

## AUTOMATED MULTISENSOR REGISTRATION: REQUIREMENTS AND TECHNIQUES

Eric Rignot, Ronald Kwok, John Curlander and Shirley Pang  
Jet Propulsion Laboratory  
California Institute of Technology  
Pasadena CA 91109 USA  
Ph 818-354-1640  
telex 675-429

### Abstract

A necessary first step in the fusion of data from a number of different remote sensors is the correction of the systematic geometric distortion characteristic of each sensor followed by a precision registration to remove any residual random offsets. This paper describes our approach to automated multisensor registration. The effects of spatially, temporally and spectrally varying factors which influence image dynamics are reviewed. A specification of the requirements for an operational algorithm is formulated using these factors. Additionally, the structure of an efficient, automated system is defined. A number of candidate image processing techniques are evaluated within this structure using a multisensor test data set assembled from the Landsat TM, SEASAT and SPOT sensors. The results are presented and discussed.

### 1. Introduction

To fully utilize the data collected by the planned NASA and ESA remote sensing spaceborne programs we must analyze the data from a number of different sensors in a multidimensional manner. However, these sensors may be on different platforms and in different orbits. In addition, the disparate viewing geometries, sensor characteristics, data collection and processing systems will introduce systematic and non-systematic geometric errors resulting in the misregistration between coincident data samples from the various sensors. Therefore a crucial and necessary first step in multidimensional analysis of such datasets is the correction of the systematic geometric distortion characteristic of each sensor followed by a precision registration to remove any residual offsets.

The traditional technique for registering disparate image data is by visual identification of tiepoints or recognizable features in the image datasets, the deformation field determined from these points is used to resample the images onto a common grid prior to data analysis. This approach is efficient as long as the data volume is small. With the anticipated data volume and data rate of the future high resolution sensors, automated registration techniques become absolutely necessary for handling the substantial amount of data which will be generated. Correlation techniques based on image intensity and developed for the operational registration of multitemporal Landsat data cannot be applied to registration of image data from different sensors due to the feature dynamics across the spectral bands. Development

of extra stages (more sophisticated segmentation techniques) in an operational processor is required before matching of the different features can proceed, therefore creating a new challenge. In addition, a large number of new techniques for determining image correspondences has appeared in the interim, the utility of which should be examined for potential incorporation into an operational algorithm.

This paper presents a summary of our approaches to an automated multisensor registration algorithm. The algorithm requirements are formulated in part 2 by considering the image/sensor dynamics and the geometric and physical characteristics of image data generated by various data processing systems. Part 3 presents the multisensor test dataset assembled for testing our image registration approaches. Part 4 describes a first type of automated registration technique based on the availability of digital elevation models (D.E.M.) of the areas to be registered. Part 5 presents a more general set of automated registration techniques where scene features are automatically extracted from the image data using segmentation techniques, and correspondence between these features across the different images to be registered is performed using matching algorithms. Conclusions and recommendations are given in part 6.

### 2. Multisensor Image Registration: Factors affecting its quality, and Requirements

A first set of factors that affect multisensor image registration comes from the sensor specific platform ephemeris and attitude. For example platform location errors directly translates into relative location error between scenes from different platforms, and attitude drift of the platform causes geometric distortions within an image frame. Depending on the stability of the sensor, such errors are usually removed by the use of tiepoints, but non-systematic errors and tiepointing introduce bias in the image location and therefore a final step of image registration for refined to sub-pixel level registration accuracy is necessary. The geometric quality of the acquired data is also affected by the presence of topography in the observed scene. For an active sensor like the SAR, quite different terrain-induced geometric distortions such as foreshortening and layover [Lewis, 1970] constitute additional difficulties. Rectification of these geometric distortions is a first step in elimination of the sensor-specific viewing effects.

Multisensor image registration is also affected by the variety of grid representations. Usually, when the platform/sensors are in

different orbits, the acquired data are initially sampled to grids which are more natural to the sensors than that for image registration. More practical grids for image coregistration are map grids where available topographic and thematic data are usually presented. Such map grids are Earth-fixed grids with known two-dimensional distortions and the mapping process is named geocoding.

