Registration Techniques for Multisensor Remotely Sensed Imagery

Leila M.G. Fonseca and B.S. Manjunath

Abstract

Image registration is one of the basic image processing operations in remote sensing. With the increase in the number of images collected every day from different sensors, automated registration of multisensor/multispectral images has become a very important issue. A wide range of registration techniques has been developed for many different types of applications and data. Given the diversity of the data, it is unlikely that a single registration scheme will work satisfactorily for all different applications. A possible solution is to integrate multiple registration algorithms into a rule-based artificial intelligence system so that appropriate methods for any given set of multisensor data can be automatically selected. The first step in the development of such an expert system for remote sensing application would be to obtain a better understanding and characterization of the various existing techniques for image registration. This is the main objective of this paper as we present a comparative study of some recent image registration methods. We emphasize in particular techniques for multisensor image data, and a brief discussion of each of the techniques is given. This comprehensive study will enable the user to select algorithms that work best for his/her particular application domain.

Introduction

Image registration is the process of matching two images so that corresponding coordinate points in the two images correspond to the same physical region of the scene being imaged. It is a classical problem in several image processing applications where it is necessary to match two or more images of the same scene. Some examples of its applications are:

- Integration of information taken from different sensors (sensor or image fusion problem).

In remote sensing, a great number of sensors for global monitoring are available, each of them with different spectral, spatial, and radiometric characteristics. It is useful to combine and analyze the image data to take advantage of their characteristics and improve the information extraction process. For example, the combination of images obtained from SPOT and Landsat Thematic Mapper (TM) satellites has been used in applications such as monitoring urban growth. SPOT images present better spatial resolution than do the TM images while the TM images have better multispectral resolution. The Intensity-Hue-Saturation transformation (IHS) can be used to merge the SPOT panchromatic band with TM multispectral bands and generate another color enhanced image with high spatial resolution (Carper, 1990). The alignment of the images is the first step in this data transformation.

Another example is combining optical and radar images. Radar images are not affected by clouds and weather conditions, and provide important complementary information about the region surveyed. For example, synthetic aperture radar (SAR) data from the Shuttle Imaging Radar-C (SIR-C) and Japanese Earth Resources Satellite-1 (JERS-1) data combined with TM optical sensor data have been used to map floodplain inundation and vegetation in the Manaus area of Brazil (Melack, 1994). SAR sensors are uniquely suited to measure floodplain inundation because they can detect flooding underneath vegetation, and they operate independently of cloud cover or solar illumination, while TM data provides additional information from the optical portion of the spectrum. The problem is that these sensors are on different platforms and in different orbits, each having different characteristics, viewing geometries, and data collection and processing systems. This makes it necessary to register the images prior to their analysis.

- Analysis of changes in images taken at different times (temporal registration and change detection).

In multitemporal image analysis, the objective is to detect changes which have occurred over a certain time period. A simple method to find changes in a pair of images is to overlay the images and detect the differences between them. Because these images are taken at different times and under different conditions, they have to be aligned prior to comparative processing.

- In computer vision, registration is necessary in extracting structure from motion, electronic image stabilization, and object recognition.

Other problems such as finding cloud heights, satellite image composite generation, weather prediction, and wind direction measurements also involve the registration process.

Two examples of image registration are shown in Figures 1 and 2. Figures 1a and 1b show balloon images from a Mojave Desert sequence taken with a CCD camera. They were part of a motion sequence with the camera attached to a floating balloon. Figure 1c shows the mosaicking of Figures 1a and 1b after registering them. Figure 2 illustrates the multisensor registration of Landsat TM and SPOT images. As the SPOT images have higher spatial resolution than Landsat TM, the features appear at different scales and registration is necessary to integrate their information. Matching of the SPOT image after the transformation is shown in the Figure 2c.

Center for Information Processing Research, Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106.

L.M.G. Fonseca is with the Instituto Nacional de Pesquisas Espaciais, Av. dos Astronautas 1758, 12227-010, São José dos Campos, SP, Brazil.

Photogrammetric Engineering & Remote Sensing,
Vol. 62, No. 9, September 1996, pp. 1049–1056.

