

A preprocessing and automated algorithm selection system for image registration

A. L. Drozd, A. C. Blackburn, I. P. Kasperovich, ANDRO Computational Solutions;
P. K. Varshney, M. Xu, B. Kumar: Syracuse University

ABSTRACT

Image registration is a technique for precisely aligning the content of two or more images. It is often used as a pre-processing stage for further analysis, such as automatic target recognition, change detection, and environmental remote sensing. However, there are many different registration algorithms available to the image analyst, and it's difficult to know which one is the best one to use for a particular pair of images. These various algorithms also have a multitude of settings and parameters that must be given proper values for best results. Consequently, it is often difficult to know which algorithm will perform the best in a given situation, under constraints of time or accuracy. We propose constructing an expert system, with rules based on experimental results, that will automatically select the appropriate registration algorithm and perform appropriate preprocessing steps to prepare the images for registration.

Keywords: Image Registration, Data Fusion, Image Processing, Expert Systems

BACKGROUND

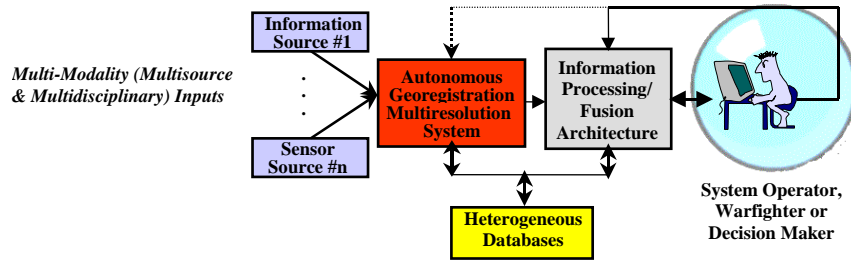
Accurate and automated geo-registration is of much interest to the US Air Force as well as other military organizations. A wealth of sensor information collected or derived from both imaging and non-imaging sensors comes in various forms. These can include: synthetic aperture radar (SAR) imagery, electro-optical/infrared (EO/IR) imagery, multispectral imagery (MSI), hyperspectral imagery (HSI), measurement and signal intelligence (MASINT), human intelligence (HUMINT), target status (detection, location, classification, identification, etc.) data confidence levels, sensor platform position, SIGINT, GMTI, and so on. We generally classify such *multi-modality* sensor data into two main categories: image intelligence (IMINT) and non-image intelligence (non-IMINT). Taken together, these can be referred to as Multi-INT. The wealth of Multi-INT sensor data is creating information overload issues. Exacerbating this is that current auto-registration technologies and techniques may not produce accurate or precise results especially for multi-modality sensor inputs. Further, such auto-registration techniques may not run in a time-efficient manner to support real time or near-real time information fusion and downstream targeting tasks. In time-critical situations, the amount of information and how it is processed can adversely affect the decision process.

SPATIAL MULTISENSOR AUTONOMOUS REGISTRATION TOOLKIT (SMART)

The overarching objective of our effort is to develop and integrate a complete suite of data registration algorithms in the form of a software toolkit for the spatial alignment of Multi-INT (image, SIGINT, GMTI and other) data collected from the battlespace. This will be accomplished by leveraging and further extending the results of our basic research and exploratory development. We will continue to diversify and enhance the initial proof-of-concept capability to support additional functionality, knowledge, information, sensors, and sources. We will continue to build upon our recent findings and current research to further identify potential problem areas and refine solutions associated with the selected registration algorithms and their use in the integrated toolkit framework. This includes a fleshing out of the expert system rule base to provide a complete, robust capability to perform Multi-INT data preprocessing and for brokering the selection of algorithms to ensure that accurate results are obtained efficiently.

To address the issues of registration performance, and overcome the limitations of individual registration algorithms, we are proposing a rule-based expert system known as the Spatial Multisensor Autonomous Registration Toolkit, or SMART.