Resolution is an important factor because generally the achievable registration accuracy between two images will be measured in terms of this quantity (rather than in terms of pixel spacing). This is of importance considering that resolution can vary tremendously from one instrument to another: the resolution of EOS imaging sensors will range from tens of meters (SAR, HIRIS) to kilometers (MODIS, HMMR) [EOS, 1987].

Among the physical characteristics of the image data from different sensors, the level of system noise (thermal, quantization etc.) is an important factor because segmentation and feature extraction techniques are usually very sensitive to noise. While all sensors are corrupted by additive thermal noise from the receiver electronics, SAR images are additionally corrupted by multiplicative noise known as image speckle. Special consideration for speckle statistics must be accounted for in the design of segmentation schemes for SAR images.

A last factor that affects multisensor image registration is the quality of the model being used to describe the scene and its digitized representation. Scenes are very difficult to model. Traditionally they are described as being composed of a set of features. An obvious requirement of a robust multisensor matching algorithm is the capability to select features which are common across the sensor data to be matched and whose frequency dependent dynamics are well understood.

In view of the above remarks the input and output requirements that define the operational domain and conditions under which the multisensor algorithm is expected to operate are formulated. They can be used as a basis for the evaluation of candidate algorithms. The input data are assumed to have already undergone a systematic geometric correction so that registration is only limited to the determination of the residual systematic and unsystematic biases. The geodetic accuracy of the input images following this geometric correction shall be better than 10-50 pixels. This requirement constrains the search space for the algorithm. Geometric differences between images are assumed to be predominantly translational. The terrain-induced and attitude-induced geometric distortions of the image data shall be rectified before input based on the best information available. The input data shall also be presented in a preselected map grid (e.g. U.T.M.) which is common to all multisensor data and the data shall be resampled to the same pixel spacing. The input data should maintain a minimum signal-to-noise ratio of 5 dB where the noise power is defined to consist of the contribution from all noise sources. In terms of output products, the registration accuracy shall be better than one resolution element.

### 3. Description of the Multisensor Test Data Set

Geocoded products from SEASAT, SIR-B, Landsat TM, SPOT

and TIMS (resampled to the same pixel spacing) were assembled as part of a multisensor test data set [Kwok et al. 1989]. Table 1. summarizes the principal characteristics of the different sensors: type of sensor (active vs passive); look angle; frequency; polarization and resolution of each instrument. Several sites were selected among the different images where the sensors have co-incident coverage, and manually coregistered. The relative misregistration was estimated to be  $\pm 2$  pixels by tie-pointing, i.e. roughly equal to the largest resolution element (typically 40m). Table 2 summarizes the characteristics of the sub-images: geographic location; sample spacing; size; revolution number; date; map projection; list of features present in the scene.

### 4. Automated Registration to Digital Terrain Data

The ability to automatically register sensor imagery to topographic maps generally requires the presence of identifiable, man-made or natural features both in the image data and the corresponding map. In mountainous and uninhabited areas such features are difficult to locate. In such areas, an alternative to inter-image matching using image neighborhoods is the registration of each image data to a simulated image generated from a digital elevation model (DEM). An illustration of the utility of the technique will be presented in the case of the registration of a SEASAT SAR image with a Landsat TM image.

### 5. Computational Approaches to Multisensor Registration

This section describes the method by which various segmentation and matching techniques can be incorporated into the general process of registration. The structure of our multisensor registration algorithm is shown in Table 3. The input data are geocoded images. The registration problem is reduced to the determination of the local relative translational displacement in the image plane between a pair of images. First, a set of image patches is automatically selected in each geocoded image to define local areas where the precision registration will be performed with a high confidence of success. The patches are then segmented to extract stable features (edges, regions, etc..) and matched them between the different images. Matching is performed in several locations of the segmented image and the results are filtered in a last operation to check their relative spatial consistency within the selected patch (local constraints). At a higher-level of processing, the combined results from each segmentation / matching scheme and from already registered neighbor patches (global constraints) are used to produce a more accurate and more reliable solution. In effect, a cooperative process can be established where the results from different steps in the algorithm are used as reinforcements with the entire process. Each one of these steps is presented in the next few sections.