0099-1112/96/6209-1049$3.00/0

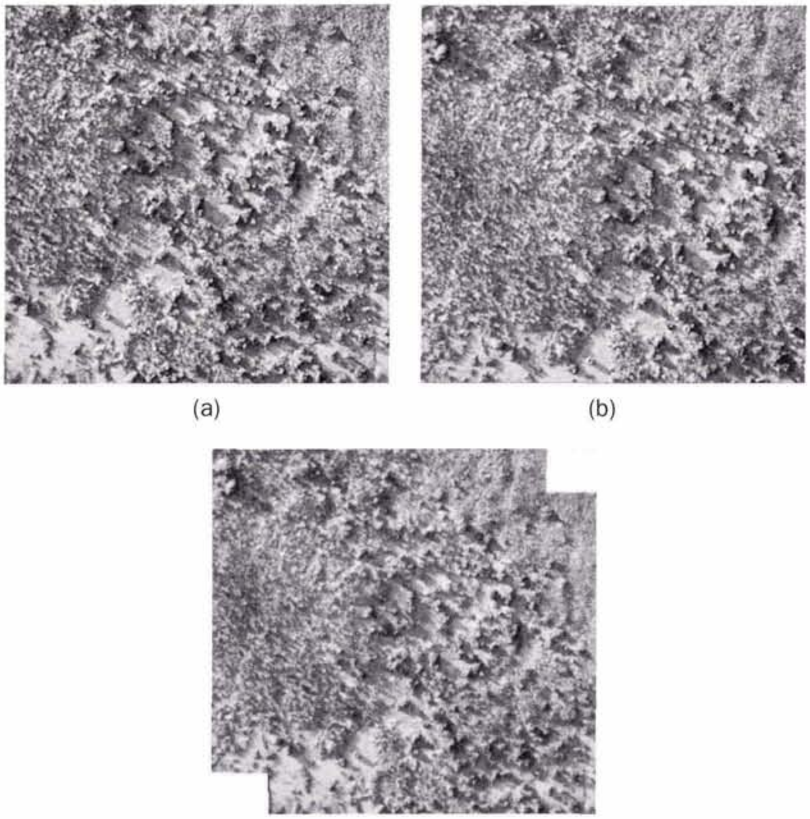


Figure 1. (a) and (b) are images from a Mojave Desert sequence; (c) is the mosaicking of (a) and (b). (Original images courtesy of JPL.)

In remote sensing applications, there is a critical need to develop automated image registration techniques which require minimum human interaction. Because the performance of a methodology is dependent on application specific requirements, sensor characteristics, and the nature and composition of the imaged area, it is unlikely that a single registration scheme will work satisfactorily for all different applications. Integration of multiple registration algorithms into a rule-based artificial intelligence system, which can analyze the image data and select an appropriate set of techniques for processing, appears to be a feasible alternative. Information such as the data type, features present in the imaged scene, registration accuracy, image variations, and noise characteristics could be provided by the user to assist in this process.

The first step in the development of such an intelligent system would be a better understanding and characterization of the various different existing techniques. This is the main objective of this paper as we present a comparative study of recent image registration methods. In selecting the methods described here, the criteria used include potential for multisensor/temporal image registration and detailed experimental evaluation. Each methodology has been categorized with respect to the type of sensor data, modality (multi or single sensor), amount of test data used in the experiments, amount of overlap tolerated, as well as type of image feature, matching techniques, and type of transformations that have been used. In addition, some observations in relation to the merits and limitations of these methodologies are also presented.

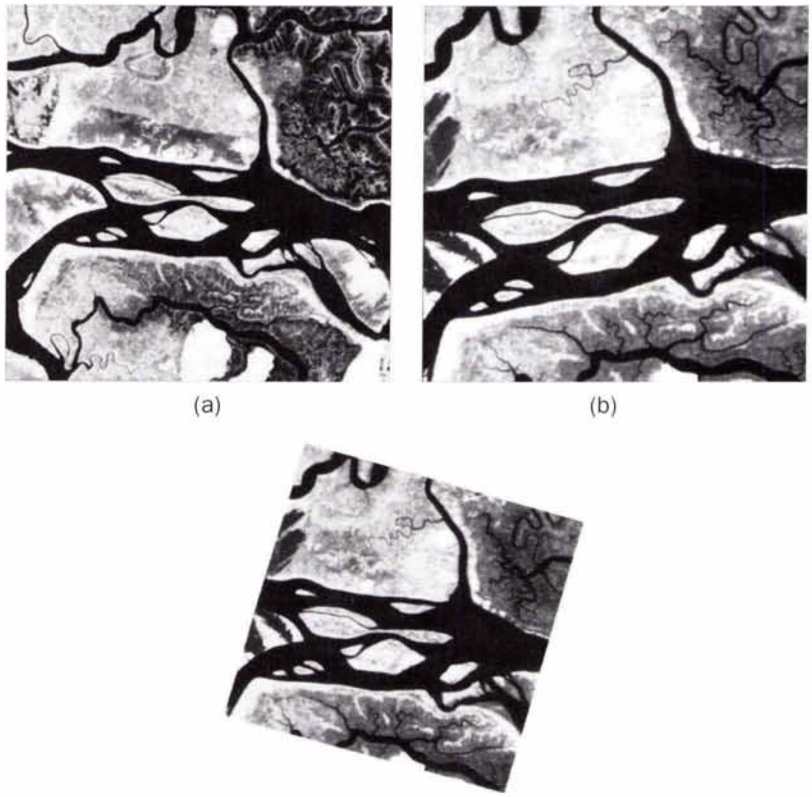
The organization of the paper is as follows. In the next section we provide the reader with an introduction to the image registration problem, describing the common tasks involved in the image registration process. Next, the description of selected algorithms for image registration proposed in the literature are presented, followed by a comparative study of these algorithms. Finally, we conclude with discussions.

The Image Registration Problem

Image registration is the process of overlaying two or more images of the same scene. The one which is registered is called the *reference image* and the one which is to be matched to the *reference image* is called the *sensed image*. The general approach to image registration consists of the following four steps:

- Feature Identification. Identifies a set of relevant features in the two images, such as edges, intersections of lines, region contours, regions, etc.
- Feature Matching. Establishes correspondence between the features. That is, each feature in the *sensed image* must be matched to its corresponding feature in the *reference image*. Each feature is identified with a pixel location in the image, and these corresponding points are usually referred to as *control points*.
- Spatial Transformation. Determines the *mapping functions* that can match the rest of the points in the image using information about the *control points* obtained in the previous step.
- Interpolation. Resamples the *sensed image* using the above mapping functions to bring it into alignment with the *reference image*.

