

Automated subpixel image registration of remotely sensed imagery

by M. D. Pritt

An algorithm is described for the automated registration of remotely sensed imagery that registers 6000 × 6000-pixel images in 8–18 minutes on an IBM RISC System/6000® workstation. The resulting registration is accurate to the subpixel level even in the presence of noise and large areas of change in the images. It is shown that the registration-mapping function for parallel projections has the form $F(x, y) = A(x, y) + h(x, y)e$, where $A(x, y)$ is an affine transformation, $h(x, y)$ is a function that depends on the topographic heights, and e is a vector that defines the epipolar lines. The algorithm determines the parameters of this equation using only the image data, without knowledge of the viewing orientations or scene point coordinates. The search for match points is then a one-dimensional search along the epipolar lines, which greatly increases the speed and accuracy of the registration.

Introduction

Image registration is the process of geometrically aligning two images of the same scene. It has applications in such diverse areas as robot vision [1], guidance systems [2], motion detection [3], medical imaging [4], and manufacturing [5]. It also has important applications in the analysis of remotely sensed imagery, such as change detection, topographic mapping, and the alignment of images with maps. (See for example [6] or [7].)

Although most current methods of registering remotely sensed imagery require human involvement, there is a need for automated techniques. According to [8], "On the average, an up-to-date Landsat image can be obtained for most areas on the Earth every 16 days. Since any newly acquired Landsat image needs to be registered with other images or maps before it can be useful, there is a great need to automate this process." Another author states, "Manual handling of data for change detection . . . is a formidable task There is a definite need for a change detector which will automatically correlate and compare two sets of imagery taken of the same area at different

©Copyright 1994 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

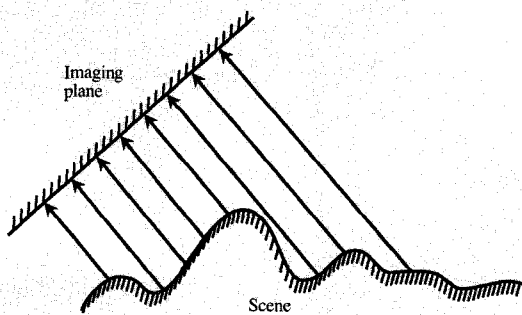


Figure 1

Parallel projection imaging geometry.

times and display the changes and their locations to the interpreter" [9]. Although much work has been done on automated registration, most relevant techniques apply only to simplistic images, and more are required for solving real-world problems [10]. Even for the comprehensive image analysis system described in [11], "... manual selection of control points is still the best method for scene registration."

This paper describes an algorithm that has been developed for the registration of remotely sensed imagery without human intervention. The algorithm has the following properties:

- It has subpixel accuracy, which is necessary for applications such as change detection [12, 13].
- It registers 6000×6000 -pixel images in only minutes on an IBM RISC System/6000® workstation.
- It is reliable, producing accurate image registration in the presence of image noise, topographic height variation, and large areas of change.
- It uses only the data contained in the images. It does not require knowledge of imaging sensor orientations or three-dimensional scene point coordinates, for example.

The algorithm is based on two main assumptions about the imagery:

- The viewing orientations differ by no more than two or three degrees.
- The imaging geometry is a parallel projection.

The first assumption is valid for image pairs used for change detection analysis, since large separation angles cause false changes in the images due to perspective and

occlusion distortions. The second assumption, in which scene points are assumed to be projected to the imaging plane in parallel lines as shown in **Figure 1**, is valid when the imaging sensor is far from the scene and has a narrow field of view.

The algorithm uses two-dimensional cross-correlation to identify pairs of corresponding points in the images. This technique is accurate and works well in the presence of noise and varying conditions of illumination. Since it does not attempt to identify features, it is applicable to many types of images from satellites such as Landsat, SPOT, and ERS-1.

The innovative feature of the algorithm is its ability to exploit the epipolar constraint to increase the speed and accuracy of the registration. The epipolar constraint is used in the registration of stereo imagery, where the orientation of the sensors is known *a priori* or can be derived from knowledge of the three-dimensional coordinates of a few scene points [3, 14, 15]. When the sensor orientations are unknown and no three-dimensional information is available, the principle cannot in general be used. However, for the special case in which the imaging geometry is a parallel projection, we demonstrate that the data necessary for exploiting the epipolar constraint can be derived from the image data without knowledge of the viewing orientations or scene point coordinates.

After briefly outlining the main steps of image registration, we describe the cross-correlation technique used for identifying pairs of corresponding points. The explicit form of the registration-mapping function is then derived, which sheds light on the image-registration problem. We then present the main steps of the registration algorithm, explaining how the orientation of the epipolar lines is derived and exploited to increase the speed and accuracy of the registration. The implementation results and plans for further work are presented, followed by the Appendix, in which two applications of registered imagery are described: change detection and interferometry.

