge, MA, USA, p. 584.

Mundy, J., A., Zisserman, (1992). Towards a New Framework for Vision. *Geometric Invariance in Computer Vision* (Mundy and Zisserman, editors), The MIT Press.

Rousseew, P., A., Leroy, (1987). *Robust Regression and Outlier Detection.* Wiley, New York.

Schaffrin, B., (1997). Class Notes on Adjustment computations. The Ohio State University.

Thepaut, O., K., Kpalma, J., Roseph, (2000). Automatic Registration of ERS and SPOT Multisensor Images in a Data Fusion Context. *Forest Ecology and Management*, 128 (2000), pp. 93-100.

Wolfson, H., (1990). On Curve Matching. *IEEE Transaction on Pattern Recognition and Machine Intelligence*, Vol. 12, No. 5, May 1990, pp. 483-489.