In general, the registration methods are different from



(a) (b) (c)

Figure 2. (a) and (b) are the Landsat and SPOT images, respectively, with scale difference; (c) shows the matching SPOT image after transformation. (Li *et al.* (1995) © 1995 IEEE, original images courtesy of JPL.)

each other in the sense that they can combine different techniques for feature identification, feature matching, and mapping functions. The most difficult step in image registration is obtaining the correspondence between the two sets of features. This task is crucial to the accuracy of image registration, and much effort has been spent in the development of efficient feature matching techniques. Given the matches, the task of computing the mapping functions does not involve much difficulty. The interpolation process is quite standard and will not be discussed in this paper.

Control Point Identification

Manual Registration

The traditional manual approach uses human assistance to identifying the *control points* in the images. In this approach, the steps of feature identification and matching are done simultaneously. The images are displayed on the screen and the user chooses corresponding features in the images which clearly appear in both the images. Candidate features include lakes, rivers, coast-lines, roads, or other such scene-dominant man-made or natural structures. Each of these features will be assigned one or more point locations (e.g., the centroid of the areas, or line endings, etc.), and these points are referred to as *control points*. These *control points* are then used in the determination of the mapping function. In order to get precise registration, a large number of control points must be selected across the whole image. This is a very tedious and repetitive task. Furthermore, this approach requires someone who is knowledgeable in the application domain and is not feasible in cases where there is a large amount of data. Thus, there is a need for automated techniques that require little or no operator supervision. Based on the nature of features used, automated registration methods can be broadly grouped into area-based and feature-based techniques.

Area-Based Registration

In the area-based methods, a small window of points in the *reference image* is statistically compared with windows of the same size in the *sensed image*. This process is illustrated in the Figure 3. Consider the *sensed image* S with M rows and N columns, and n windows W_z, $z = 1, ..., n$, with K rows and L columns extracted from the *reference image* R and centered at the point (a_z,b_z). Let S_{ij} *denote the* K by L subimage of S with its upper left corner coordinates (i,j), where

$$S_{ij}(l,m) = S(i + l, j + m), \tag{1}$$

for $0 \leq l \leq K - 1$, $0 \leq m \leq L - 1$, and $0 \leq i \leq M - K$, $0 \leq j \leq N - L$. Each window W_z is compared to every subimage S_{ij} in the image S. After finding the subimage S_{ij} which best matches W_z, their centers (a_z,b_z) and $(i + (K - 1)/2, j + (L - 1)/2)$ are taken as the control points. Then the control points can be used to calculate the transformation parameters.

The comparison uses similarity metrics which measure the similarity between two given windows, and is maximized over all potential candidate matches. Common simi-

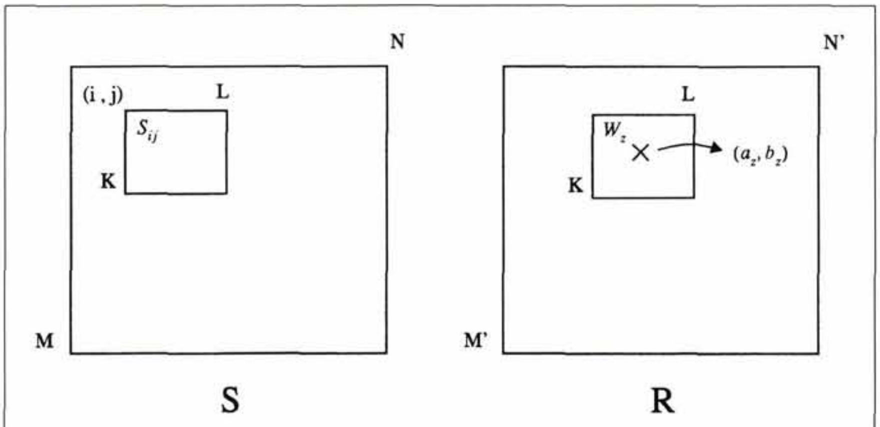


Figure 3. Area-based registration.

larity measures are the *normalized cross correlation, correlation coefficient*, and the *sequential similarity detection* (Barnea and Silverman, 1972). For the window W_z and the subimage S_{ij}, defined above, *the normalized cross correlation* can be mathematically represented by

$$R(i,j) = \frac{\sum_{l=0}^{K-1} \sum_{m=0}^{L-1} W_z(l,m) S_{ij}(l,m)}{\sqrt{\sum_{l=0}^{K-1} \sum_{m=0}^{L-1} W_z^2(l,m) \sum_{l=0}^{K-1} \sum_{m=0}^{L-1} S_{ij}^2(l,m)}} . \quad (2)$$

The best match occurs when the value $R(i,j)$ is maximum.