Image registration

There are many approaches to solving the problem of image registration. One approach seeks to minimize the mean square difference between the pixel intensity values in a pair of images through the use of gradients [16]. Most approaches, however, consist of the following three steps:

1. Identify a set of *match points*, which are pairs of points that correspond to the same point in the scene. (Match points are also referred to as *corresponding points*, *control points*, *conjugate points*, and *tie points* by various authors.
2. Compute the *registration-mapping function*, which is the function that assigns to each point (x, y) in the first

image its corresponding point $F(x, y)$ in the second image. Note that if $[(x, y), (u, v)]$ is a match point, then $(u, v) = F(x, y)$.

3. Resample the second image using the mapping function $F(x, y)$ to bring it into alignment with the first image.

Traditional registration techniques model the registration-mapping function as a first- or second-degree polynomial whose coefficients are computed by a least-squares fit to the match points. However, globally defined polynomials cannot accurately model the local distortions in the mapping function which result from topographic height variation [17, 18]. More general functions, such as locally defined polynomials or splines, are necessary.

The most difficult step of image registration is the accurate identification of match points. The traditional approach uses human assistance to identify these points. Automated methods follow one of two main approaches [10, 11]. *Area-based* approaches use correlation techniques to match the pixel intensity patterns of one image with those of the second image. *Feature-based* approaches seek to identify features in the intensity patterns, such as edges, corners, line intersections, and closed boundary regions. Feature-based methods are faster than area-based methods [10], but they tend to be suited for only special types of images. Area-based methods, on the other hand, tend to be more robust, but they fail for images taken from widely separated viewing angles, since perspective and occlusion distortions degrade the correlations [3, 11].

Cross-correlation

Since the registration algorithm described in this paper is designed for images taken from similar viewing orientations (where perspective and occlusion distortions are minimal), and since robustness and applicability to many types of imagery are important requirements, an area-based method based on two-dimensional cross-correlation is used for identifying match points. The high computational cost is minimized through the use of several constraints, including the epipolar constraint described in a later section.

The cross-correlation approach assumes that the complicated transformation that relates two images of the same scene is locally composed of plane translations. Small templates or patches from one image appear to be merely displaced in the second image without perspective or rotational distortions. While this assumption is not strictly true, it is nonetheless a valid approximation for images which have undergone only a small change in viewing orientation. For example, if an image undergoes a small planar rotation, small templates from the top and bottom edges of the image appear to shift horizontally, and templates from the right and left edges appear to shift vertically.

The two-dimensional cross-correlation approach proceeds as follows. A small $N \times N$ -pixel template $f_1(i, j)$ is extracted from the first image, centered at some point (a, b) , and a larger $M \times M$ -pixel template $f_2(i, j)$ is extracted from the second, centered at some point (c, d) . The value N should be large enough to provide good correlation peaks, but not so large that the templates contain significant rotational distortions. The value M should be chosen so that $(M - N)/2$ is the maximum anticipated displacement of the two templates. The two-dimensional cross-correlation function $R(u, v)$ is then defined by

$$R(u, v) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f_1(i, j) f_2(i + u, j + v) \quad (1)$$

for $0 \leq u, v \leq M - N$. The peak of the correlation, which may be determined with subpixel accuracy through quadratic interpolation of the correlation values, determines the actual match point. If the peak occurs at the point (u, v) , the match point is

$$[(a, b), (c + u - (M - N)/2, d + v - (M - N)/2)]. \quad (2)$$

The correlation peak is more pronounced if the means of the pixel intensities of the templates are zero. This can be accomplished by subtracting the means before calculating $R(u, v)$.

The correlations $R(u, v)$ should not be normalized, as in [19]. Featureless image regions yield high normalized correlations, but the peaks are broad and poorly defined. A better measure of the correlation is the sharpness of a peak, which can be measured quickly, albeit roughly, as follows. Let p be the height of the peak, and let q be the maximum height of the correlation surface at a fixed distance of 2 pixels from the location of the peak. Then the ratio q/p is a measure of the peak sharpness, with 0 indicating an ideal peak and 1 indicating no peak at all. If this ratio exceeds a preassigned threshold (e.g., 0.5), the match point should be rejected as unreliable.

Accelerating the cross-correlation computations

FFTs may be used to accelerate the calculation of the cross-correlations. The first template $f_1(i, j)$ must be placed in an $M \times M$ -pixel array and zero-padded to avoid end effects. Let $F_1(a, b)$ and $F_2(a, b)$ denote the two-dimensional Fourier transforms of $f_1(i, j)$ padded and $f_2(i, j)$, respectively. Then the cross-correlation function $R(u, v)$ is the inverse Fourier transform of the product $F_1(a, b) F_2^*(a, b)$, where $*$ denotes the complex conjugate. The correlation function thus can be computed using three two-dimensional FFTs of size $M \times M$.