**Spatial Multi-sensor Autonomous Registration Toolkit (SMART)
Conceptual Diagram**



SMART technologies will directly benefit a broad range of applications. The military application is the increased understanding of the air and ground stationary and moving targets in the battlespace environment. For example, *SMART* could be used to aid command-level decision tasks by producing reliable and consistent information that reduces target uncertainty. Once the registration capability (as part of an overall fusion and signal processing system) has been exercised (theater-level, order of battle, etc.), data may be fed back to command teams in the form of initial or revised recommendations. Tactical decisions can then be made to appropriately deploy military assets and react to or guard against hostile targets. The proposed capability can provide a means to derive important information about threat status and warfare-engagement issues. This methodology is clearly in support of real time warfighter and time-critical moving target detection scenarios.

The ability to incorporate large amounts of electronic flight information that can be managed with great ease and minimal training has the potential of revolutionizing the general aviation sector. The algorithms and techniques to be developed are also applicable to any commercial industry that evaluates large amounts of data from real-time to historical data in order to determine the current state of the environment and analyze or forecast future trends. Cross discipline applications include biologists, chemists, and other scientists, the virtual operating room, simulation control, commercial aircraft, air traffic control, and salvage, search and rescue management. Other potential civilian/dual use applications include: financial markets, traffic analysis in major cities, manufacturing and operations research applications, medical diagnostics, meteorology and weather forecasting; concealed weapons detection, computer vision, intelligent auto vehicle dashboard/head-up displays, airborne hyperspectral analysis, environmental monitoring, ordnance detection, and geographical information system (GIS) applications. Potential applications also include remote sensing, oil exploration, security systems, and extensions of this technology to integrating data beyond geo-registration domain to medical diagnostics, biophysical and drug-discovery applications, financial industry, internet search engines, application-specific information search engines, and other commercial applications such as emergency control centers, law enforcement, public safety, traffic control and manufacturing process control. The resulting technology may also have applications for such areas as traffic pattern monitoring and the identification of illegal activities in support of homeland defense.

We recognize that aside from the algorithm development and implementation, the viability of the software is the major key to the success of concepts like multisource autonomous data registration. In our approach, the design and implementation of the software toolkit will include an expert system based capability to select the most suitable algorithms. We also recognize that based on the problem to be solved, methodologies like control point mapping, wavelet coefficients, cross-correlation, and mutual information similarity measures, all represent methods that may be required to accomplish registration. Other methods developed by contractors under the DDB program to register Multi-INT data provide another set of algorithms to be included in our framework. In our expert system approach, rules will be used to guide the process of rapidly determining the best way to solve the automatic registration problem and guide the user through the solution process.

REGISTRATION ACCURACY

There are many factors that can affect registration accuracy. These are generally modality dependent and include:

- Radiometric/environmental and geometric factors (lens aberrations, distortions).
- Characteristic image content (extremes in sunlight/shadows and speckling/artifacts as in SAR images).

- Aspect (view) angle and the degree of orthogonality—shallow elevation angles and long range (horizon) perspectives.
- Changing pixel resolutions for variable footprints over a given field of view—poses difficulties for some registration methods.
- Spatio-temporal image spacing.
- The effects of *artifact noise* and *speckling* in a pair of images for a given modality (modality dependent).
- Image deformations.

Limitations may arise with certain registration methods if the following conditions exist:

- Relative image sizes, scaling, and pixel densities are disproportionate.
- Time-variant contrast changes (extremes in shadows, sunlight) for images taken at different times.
- Inappropriate choice of registration algorithm for selected image pairs.
- When image sizes and the corresponding number of pixels are too small.
- When rigid body transformation is violated.
- The search space is too large resulting in computational inefficiency.
- Inaccurate or unknown image resolutions and scaling factors (supporting sensor data is sparse or incomplete).
- Insufficient overlap of images can detract from finding the global optimum efficiently and correctly within the search space—an intelligent way to reduce and determine the search space may be used by first applying coarse registration techniques for multi-modality images to reduce the search space.
- Sensitivity to orientation or angular displacements—the inability to handle elastic body problems which may involve angular perspectives—develop transformation models that account for rotations along different axes—at most, we can accommodate *affine* and *projective* transformations.