The *correlation coefficient* measures similarity between two windows on an absolute scale ranging from $[-1,1]$: i.e.,

$$C(i,j) = \frac{\sum_{l=0}^{K-1} \sum_{m=0}^{L-1} (W_z(l,m) - \mu_w)(S_{ij}(l,m) - \mu_s)}{\sqrt{\sum_{l=0}^{K-1} \sum_{m=0}^{L-1} (W_z(l,m) - \mu_w)^2 \sum_{l=0}^{K-1} \sum_{m=0}^{L-1} (S_{ij}(l,m) - \mu_s)^2}} , \quad (3)$$

where μ_w and μ_s are means (average intensity value) of the window W_z and subimage S_{ij}, respectively.

Computationally much simpler is the *sequential similarity detection* algorithm proposed by Barnea and Silverman (1972). It is based on the absolute difference between the pixels in the two images,

$$E(i,j) = \sum_{l=0}^{K-1} \sum_{m=0}^{L-1} | W_z(l,m) - S_{ij}(l,m) | , \quad (4)$$

and on the following sequential search method. The similarity measure defined in Equation 4 is accumulated for pixel values in a window area until a predetermined threshold value is exceeded; the number of pixels examined when this threshold is exceeded is recorded as a rating of the test; when all subimages have been examined, the subimage with the largest rating is considered the best match.

Area-based correlation methods can be efficiently implemented in the Fourier transform domain using the Fast Fourier Transform (FFT). Some of the properties of the Fourier transform can be used to achieve invariance to translation, rotation, and scale.

In the past, the correlation method was limited to the registration problems in which the images were misaligned by small rotational differences. Recently, Zheng and Chellapa (1993) proposed a novel solution to this problem by correcting the image rotation first, using the illuminant direction estimation method, and then determining corresponding control points using area-based correlation in the spatial domain. However, the correlation-based algorithms are not appropriate to register images taken from different acquisition systems. The gray-level characteristics of the images to be matched can vary from sensor to sensor, and the correlation measures become unreliable. In this context, feature-based methods, which extract and match the common features from two images, have been shown to be more suitable.

Feature-Based Registration

In feature-based registration, two tasks are involved: feature extraction and feature matching. They seek to extract and match the common features from the two images. In the feature extraction phase, the image is represented in a compact form by a set of features, and the matching process is carried out in the feature space. Feature proprieties invariant to scaling, rotation, and gray-level modification are often used in the matching process.

For the feature extraction purpose, the image can be represented either in the *spatial domain* or in the *transform domain*. In the *spatial domain*, the candidate features commonly used in digital imagery include edges, regions, lines, line endings, line intersections, region centroids, curvature discontinuities, etc. One important requirement is that the features must be robust to changes in sensor geometry, wavelength, and noise characteristics. Two particular types of features – region boundaries and edges – have been extensively used. In general, edges and region boundaries are extracted using edge detection and segmentation techniques such as the Canny's operator (Canny, 1986), the Laplacian of the Gaussian operator (Marr and Hildreth, 1980), or any region growing methods (Ballard and Brown, 1982).

The feature matching algorithms make use of attributes such as shape (perimeter, invariant moments, ellipticity, thinness), color, texture, etc., and by relations defined by the spatial arrangement. Each feature in one image is compared with potential corresponding feature in the other image. A pair of features with similar attributes are accepted as matches. In computing the mapping function, centroids of closed boundary regions, salient points along the contours, or locations of maximum curvature can be used as control points.

In the *transform domain*, the image can be decomposed and represented as a set of transform coefficients. For example, when the Fourier Transform is used to represent the image in the frequency domain, the edge information in the image can be acquired by considering the high-frequency content in its Fourier transform. However, the problem with global transformations such as the Fourier transform is that they don't have spatial localization information necessary for matching. Spatial transformations such as the wavelet transform (Chui, 1992; Mallat, 1989) are more suitable for this purpose.

In general, feature-based techniques often require sophisticated image processing for feature extraction, and depend on the robustness of feature detection for reliable matching. Because they do not directly depend on the pixel values, they offer the potential to carry out multisensor image registration where the image data can be significantly different in the different modalities.

Spatial Transformations

The registration process, in simple terms, is a transformation of image coordinates. These spatial transformations are performed using *mapping functions*. In the traditional registration approach, the *mapping functions* are determined using the *control points* acquired from the matching process. This approach is commonly used in situations where an accurate set of control points can be determined *a priori*. In situations where the matching of the features is difficult, approaches like the *point mapping with feedback* can be used (Brown, 1992). In point mapping with feedback, the feature matching process and the determination of the optimal mapping function are accomplished simultaneously. The optimal spatial transformation between the images can be determined by an evaluation of all possible pairs of feature matches in an iterative fashion. Examples of these techniques include clustering (Stockman *et al.*, 1982), relaxation (Ranade and Rosenfeld, 1980), and matching of convex hull edges of the two sets (Goshtasby and Stockman, 1985). Furthermore, these techniques can be used as the last step in the feature matching process to check for global consistency of the matches (Li *et al.*, 1995). In this way, the mismatches occurred in the matching process can be rejected and thus can ensure correct registration.