There are two ways to further reduce the time spent computing the cross-correlations. The first method is to

minimize the number of match points that must be identified. This is accomplished with the epipolar constraint, as described in the next section. The second method is to keep the search distance $(M - N)/2$ as small as possible.

One way to minimize the search distance is to perform a multistage registration in two or more stages, in which a coarse registration is first performed on the images with a low resolution [20]. Since the low-resolution images are smaller, the search distances are smaller as well. By using the results of this coarse registration, the search distances for match points in the higher-resolution images can be constrained.

A second way to reduce the search distances, which is used by the current algorithm, is to use the first few match points that are identified to constrain the search for further match points. As more match points are found, they define an increasingly accurate registration of the image, which predicts the locations of further match points. This method is similar to the multistage approach, but it does not require the images to be resampled to a lower resolution.

Epipolar constraint

The image registration algorithm presented here adopts the *epipolar constraint* principle from the analysis of stereo imagery. A stereo imaging system consists of a pair of optical sensors with their viewing axes mutually parallel and separated by a horizontal distance known as the *baseline* [3, 10]. The viewing axes of the sensors are perpendicular to the baseline, and the image scanlines are parallel to the baseline. Since the displacement between the sensors is purely horizontal, the positions of corresponding points in the two images differ only in horizontal displacement. The search for match points is thus reduced to a one-dimensional search along the horizontal lines, referred to as the *epipolar lines*. Human binocular vision is an example of a stereo imaging system, and the human brain registers the disparate images from the two eyes using the epipolar constraint [21].

If the viewing axes of the imaging sensors are not parallel, the match points are still displaced along epipolar lines, although the lines may no longer be parallel. For parallel-projected images, however, the epipolar lines are parallel to one another for *any* relative orientation of the sensors. Expressed in mathematical terms, the form of the registration mapping function for parallel-projected images is

$$F(x, y) = A(x, y) + h(x, y)e, \quad (3)$$

where A is an affine transformation (a 2×2 matrix transformation followed by a translation), e is a fixed vector, and $h(x, y)$ is a scalar-valued function. The vector e determines the direction of the epipolar lines. The affine transformation A and the epipolar vector e depend on the

orientations of the two imaging planes of the sensor relative to the scene, while the function $h(x, y)$ depends on the topographic heights of the points of the scene.

Equation (3) can be derived as follows. Let f and g be the vector-valued functions which define the two images of the scene, in the sense that a point (u, v, w) in the scene maps to the point $f(u, v, w)$ in the first image and to the point $g(u, v, w)$ in the second image. Since the imaging sensors are assumed to be parallel projections, they can be modeled as coordinate systems in which points are projected to the XY plane by the natural projection $p(x, y, z) = (x, y)$. After a fixed coordinate system is erected at the scene, the orientation and position of the first sensor can be given by a coordinate transformation matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (4)$$

and a translation vector $(a_{14}, a_{24}, a_{34})^T$. The imaging function f thus has the following form:

$$f(u, v, w) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} + \begin{bmatrix} a_{14} \\ a_{24} \end{bmatrix}. \quad (5)$$

If we define the matrix S by

$$S = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (6)$$

and the vectors s and a by

$$s = \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}, \quad a = \begin{bmatrix} a_{14} \\ a_{24} \end{bmatrix}, \quad (7)$$

the function f can be written in the form

$$f(u, v, w) = S \begin{bmatrix} u \\ v \end{bmatrix} + ws + a. \quad (8)$$

Similarly, the second imaging function g can be written in the form

$$g(u, v, w) = T \begin{bmatrix} u \\ v \end{bmatrix} + wt + b, \quad (9)$$

where T is a 2×2 matrix and t and b are vectors.

Let $h(x, y)$ be the scalar-valued function which assigns to each point (x, y) in the first image the z -coordinate of the (non-occluded) point in the scene which maps to (x, y) under the function f . Note that $h(x, y)$ is *not* the topographic height function; the domain of $h(x, y)$ is the

image plane, while the domain of the topographic height function is the scene plane. Additionally, let (x, y) be a point in the first image, and let (u, v, w) be the non-occluded point in the scene which is mapped to (x, y) by the function $\mathbf{f}$. Since $\mathbf{f}(u, v, w) = (x, y)$ and $\mathbf{g}(u, v, w) = \mathbf{F}(x, y)$, the function $\mathbf{F}(x, y)$ can be determined by solving for the vector (u, v) in Equation (8) and substituting the resulting expression into Equation (9). Solving for this vector in Equation (8) gives

$$\begin{bmatrix} u \\ v \end{bmatrix} = S^{-1}\mathbf{f}(u, v, w) - wS^{-1}\mathbf{s} - S^{-1}\mathbf{a}. \quad (10)$$