#### 5.1 Automated Selection of Tie-points

Patches must be selected that define local areas where fine registration can be performed with a high confidence of success. They must correspond to stable strong features that can be unambiguously identified in the different sensed images and matched rapidly with a pixel-level accuracy. Possible features include

rivers, lakes, coast-lines, or any type of large natural or man-made structure that presents some element of similarity across the different images to be registered. One possible technique has been described in [Davis and Kenue, 1978] based on binary edge-maps. However this still remains an important area of investigation, the difficulty being to formulate an approach without *a priori* knowledge of the scene content.

## 5.2 Segmentation

Two types of segmentation techniques were investigated for feature extraction: 1) those based on edge operators and 2) those on the stationarity properties of regions.

**5.2.1 Region-Analysis** Although the dynamics of the spectral response of different targets can vary a lot from one instrument to the other, one can still expect to detect regions such that their shape, spatial extent or location in the image plane is largely preserved across the spectrum. The goal is to design techniques that are compatible from one sensor to another. Usually, most segmentation techniques designed for optical images are not effective for SAR images. Acceptable results were obtained however using an unsupervised technique based on a clustering algorithm that segment the images into several regions of similar intensity and texture [Kwok et al., 1989]. An example of such segmentation is shown on Fig. 1 together with the original images from SEASAT, Landsat, and SPOT. The results obtained by matching the region boundaries are not as good as those obtained using other techniques but they can still be significantly improved, and always provide a complement of information to results from other techniques.

**5.2.2 Edge Detection** An extensive literature exists on the subject of edge detection in optical images since it is a basic operation in image processing. However, in the case of SAR images the detection process is complicated since the images are corrupted by speckle noise. Techniques based on an approximation of the first and second directional derivatives (e.g. Sobel, or Robert operators) perform poorly, especially in terms of localization of the edges since they tend to produce large responses. Statistical edge operators in a lot of cases suffer from the same limitation. The problem is solved by regularization techniques, specifically using a two-dimensional Gaussian smoothing operator as in a Marr-Hildreth operator [Marr and Hildreth 1980] and a Canny edge detector [Canny 1983]. These operators have good detection and localization properties without multiple responses to a single edge, the three performance criteria for evaluation of edge detection algorithms. Theoretically, these techniques are compatible to almost all types of remote sensor data. Their performance with optical data have been documented in the literature [Marr and Hildreth 1980, Canny 1983]. In the case of SAR imagery their performance was quantitatively compared in [Kwok and Rignot, 1989] where it was shown that the gradient operator outperforms the Laplacian operator in both detection and localization of edges in image speckle. Several parameters affect the quality of the gradient operator, such as the filter spatial width of the gradient operator, the selection of the thresholding of the gradient magnitude to eliminate insignificant edges, post-processing of the edge-maps such as contour filling and thinning techniques [Kwok et al., 1989]. For illustration, an example of edge-map using SEASAT, Landsat TM and SPOT data and the

Canny edge detector is presented on Fig 2.

## 5.3 Image and Feature Matching

**5.3.1. Binary Cross-correlation** Edges or region boundaries can be cross-correlated between the two images to be registered. Since the edge-maps (or region boundaries) can be represented as binary images, the process is fast and can be efficiently implemented on an array processor or any other vectorizing computer.

**5.3.2. Distance Transform, Chamfer Matching** The basics of the distance transform and Chamfer matching technique are described in [Barrow et al. 1977]. This algorithm matches features (edges or region boundaries) by minimizing a generalized distance between them. This is highly efficient since the computational cost is proportional to a linear dimension. The feature-map from the source image is transformed into a distance-map assuming each grey level represents the distance to the nearest feature-point. The quality of the match is then measured by summing the distance-map values at each location of a feature-point in the target image. Even when several target-images are being coregistered, only one distance-map needs to be computed. The algorithm is by design more robust to distortion or residual rotation effects than a binary cross-correlation.