In our research so far, we have clearly identified the advantages and some of the drawbacks associated with the candidate algorithms. For instance, whereas MI-based algorithms are robust and generally well suited to the multi-modality image data registration problem, some problems have been encountered, such as: inefficient computation when the search space is too large; inaccuracies when angular displacements or orientations are introduced in one of the images, but not the other (rigid vs. non-rigid body problem); potential errors arising from disparate image sizes, scaling, and drastically different pixel densities between images; and slow convergence when image pairs do not overlap sufficiently. Certain limitations were also observed with the other registration methods (e.g., feature-based) that are being considered for the toolkit. Clearly, some techniques can be implemented to make the algorithms themselves more robust, which will be further addressed in our future research.

Another observation was that in order to reduce computational time and improve registration accuracy, it was necessary to preprocess the images in some suitable manner i.e., prescreening images and then applying one or more registration algorithms in a prescribed order depending on the “contents” of the images in conjunction with certain user-specified constraints. There are yet other factors that were found to detract from achieving accurate registration results. Many of these were outlined earlier. From the tests performed and the observations made to date, we have developed and continue to refine an initial expert system rule base, which is being used to develop an automated geo-registration toolkit. We have also identified and developed potential solutions to these difficulties. These solutions are currently being tested.

Some of the concerns to be addressed and questions that will be raised in order to further measure the feasibility of our approach include:

- What is the state of the art in multisensor registration and information fusion technologies as pertaining to Multi-INT data?
- What are the best methods for the development of generic algorithms for correlating salient features across multiple data sources?
- In implementing an expert system approach, one main question is how to develop the rules for the automatic selection of appropriate registration algorithms and to select a suitable similarity measure? Is it a function of the type of sensor or the type of scene (i.e., its content)?
- What are the preferred registration architectures, methods, and algorithms and what are the associated advantages and limitations of each?

- What factors contribute to registration error and uncertainty when disparate sensor data sets are considered?
- What is the most effective way of computing uncertainty for measured target images, and what methods should be used to reduce uncertainty and registration errors for multisource data sets?
- What technical factors come into play when attempting to optimize image registration accuracy? Confined search space? Heuristics testing? Pyramidal resolutions? Pixel density? Image sizes? Image superposition?
- How do we make use of photogrammetry methods to handle elastic vs. rigid body image registration?
- Are there viable image pair preprocessing or prescreening methods that could be used to improve registration accuracy and precision?
- What registration performance/efficiency measures or figures of merit are most useful?
- What are the best methods to determine user needs and to present customized information so that utility to the user is maximized and information overload is prevented?
- What should the registration toolkit provide as a minimum?
- How do we ensure a flexible, scalable hardware/software architecture to permit the new capability to be readily installed within existing or future onboard sensor fusion and signal processing architectures? Will onboard architectures and computing hardware limit the usefulness of the new technology to be inserted?

PREPROCESSING

Typically, images must be pre-processed in some way before they can be registered. These steps include:

Detection of initial conditions – based on the image metadata, how much initial scaling, rotation and translation must be passed to the registration algorithm? This is done because, typically, the reference image does not “perfectly overlap” the test image. Typically, the reference image will be of different scale, cover a larger area, or be oriented differently from the test image.

Determination of search space – what limits do we want to place on scaling, rotation and translation? Frequently, information about problem error bounds are available – for example, the error bounds for the GPS system used to collect the image coordinates.

Noise detection and suppression – do we need to remove noise from the image? If so, what approach should we apply? Mean filtering, median filtering, Gaussian blurring, wavelet decomposition, and the Lee filter all provide different ways of removing noise from an image, but all affect image content (usually by softening edges).

Feature extraction – techniques for feature extraction abound. Canny edges, Harris corners¹, Gabor filtering², and wavelet decomposition³ are all popular methods.

Image enhancement – do we need to perform sharpening, contrast enhancement, or histogram equalization? The answer to this question lies in whether we are using an intensity-based or feature-based algorithm. Feature-based algorithms typically prefer higher sharpness and contrast prior to feature extraction; intensity-based algorithms can be thrown off by any processing that changes the range or distribution of pixel intensities.