Mapping Functions

Given n control points in two images of the same scene $[(X_i,Y_i), (x_i,y_i)]$, $i = 1, ..., n$, the *mapping functions* that register the images can be formulated in the following way:

$$X_i = f(x_i,y_i), \quad Y_i = g(x_i,y_i) \tag{5}$$

where (x_i,y_i) corresponds to a point in the *reference image* and (X_i,Y_i) corresponds to a point in the *sensed image*. The transformation function maps locations of points in the *reference image* to corresponding locations in the *sensed image*.

The functions used to align two images may be global or local. A global transformation is given by a single equation which optimally registers all the pixels in the two images. Local transformations map the images depending on the spatial location – the map is composed of several equations for each segment of the image that is considered (Goshtasby, 1988a). Local transformations are usually more accurate but are also computationally more demanding.

The two-dimensional *affine* transformation is frequently used to obtain a mapping between the two image coordinate systems. This transformation is sufficient to match two images with rigid-body distortion (Brown, 1992), and is composed of the Cartesian operations of scaling (s), translation (t_x,t_y), and rotation (θ) as follows:

$$X_i = t_x + s(x_i\cos\theta - y_i\sin\theta), \quad Y_i = t_y + s(x_i\sin\theta + y_i\cos\theta). \tag{6}$$

Other transformations such as the *general affine* transformation, the *projective* transformation, and the *perspective* transformation are also used to account for more general spatial distortions (Brown, 1992). But if the images have significant distortion between them, higher order polynomial fitting may be necessary. The parameters of the polynomials are determined by requiring that polynomials overlay the control points as closely as possible. This is accomplished by minimizing the sum of squared errors in the overlaying process. The least-squares technique that is used to determine the parameters averages a local geometry equally all over the image. For this reason, they cannot account for local geometric distortions such as sensor non-linearities, atmospheric conditions, and local three-dimensional scene features observed from differents viewpoints.

To overcome the problem of geometric distortions, several other types of local sensitive mapping functions have been proposed (Flusser, 1992; Goshtasby, 1986; Goshtasby, 1987; Goshtasby, 1988a; Goshtasby, 1988b). The most accurate results of registration of images with local distortions are obtained using the surface spline mapping functions. However, due to its extreme computing complexity, it is not suitable for practical applications. Flusser (1992) has proposed an approach for the determination of a mapping function sensitive to non-linear geometric distortions, with comparable accuracy but faster than surface spline mapping functions.

Registration Algorithms

In this Section we give an overview of some recent image registration algorithms. Our choice of algorithms is primarily dictated by the following criteria:

- applicability for registering remotely sensed image,
- shown good results on satellite images, and
- good potential for registering multisensor or multitemporal images.

The grouping of algorithms is based on the broad classification into area-based and feature-based techniques discussed in the previous section. However, some of the algorithms make use of both area-based and feature-based features in obtaining a registration.

Area-Based

- **Cideciyan *et al.*** (1992) have used the Fourier transform to decouple the translation from scaling and rotation, and the log-polar mapping of the Fourier magnitude to map the rotation and scaling into independent shifts. The cross-correlation of the log-polar Fourier magnitude is performed to obtain the optimal rotation and scaling parameters. The *sensed image* is rotated and scaled and a cross-correlation of this image transform is calculated to find the translation parameters. A local adaptive search strategy is used to find the optimal registration parameters in a parameter space.
- **Kher and Mitra** (1993) have performed the registration of SAR images, using a two-dimensional cepstrum technique. They separate the estimation of translation from that of rotation. The translation is estimated by inspecting the impulse train in the cepstrum domain, and the rotation parameters are obtained by evaluating the power spectrum of the two images for various angles of rotation. Morphological filters have been used to reduce the speckle noise.
- **Zheng and Chellapa** (1993) have used a Gabor wavelet transform to extract the feature points and propose a multiresolution strategy to match them. In their approach, image rotation is obtained by taking the difference of the estimated illuminant directions, under the assumption that illumination is constant during the time the image pair is taken or the images are taken at about the same time. In estimating the illuminant direction, they use techniques from shape-from-shading developed in the computer vision literature. The feature points, extracted by a Gabor wavelet decomposition (Manjun-

ath *et al.*, 1992), are matched with each other in a hierarchical manner using area correlation in the spatial domain. This method has been tested on different types of data and has shown good performance in registering images with large rotation and translation.