Using our data set, we found the binary cross-correlation was up to 4 times faster than Chamfer matching, and both of them are far more efficient than area-correlation. The binary cross-correlation consistently showed a gain in the quality of the matches between 10 and 20 % compared to Chamfer matching. The quality metric used during Chamfer matching does not perform as well as expected with multisensor data because of the bias introduced in the total distance between feature points by non-matchable edges (edges that do not have a correspondence in the other image and that always exist in multisensor image registration). Non-matchable edges do not affect the binary cross-correlation unless they are much more numerous than matchable edges. Subsequent tests also indicated that Chamfer Matching is indeed more robust to rotation effects, but the performance of the two techniques is comparable when the binary edges of the cross-correlation are thickened from 1 to 3 pixels. Based on a maximum registration accuracy of 1 resolution cell ( $\approx 40m$ ), tolerance to rotation is about 3 degrees.

**5.3.3 Dynamic Programming based on an Autoregressive Model.** The method is based on the combination of dynamic programming and of an autoregressive model. It was used to register severely distorted optical images to a reference map [Maitre and Wu, 1989] without any *a priori* knowledge of the distortion. The good performance of the technique with optical images needs to be further investigated with other types of multisensor data. In that case, with an expected deformation that is mostly translational, the optimization process would be simpler. The *a priori* inconveniences of this method are the high complexity of the process (of the order of  $O(NM^2)$  where N and M are the the number of features in the source and the target images, typically several thousands in a 512 x 512 image), and the sequential tracking of the feature-points (i.e. no straightforward implementation on a parallel machine).

### 5.4 Constraint Filtering

In practice, matching is performed on small areas (typically 256 x 256 pixels or less) such that the distortion is small, the time of computation is reduced and the number of estimates of the global shift of a selected patch is increased. The results must then be filtered to eliminate false matchings. A clustering technique is used where the cluster center defines the estimate of the global shift of the patch.

### 5.5 Results of the Registration Process

When the 48 coincident sub-patches from SEASAT SAR and SPOT were coregistered the rate of success of the binary cross-correlation of edges was 87 %. The use of local constraints improved this result to 92 % correct registration. Registration fails in the case of a river bed which produces very dissimilar responses between the optical and the SAR sensors. When SEASAT SAR data are coregistered to Landsat TM data the rate of success is 86 % after filtering. Matching fails in the case of the river bed and in the case on an urban scene. Registration was qualitatively more difficult than with SPOT because of the lower resolution of the Landsat images.

### 6. Conclusions

While entering in a new era of remote sensing (permanent platforms types) it is of considerable importance to stress the necessity of developing automated multisensor registration techniques. We presented the status of our current work in the development of such tools. Input and output requirements were presented after examination of the geometric and radiometric image dynamics across a common set of active and passive instruments. A structure for the multisensor registration algorithm has been defined, and candidate techniques which fit into this framework have been tested using a multisensor dataset assembled from Landsat TM, SPOT, SEASAT and SIR-B SAR, and TIMS. Further efforts should be focused in that direction.

### 7. Acknowledgments

This work was carried out under contract with the National Aeronautics and Space Administration at the Jet Propulsion Laboratory, California Institute of Technology.

### 8. References

Barrow, H. G., J. M. Tenenbaum, R. Bolles, H. C. Wolf. Parametric Correspondence and Chamfer Matching: Two New Techniques for Image Matching. *Proc. DARPA Image Understanding Workshop*, May 1978.

Canny, J. F. Finding edges and lines in images. *Artif. Intell. Lab., Mass. Inst. Technol., Cambridge, MA. Tech Rep. AI-TR-720*. June 1983.

Davis, W. A. and S. K. Kenue, Automatic Selection of Control Points for the Registration of Digital Images, *Proc. of the Int. Joint Conference on Pattern Recognition*, 4<sup>th</sup>, Kyoto, Japan, Nov 7-10 1978.

EOS Earth Observing System Reports, Vol. I and Vol II, 1987.

Kwok, R. and E. Rignot. Comparison of the  $\nabla^2 G$  and  $\nabla G$  Operators for Edge Detection in Speckle Noise. Submitted to *IEEE Trans. on Acoust., Speech, and Signal Proc.*

Kwok, R., E. Rignot, J. C. Curlander and S. Pang. Multisensor Image Registration: A Progress Report, JPL D-6097, September 1989.