REFERENCE IMAGE SELECTION

Another important part of registration is selection of the proper reference image. A number of different factors control the suitability of a particular reference image, including:

Modality – Is there a reference image that is the same modality as the test image? Or a reference image of a compatible modality?

Resolution/scale – A reference image that is closer in scale (meters per pixel) to the test image is typically more desirable than one that is not. Differences in scale can always be resolved using interpolation; the smaller the difference, however, the less damage done to the image by interpolation.

Time of day – If multiple reference images are available, one typically would like to select the one that most closely matches the time of day of the test image when dealing with EO imagery. The direction of shadows, for example, is different before and after noon, and might contribute to inaccuracy. That said, one cannot simply subtract the time of the test image from the time of the reference image and assume that the image collected closer in time is better. For

example, a test image collected at four in the evening in January in New York will be more similar to an image collected at noon than an image collected at six in the evening, since six in the evening is after sunset. Consequently, what we have done is establish a “daylight” score, based on astronomical calculations of sunrise and sunset time based on latitude, and choose the reference image that has the closest daylight score.

Time of year – In regions with large seasonal climate variations, snow and leaves can cause substantial problems for EO imagery. Consequently, we have been developing a season score for each image, based again on its latitude and the month/day it was collected.

Percentage of overlap – How much of the test image is overlapped by the reference image?

Region of overlap – Does this test image overlap an area of interest in the test image that’s been selected by the user? Can we use scene content measurement techniques (like entropy) to determine whether the reference image overlaps a high- or low-feature region of the test image? Obviously, if the overlap is 100%, this is a non-issue.

Perspective/projection – Frequently, a reference image is collected in “ideal” circumstances – orthorectified, radiometrically corrected, photogrammetrically corrected, georeferenced, etc. – while test images are collected in a more haphazard way. One of the significant ways a test image might differ from a reference image is look angle, or perspective. If a reference image closer in perspective is available, it likely will mean

Final reference image selection is made by scoring each of these criteria from 0-100 points, then summing the points.

AUTOMATIC REGISTRATION ALGORITHM SELECTION

Automatic registration algorithm selection involves not only selecting what family of algorithms to use (feature-based, intensity-based, etc.) but also which implementation of those algorithms. For example, one can use Mutual Information as a similarity metric, but select from a variety of different search algorithms and optimizers. Hence, the problem of “which registration algorithm is the best” is necessarily somewhat open-ended.

Part of the problem is trying to determine criteria that one can use to try to predict registration algorithm performance.

A number of techniques have been recommended:

1. Scene content measurement techniques, such as entropy, Fisher information, and extracted feature counts.
2. Scene content description, such as “Trees”, “Roads”, “Body of Water” as entered by a user.
3. Modality compatibility – give the sensor types of my test and reference images, what algorithm will likely work best?
4. Run-time constraints – will the algorithm complete in the amount of time we have available?
5. Accuracy requirements – how precise a registration does the user require? Is the user willing to accept lower precision in return for faster execution?
6. Character of the algorithms themselves – experience gained through testing. (See Table 1)

In the end, the goal is to develop a set of rules that can be encoded into an expert system that will have the ability to select the best algorithm. Initial rules were derived, tested, and evaluated for the MSTAR data problem. It was observed that the MSTAR data were not converging when only MI was used and run times were lengthy. The same converged very fast and accurately when the Fourier-Mellin method was applied as a preprocessor and then MI was used. Two other types of images (one of them was IKONOS data of the Syracuse University quad area and the other one of the lake and the adjoining land tract) were used and there was good improvement in the speed and convergence rate.

The initial set of rules was developed based on these observations and for a few selected user inputs. A preliminary rule base was rapidly prototyped using Gensym’s G2 Real Time Expert System. The rapid prototyping was achieved using simulated data in which simulated user inputs (randomized) were automatically fed into the G2 system and where each input parameter or variable was constrained in the following manner:

- Run Time Constraint (RTC) (optional/default)
- Accuracy Requirements:
 - Coarse Precision (CP)
 - Moderate Precision (MP)
 - High Precision (HP)
- Image Content (a priori knowledge):
 - Clustered Buildings
 - Isolated Buildings
 - Roads, Rivers or Railway Tracks

- Agricultural Area
- Large Water Bodies (e.g. lakes and oceans)
- Other...
- Image or Data Type (a priori knowledge):
 - SAR
 - IKONOS
 - MSS
 - LANDSAT
 - HSI
 - Other...