Feature-Based: Spatial Domain Features

- **Flusser and Suk** (1994) propose a method for registering images with general affine distortion. They also use the centers of gravity of the closed-boundary regions as control points. Each region extracted in the images is represented by affine-invariant moment-based features which describe the shape of the regions. The correspondence is established in the following way: (1) Three pairs of the most likely corresponding regions are found using the four-dimensional Euclidean feature space (moments up to the fourth order), (2) the parameters of the affine transformation are calculated and the sensed image is brought into alignment with the reference image, and (3) region-to-region correspondence is then established by the nearest-neighbor rule. The performance of their algorithm is demonstrated by registering SPOT and Landsat TM images taken in different years.
- **Goshtasby *et al.*** (1986) have used the centers of gravity of the closed-boundary regions as control points, and used a clustering technique for determining the corresponding control points in the images. The corresponding regions are refined so that the centers of gravity of the regions can correspond to each other more accurately. The authors observe that, for images with significant geometric distortions, features such as line intersections and line ends should be used instead of the centers of gravity of the regions.
- **Li *et al.*** (1995) have used region boundaries and other strong edges as matching primitives. Chain-code correlation and other shape similarity criteria such as moments are used to match closed contours. For the open contours, salient segments such as corners are detected first and then used in the matching process. An active contour method is used to match optical images with radar images. A consistency checking is conducted in the transformation parameter space to eliminate some false matches that occurred in the matching process. The authors have tested their method on registering different types of data such as Landsat-TM, SPOT, and Seasat SAR.
- **Ton and Jain** (1989) have used the centroids of objects, which are very visible in Landsat images and frequently appear in the area surveyed, as control points. A relaxation algorithm is used to match the control points. The method assumes that at least half of the control points are present in both images. The method is tested registering multitemporal images.
- **Toth and Schenk** (1992) have proposed a similar scheme to Li *et al.* (1995) where edges are extracted and matched in two stages using shape attributes. Application-specific constraints are used to resolve the ambiguity of multiple matches. Their method is used in registering map and optical images, and multispectral images.
- **Ventura *et al.*** (1990) have reported a registration method which is based on finding a match between objects with a limited set of proprieties and relations (shape descriptors and positional relations). For each pair of objects in the two images, the difference of the attribute values are computed and partitioned in labeled sets (small, very small, high, low, etc.). Then a Multi-Value Logical Tree (MVLT) decides if the pair of objects are "similar," "different," or "quasi-similar" based on the values of the labels. They have used geometrical relations to identify corresponding features to solve the problem where a feature in one image has N similar features in the other image. Experimental results in registering Landsat TM images with topographic maps are provided.
- **Wu and Maitre** (1990) report a contour-based method for registering SAR images and optical images. They have used long coastal lines, which have a high visual contrast in both images, as primary features to match the two images. To overcome the problem of speckle noise, they propose a multiresolution approach to detect the contour lines in the radar images: edges detected at the lower resolution level are used as heuristic information for the next higher resolution level. On the other hand, the contour lines in the optical image are detected using the split and merge algorithm. The inflection points along the contours are detected, and a combinatorial search is made to match these points.

Feature-Based: Transform Domain Features

- **Chandra** (1992) uses the Rationalized Haar transform (RHT) to extract feature points in the images. Windows in the images are represented by a predetermined number of highest RHT coefficients, which are correlated to determine the matching points. These coefficients are expected to provide a set of features which are efficient for registration purposes. Experimental results in registering multispectral images are provided.
- **Djamdji *et al.*** (1993) describe an image registration approach based on multiresolution analysis of images which is obtained by means of a discrete wavelet transform (DWT). The parameters of the polynomial transformation are estimated in an iterative way over the different levels in the transform domain. However, detailed information about the matching process is not provided in the paper. Experimental results on images from different sensors with small translations and rotations are provided.
- **Le Moigne** (1994) also uses a wavelet transform. Following a multiresolution decomposition of the data, characteristic features are extracted from the wavelet coefficients. The rotational transformation is computed in an iterative way over the different levels in the transform domain. Experimental results using rotated TM images are provided. A generalization of this method for registering multisensor images is also suggested.

A Comparative Study

A comparative study of the various registration methods is presented in Table 1 and Table 2. The following information is used in making the comparisons: (1) the modality of sensors (single or multisensor); this informs if the method has been tested on registering images from different sensors (multi) or from the same sensor (single); (2) type of sensor used in the experiments; (3) amount of overlap between the images tolerated by the method; (4) amount of test data on which results are shown; (5) types of features used in the matching process; (6) matching techniques used; (7) type of spatial transformations; and (8) observations concerning the limitations and merits of the methodologies.

Most of the methods described have presented subpixel accuracy. Information about the time complexity of the algorithms could not be obtained for all methods, so this information is not reported here.

Discussion

In the context of satellite image data, one can make the following observations:

- Images are from different sensors and usually have different spatial resolution;
- Images have different spectral characteristics, so that contrast information is different for the same imaged object;
- Radar images have speckle noise which affects information extraction; and
- Images taken at different times or under different conditions present changes that can affect the matching process.

Some of these issues are addressed in the registration algorithms discussed in the previous section. For example, in order to overcome the problem of different spatial resolution, most of the algorithms resample the images to bring back the two sets of data to the same resolution. Another alternative is to use attributes invariant to scale changes. Rignot *et al.* (1991) have suggested that the input data should be corrected for geometric distortions and geocoded onto a preselected grid common to all sensors, and resampled to the same pixel size for the registration process to obtain good results.