Lewis, A. J. and H. C. Mac Donald, Interpretive and mosaicking problems of SLAR imagery, *Remote Sens. Environ.*, 1, 213, 1970.

Maitre, H. and Wu, Y., Dynamic programming algorithm for elastic registration of distorted pictures based on autoregressive model, *IEEE Trans. on Acoust., Speech, and Sign. Proc.*, Vol. 37, No. 2, Feb 1989.

Marr, D. and E. Hildreth, Theory of edge detection. *Proc. Royal Soc. London, ser. B*, vol 207, pp 187-217, 1980.

Ulaby, F., R. K. Moore and A. K. Fung, *Microwave Remote Sensing: Active and Passive*. Vol. III. Artech House, Inc., 1986.

NAME OF SENSOR	TYPE	FREQUENCY	RESOLUTION (ORIGINAL DATA)
SEASAT	SAR ACTIVE	L BAND, HH POLARIZATION 23° LOOK-ANGLE	25 m
SIR-B	SAR ACTIVE	L BAND, HH POLARIZATION 15 TO 80° LOOK-ANGLE	25 m
TIMS	RADIOMETER PASSIVE	THERMAL-INFRARED	30 m
LANDSAT TM	OPTICAL PASSIVE	7 BANDS SPECTRAL RANGE: 1. 0.4 - 0.45 μm    0.45 μm    0.53 μm    0.65 μm 2. 0.65 - 0.68 μm    0.68 μm    0.73 μm    0.8 μm 3. 0.85 - 0.88 μm    0.88 μm    0.93 μm    1.1 μm 4. 1.1 - 1.25 μm    1.25 μm    1.65 μm    2.1 μm	28.5 m (pixel spacing) IFOV
SPOT	OPTICAL PASSIVE	3 BANDS SPECTRAL RANGE: 1. 0.45 - 0.68 μm 2. 0.68 - 0.8 μm 3. 0.85 - 1.1 μm	20 m (pixel spacing) IFOV

Table 1.

IMAGE FRAMES LOCATION	SEASAT	LANDSAT	SPOT	TIMS	FEATURES IN PATCHES
ALTAMAHA RIVER, GEORGIA (pixel size = 20 m)	Rev. 407 Date: Jul 78 Size: 5 K x 5 K Map Proj.: UTM # of patches selected for testing: 12	Date: Jul 84 Size: 3 K x 4 K Rotated to North # of patches selected for testing: 12	Date: Sept 84 Size: 3 K x 5 K Rotated to North # of patches selected for testing: 12		Rivers, Lakes, Fields, Roads, Coasts.
WIND RIVER BASIN, WYOMING (pixel size = 30 m)	Rev. 781 Date: Aug 78 Size: 5 K x 5 K Map Proj.: UTM # of patches selected for testing: 5	Date: Jun 84 Size: 5 K x 5 K Rotated to North # of patches selected for testing: 13			Mountains, Rivers, Lakes, Fields, Roads, Cities.
YUMA, ARIZONA (pixel size = 25 m)	Rev. 681 Date: Aug 78 Size: 3 K x 3 K Map Proj.: UTM # of patches selected for testing: 13	Date: Jun 84 Size: 3 K x 4 K Rotated to North # of patches selected for testing: 13			Mountains, Rivers, Fields, Roads, Cities, Dunes.
DEATH VALLEY, CALIFORNIA (pixel size = 25 m)	Rev. 882 Date: Aug 78 Size: 5 K x 5 K Map Proj.: UTM # of patches selected for testing: 4	Date: Nov 82 Size: 3 K x 4 K Rotated to North # of patches selected for testing: 4		Date: Jul 83 Size: 1 K x 1 K Rotated to North # of patches selected for testing: 4	Mountains, Fields.

Table 2.

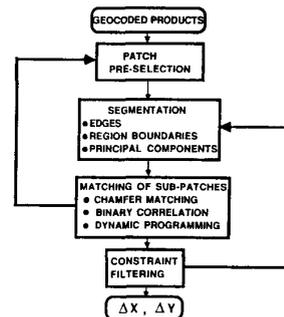


Table 3.