Table 1: Advantages and Drawbacks of Selected Registration Methods-Basis for General Rules

Candidate Registration Method	Advantages/When It Works	Drawbacks/When It Fails	Comments
Correlation based	<p>Moderate computational complexity</p> <p>Can exactly align mutually translated images and can be successfully applied when slight rotation and scaling are present using exhaustive search.</p> <p>Can be implemented in frequency domain to speed up computations.</p> <p>The distinctive information is provided by graylevels/colors rather than by local shapes and structure.</p>	<p>Flatness of the similarity measure maxima (due to the self-similarity of the images)</p> <p>Reference and sensed images must have somehow 'similar' or identical intensity functions</p> <p>When search space is large, an efficient global optimizer is required.</p>	<p>Useful for intramodality image registration.</p> <p>Easy hardware implementation.</p> <p>Useful in real time applications.</p> <p>The maximum can be sharpened by preprocessing or by using the edge or via vector based correlation.</p> <p>Computational load will increase dramatically when large rotation and scaling differences are involved using exhaustive search.</p> <p>Pyramidal image representations and sophisticated optimization algorithms can be used to find the maximum of the similarity metric and speed up computations.</p>
Mutual information based	<p>Particularly suitable for registration of images from different modalities.</p> <p>The distinctive information is provided by graylevels/colors rather than by local shapes and structure.</p>	<p>Intensities of reference and sensed images must be statistically related.</p> <p>A robust and efficient global optimizer is required</p> <p>High computational load</p>	<p>Often the speed up of the registration is implemented, exploiting the coarse-to-fine resolution strategy (the <i>pyramidal</i> approach).</p> <p>MI methods work with the entire image data and directly with image intensities.</p> <p>Pyramidal image representations and sophisticated optimization algorithms can be used to find the maximum of the similarity metric and speed up computations.</p>
Feature-point based	<p>Typically applied when the local structural information is more significant than the information carried by the image intensities.</p> <p>Allows for the registration of images of completely different nature (like aerial photograph and map) and can handle complex between-image distortions.</p> <p>Usually low computational load, depending on the feature extraction and pairing algorithms</p>	<p>Respective features might be hard to detect and/or unstable in time.</p> <p>Appropriate invariant and modality-insensitive features can result in good registration results.</p>	<p>Should have discriminative and robust feature descriptors that are invariant to all assumed differences between the images.</p>
Wavelet based	<p>Can be applied for registration of images from different modalities.</p> <p>Typically applied when the local structural information is more significant than the information carried by the image intensities.</p> <p>Allows for the registration of images of completely different nature (like aerial photograph and map).</p>	<p>Respective features might be hard to detect and/or unstable in time.</p>	<p>It is a combination of feature-based and intensity based method. Instead of the image intensities, wavelet coefficients are calculated followed by correlation based method.</p>

The specific rules are summarized in Figure 1. The flow diagrams shown in Figure 2 and Figure 3 show how these rules were implemented and demonstrated in the present automated brokering scheme. A screen capture of the G2 rule base GUI is shown in Figure 4.

Figure 1: Example algorithm selection rules

1. If **no RTC** (this includes the case when a very large amount of time is allowed for processing) and **desire HP**
 - (a) Run Mutual Information (MI)
 - (b) During the run, if Runtime > N (e.g. 1/2 of the RTC or a default time which ever is lower). **Use Pyramidal Decomposition** to speed up search (the number of level is to be determined later)
2. If **no RTC** (this includes the case when very large amount of time is allowed) and **desire MP or CP**, then use the image content and image size information
 - (a) If image content consists of roads & buildings, then use **Feature Based**.
 - (b) If image content consists of land & large water bodies, then use **Fourier Based**.
3. If **severe RTC and desire HP**, then use preprocessing using rules 2a & 2b followed by MI.
4. If **severe RTC and desire MP or CP**, then use, rules 2a & 2b.
5. If **moderate RTC and desire HP**, then
 - Start MI
 - If Runtime > N use rule 1b
 - If Runtime > N still true, then use rule 3.
6. If **moderate RTC and desire MP or CP**, then
 - Use, rules 2a & 2b.