TABLE 1. REGISTRATION METHODS COMPARISON—PART I. THE WORDS SMALL, MEDIUM, AND LARGE MEAN THAT THE IMAGES CAN BE OVERLAPPED LESS THAN 30 PERCENT, 50 PERCENT, AND GREATER THAN 50 PERCENT, RESPECTIVELY

Reference	Sensor	Data types	Overlap	Feature	Feature Matching
Chandra (1992)	single	optical sensor	large	wavelet transform (highest RHT coefficients)	cross-correlation
Cideciyan *et al.* (1992)	single	visual camera	large	area-based	cross-correlation
Djamdji *et al.* (1993)	multi	SPOT, TM Landsat-MSS	large	local maxima of wavelet coefficients	multiresolution strategy
Flusser/Suk (1994)	multi	SPOT, TM	medium	closed-boundary (centroids)	shape similarity, nearest neighbor rule
Goshtasby *et al.* (1986)	multi	HCMM, TMS Landsat2-MSS	medium	closed-boundary (centroids)	clustering
Kher/Mitra (1993)	single	SAR images	large	area-based	analysis in the Fourier domain and Cepstrum domain
Le Moigne (1994)	single	Landsat-TM, visual camera	large	maxima of wavelet coefficients	cross-correlation, multiresolution strategy
Li *et al.* (1995)	multi	Landsat-TM Seasat, SPOT	medium	closed-boundary, salient contours	chain code correlation, shape similarity
Ton/Jain (1989)	single	Landsat-TM	medium	regions (centroids)	relaxation
Toth/Schenk (1992)	multi	Landsat-TM map	medium	curvature changes	shape similarity
Ventura *et al.* (1990)	multi	Landsat5-TM SPOT, map	medium	centroids of small regions or contour points or large regions	structural matching using Multi Value Logical Tree
Wu/Maitre (1990)	multi	Seasat SAR SPOT	medium	inflection points on the edge curve	multiresolution strategy, point matching
Zheng/Chellapa (1993)	single	visual camera optical sensor	small	Gabor wavelet transform	correlation coefficient, multiresolution strategy

TABLE 2. REGISTRATION METHODS COMPARISON – PART II. THE SYMBOLS R, T, AND S ARE ABBREVIATIONS OF ROTATION, TRANSLATION, AND SCALING, RESPECTIVELY.

Reference	Number of examples	Transformation	Merits/Limitations
Chandra (1992)	1	—	– comparable performance to the area correlation methods, even in the presence of noise.
Cideciyan *et al.* (1992)	6	affine (RTS)	– good performance on noisy images. – region of interest must be visible completely in both images.
Djamdji *et al.* (1993)	5	polynomial	– large amount of disk space needed for the processing.
Flusser/Suk (1994)	1	general affine	– images with distinctive features are used.
Goshtasby (1986)	2	affine (RTS) polynomial	– images with distinctive features are used.
Kher/Mitra (1993)	1	affine (RT)	– region of interest must be visible completely in both images. – noise-tolerant and computationally efficient.
Le Moigne (1994)	2	affine (R)	– images with only rotational difference are used.
Li *et al.* (1995)	6	affine (RTS)	– tested on different multisensor images. – well-defined, strong contours must be detected from the optical image to detect edges in the radar images. – displacement between the radar and optical images must be less than 5 pixels (active contour model approach).
Ton/Jain (1989)	3	affine (RT)	– images with specific features such as water regions and, oil and gas pad are used.
Toth/Schenk (1992)	3	—	– specific features (map) are used: water regions, agricultural regions or roads.
Ventura *et al.* (1990)	1	polynomial	– can be adapted to a given application. – flexibility in the choice of the attributes to match. – the definition of the logical rules and the choice of the attributes must be supported by the expert's advice.
Wu/Maitre (1990)	1	affine (RTS)	– specific features are used: coastal lines.
Zheng/Chellapa (1993)	8	affine (RTS)	– tolerate large rotation and translation.

In order to overcome the problem of speckle noise in radar images, two main methods have been suggested: (1) noise removal techniques before feature extraction, and (2) incorporating multiresolution techniques into the feature extraction or matching processes. Various techniques for the speckle removal such as morphological filters, geometric filters, and statistic filters have been used. Because noise removal usually results in blurred edges, a comparative study about the performance of these filters could aid in the selection of the most suitable filter for this task. There are other approaches which use digital elevation model (DEM) as ancillary data for registration between radar images and optical images (Rignot *et al.*, 1991; Takeuchi, 1993).

In summary, multisensor image registration is a difficult problem, and it is unlikely that a single algorithm will be developed to work satisfactorily on all types of data. A possible solution is to integrate multiple registration algorithms into a rule-based artificial intelligence system so that appropriate methods for any given set of multisensor data can be automatically selected. Such an expert system would use information about the sensor specifications, ancillary data (such as DEM), and the nature of the physical phenomena represented in the image and image distortions. A prototype system which integrates some of the algorithms described in this paper is currently under development at UCSB.

Acknowledgments

This work was partially supported by a fellowship to Leila M.G. Fonseca from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) of the secretary for science and technology of Brazil, by a grant from NASA (award number:NAGW 3951) and the UC MICRO program. The authors would like to thank Dr. George Campbell for providing the registration result shown in Figure 1.

References

Ballard, D.H., and C.M., Brown, 1982. *Computer Vision*, Prentice-hall, Englewood Cliffs, New Jersey.

Barnea, D.I., and H.F. Silverman, 1972. A class of algorithms for fast digital image registration, *IEEE Trans. Comput.*, 21(2):179–186.

Brown, L.G., 1992. A survey of image registration techniques, *Computing Surveys*, 24(4):325–376.

Canny, J.F., 1986. A computational approach to edge detection, *IEEE Trans. on Pattern Anal. and Machine Intell.*, 8(6):679–698.

Carper T.W., T.M. Lillesand, and R.W. Kiefer, 1990. The use of intensity-hue-saturation transformation for merging panchromatic and multispectral image data, *Photogrammetric Engineering & Remote Sensing*, 56(4):459–467.