Figure 2: Flow Diagram Showing the Implementation of Rules and Constraint-Based Decisions

User Inputs / *a priori* Knowledge:

Run Time Constraint (RTC) (opt.)

Accuracy Requirements

Image Content

Image or Data Type

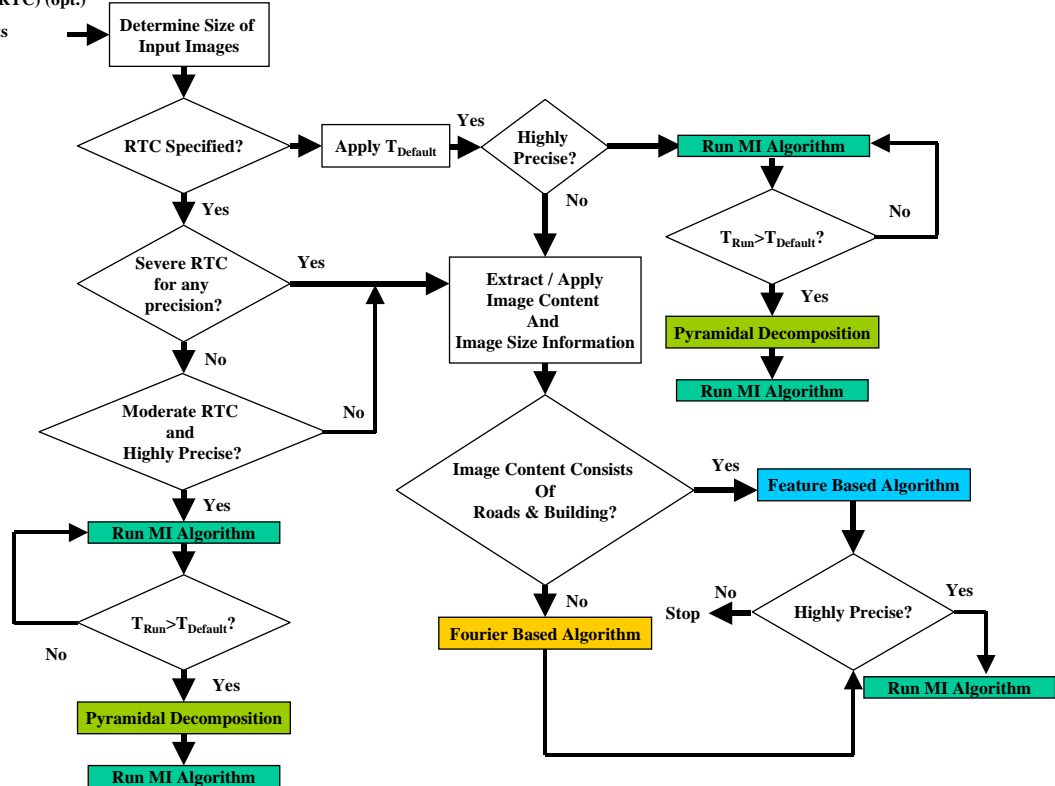


Figure 3: A Possible Additional Prescreening Test for Image Pairs

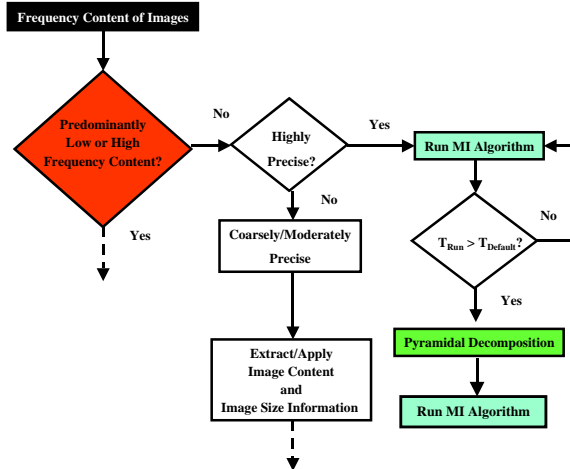


Figure 4: Screen Capture of Preliminary Expert System Tool for Autonomous Geo-Registration