Chandra, D.V.S., 1992. Fast registration of multispectral images using orthogonal transform features, *Proceedings of the 35th Midwest Symposium on Circuits and systems*, Washington, D.C., 1: 703–706.

Chui, C., 1992. *An Introduction to Wavelets*, (C. Chui, editor), Academic Press, New York.

Cideciyan, A.V., S.G. Jacobson, C.M. Kemp, R.W. Knighton, and J.H. Nagel, 1992. Registration of high resolution images of the retina, *Proceedings of the SPIE – The International Society for Optical Engineering*, Medical Imaging VI: Image Processing, 1652:310–322.

Djamdji, J.P., A. Bijaoui, and R. Maniere, 1993. Geometrical registration of images: the multiresolution approach, *Photogrammetric Engineering & Remote Sensing*, 59(5):645–653.

Flusser, J., 1992. An adaptive method for image registration, *Pattern Recognition*, 25(1):45–54.

Flusser, J., and T. Suk, 1994. A Moment-based approach to registration of images with affine geometric distortion, *IEEE Transactions on Geoscience and Remote Sensing*, 32(2):382–387.

Goshtasby, A., 1986. Piecewise linear mapping functions for image registration, *Pattern Recognition*, 19(6):459–466.

———, 1987. Piecewise cubic mapping functions for image registration, *Pattern Recognition*, 20(5):525–533.

———, 1988a. Registration of Images with geometric distortions, *IEEE Trans. on Geoscience and Remote Sensing*, 26(1):60–64.

———, 1988b. Image registration by local approximation methods, *Image and Vision Computing*, 6(4):255–261.

Goshtasby, A., and G.C. Stockman, 1985. Point pattern matching using convex hull edges, *IEEE Trans. Syst. Man Cybernetics*, 15(5): 631–637.

Goshtasby, A., G.C. Stockman, and C.V. Page, 1986. A region-based approach to digital image registration with subpixel accuracy, *IEEE Trans. on Geoscience and Remote Sensing*, 24(3):390–399.

Kher, A., and S. Mitra, 1993. Registration of noisy SAR imagery using morphological feature extractor and 2-D cepstrum, *Proceedings of the SPIE – The International Society for Optical Engineering*, Applications of Digital Image Processing XV, San Diego, California, 1771:281–291.

Le Moigne, J., 1994. Parallel registration of multi-sensor remotely sensed imagery using wavelet coefficients, *Proceedings of the SPIE – The International Society for Optical Engineering, Orlando*, Florida, 2242:432–443.

Li, H., 1993. *Multi-Sensor Image Registration and Fusion*, Ph.D. thesis, University of California, Santa Barbara.

Li, H., B.S. Manjunath, and S.K. Mitra, 1995. An approach to multisensor image registration, *IEEE Transactions on Image Processing*, 4(3):320–334.

Mallat, S.G,, 1989. Multifrequency channel decompositions of images and wavelet models, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(12):2091–2110.

Manjunath, B.S., R. Chellappa, and C. Malsburg, 1992. A feature based approach to face recognition, *Proc. IEEE Conf. Comput. Vision Pattern Recognition*, Champaign, Illinois, pp.373–378.

Marr, D., and E. Hildreth, 1980. Theory of edge detection. *Proceedings of Royal Society of London*, B207:187–217.

Melack, J.M., L.L. Hess, and S. Sippel, 1994. Remote sensing of lakes and floodplains in the Amazon Basin, *Remote Sensing Reviews*, 10:127–142.

Ranade, S., and A. Rosenfeld, 1980. Point pattern matching by relaxation, *Pattern Recognition*, 12:269–275.

Rignot, E.J.M., R. Kowk, J.C. Curlander, and S.S. Pang, 1991. Automated multisensor registration: Requirements and techniques, *Photogrammetric Engineering & Remote Sensing*, 57(8):1029–1038.

Stockman, G., S. Kopstein, and S. Benett, 1982. Matching images to models for registration and object detection via clustering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4(3): 229–241.

Takeuchi, S., 1993. Image registration between SAR and TM data using DEM and slant range information, *Proceedings of IGARSS – IEEE International Geoscience and Remote Sensing Symposium*, Tokyo, Japan, 3:1351–1353.

Ton, J., and A.K. Jain, 1989. Registering Landsat images by point matching, *IEEE Trans. on Geoscience and Remote Sensing*, 27(5):642–651.

Toth, C.K., and T. Schenk, 1992. Feature-based matching for automatic image registration, *ITC Journal*, (1):40–46.

Ventura, A.D., A. Rampini, and R. Schettini, 1990. Image registration by the recognition of corresponding structures, *IEEE Trans. on Geoscience and Remote Sensing*, 28(3):305–314.

Wu, Y., and H. Maitre, 1990. A multiresolution approach for registration of a SPOT image and a SAR image, *Proceedings of International Geoscience and Remote Sensing Symposium*, pp. 635–638.

Zheng, Q., and R. Chellappa, 1993. A computational vision approach to image registration, *IEEE Transactions on Image Processing*, 2(3):311–326.

(Received 29 December 1994; accepted 13 March 1995; revised 4 May 1995)