REGISTRATION TESTBED

To help us answer our many questions about registration performance, and how algorithms perform under difference circumstances, we have begun to construct a registration testbed. This testbed will allow us to perform parameter sweeps of a number of different registration algorithm combinations against a number of large image databases. One of the large image databases we have is the MAD98 database, which was created as part of the Dynamic Database (DDB) project. MAD98 contains SAR, Hyperspectral EO, and infrared test imagery, as well as EO orthophotos to use as reference images. We also have ground truth information available for the SAR images.

Have access to a number of different algorithm families for registration:

1. Mutual Information (MI) – our implementation was developed by Dr. Hua-Mei Chen of the University of Texas; it is implemented primarily in MATLAB. For his MI toolkit, we have access to a number of different interpolation techniques (used by an MI algorithm to estimate which pixel in the test image corresponds to a pixel in the reference image, once the test image has been scaled, rotated and translated), such as Linear interpolation, generalized partial volume estimation (GPVE), partial volume estimation (PVE), cubic interpolation, b-spline and nearest neighbor. We have the ability to use rigid body translation; we are looking

to add the affine translation at some point in the future. We are, however, limited to his current optimization code.

2. Earth and Remote Sensing Software (ERSS) – ERSS is a commercial product, produced by A.U.G. Signals of Toronto, Canada. The algorithms contained in ERSS provide unique solutions in the areas of image restoration, classification and detection, registration, image fusion, and hyperspectral image data analysis. To provide a more comprehensive set of tools for processing and analysis, several other popular tools and algorithms have been integrated within the application. This module is an effective tool for the automatic registration of multi-sensor, multi-temporal images from either Electro-Optical or microwave Synthetic Aperture Radar sensors using feature-based GCPs which are selected automatically.
3. Precision Image Control and Alignment (PICA) – PICA is currently under development by BAe Systems' Advanced Information Technologies branch for the US Air Force. It offers a variety of implementations of registration algorithms, including pixel-based Mutual Information, Constant False-Alarm Rate (CFAR), feature-matching using Harris corners or Canny Edges, and correlation.

We are in the process of building some middleware to perform tests and parameter sweeps using these algorithms. The goal is to, using as many images as possible, register the images and evaluate the results to derive some additional rules for the expert system.

PERFORMANCE METRICS

We are currently testing our images against a data set for which we have ground truth, so for performance, we are simply measuring the difference between the results of registration and the “correct answers” for a particular image. As we progress, however, we will begin to use images for which we do not have ground truth. For these situations, we are evaluating the use of Fitness Distance Correlation⁴ (FDC) to generate a measure of registration quality. FDC creates a fitness surface, which represents the output of a mutual information metric at various degrees of rotation and translation. The surface is then analyzed to find the distance from the found solution to the global maximum.

ACKNOWLEDGEMENT

This effort has been funded by the Air Force Research Laboratory's Sensors Directorate (SNAR) under contract FA8650-04-C-1674, a Small Business Innovation Research (SBIR) Phase II program, and F33615-03-M-1494, a SBIR Phase I program.

¹ C. Harris and M. Stephens, "A Combined Corner and Edge Detector," in Proc. 4th Alvey Vision Conference, 1988, pp. 147-152

² Kyrki, V., Kamarainen, J.-K., Kalviainen, H. Simple Gabor Feature Space for Invariant Object Recognition, Pattern Recognition Letters 25, 3 (2004) 311-318.

³ H. H. Li and Y.-T. Zhou, "A Wavelet-Based Point Feature Extractor For Multi-Sensor Image Registration," in *SPIE Aerosense Wavelet Applications III*, Orlando, FL, 1996, pp. 524-534.

⁴ Kumar, Bhagavath, Varshney, Pramod, and Drozd, Andrew, "An approach for performance comparison of image registration methods", to appear in the Proceedings of the SPIE Defense and Security Symposium, Orland, FL, 2006.