Substituting this expression into Equation (9) yields

$$\mathbf{g}(u, v, w) = TS^{-1}\mathbf{f}(u, v, w) + (\mathbf{b} - TS^{-1}\mathbf{a}) + w(\mathbf{t} - TS^{-1}\mathbf{s}). \quad (11)$$

Since $w = h(x, y)$, Equation (11) is equivalent to Equation (3), where $A(x, y)$ is the affine transformation

$$A(x, y) = TS^{-1} \begin{bmatrix} x \\ y \end{bmatrix} + (\mathbf{b} - TS^{-1}\mathbf{a}), \quad (12)$$

and $\mathbf{e}$ is the epipolar vector $\mathbf{t} - TS^{-1}\mathbf{s}$.

The preceding derivation required the matrix S to be invertible. We now investigate the conditions under which this is the case. Since the matrix (4) is a coordinate transformation, its rows are orthonormal. In particular, the third row vector is the vector cross product of the first two rows, which implies $a_{33} = a_{11}a_{22} - a_{12}a_{21}$. The right-hand side of this identity is nonzero if and only if S^{-1} exists. Note that the third row vector of (4) is the line-of-sight vector from the scene to the imaging sensor relative to the scene coordinate system. Thus, the matrix S is invertible if and only if this line-of-sight vector is not horizontal relative to the scene—a quite reasonable assumption for remotely sensed imagery!

Equation (3) sheds light on the registration problem and completely characterizes the local distortions in registration caused by topographic height variation. Suppose, for example, the scene is planar. Then the height function $h(x, y)$ is constant, and the registration-mapping function is an affine transformation. The images can be registered by identifying at least three match points and performing a least-squares fit to the coefficients of the affine transformation. On the other hand, the local distortions due to height variation in nonplanar scenes are merely displacements along parallel lines. Furthermore, these displacements are directly proportional to the heights. **Figures 2(a) and 2(b)** show two images of a scene composed of a vertical cylinder on a planar square. **Figure 2(c)** shows the result of applying an affine transformation to the second image to register the planar squares. **Figure 2(d)** shows this transformed image

overlaid on the first image. The epipolar lines, which are shown as heavy black lines in **Figure 2(a)**, pass through corresponding points on the cylinder and show that the displacements of the points along the lines are proportional to their heights on the cylinder.

Determination of the registration-mapping function

The algorithm described in this paper determines the registration-mapping function $\mathbf{F}$ in four steps. The first step determines the affine transformation $A(x, y)$, and the second step determines the orientation of the epipolar vector $\mathbf{e}$. The third step determines the height function $h(x, y)|\mathbf{e}|$ on a grid of points, and the fourth step fits a surface to these grid points to determine the function $\mathbf{F}(x, y)$. The following paragraphs explain each step in more detail.

Step 1 A set of match points is identified using the cross-correlation technique described previously, and an iterative least-squares fit of the coefficients of the affine transformation $A(x, y)$ to these points is performed. Each iteration eliminates the match point with the largest fitting error until the maximum fitting error of the remaining match points falls below a preassigned threshold. The match points which remain correspond to points in the scene which lie in the same plane, thus making possible the calculation of $A(x, y)$.

Step 2 Each match point $[(x, y), (u, v)]$ that was eliminated by the least-squares iteration of Step 1 is replaced with the vector $(u, v) - A(x, y)$. A least-squares fit of these vectors to a line passing through the origin is then performed, yielding the direction of the epipolar vector $\mathbf{e}$. [The magnitude of $\mathbf{e}$, which is not necessary for the determination of $\mathbf{F}(x, y)$, cannot be determined unless the orientation of the image planes relative to the scene is explicitly known.]

Step 3 At each point (x, y) of a grid conceptually placed over the first image, the cross-correlation surface is computed for templates centered at (x, y) in the first image and $A(x, y)$ in the second image. The displacement $h(x, y)|\mathbf{e}|$ is then determined by searching for a peak along the epipolar line, whose direction was determined in Step 2. To further reduce the probability of obtaining a false match point, a smoothness constraint is enforced which does not allow the absolute values of the second derivatives of the function $h(x, y)|\mathbf{e}|$ to exceed a preassigned threshold.

Step 4 The definition of the scalar-valued function $h(x, y)|\mathbf{e}|$ is extended to the entire image by interpolating a surface through the displacements determined in Step 3.

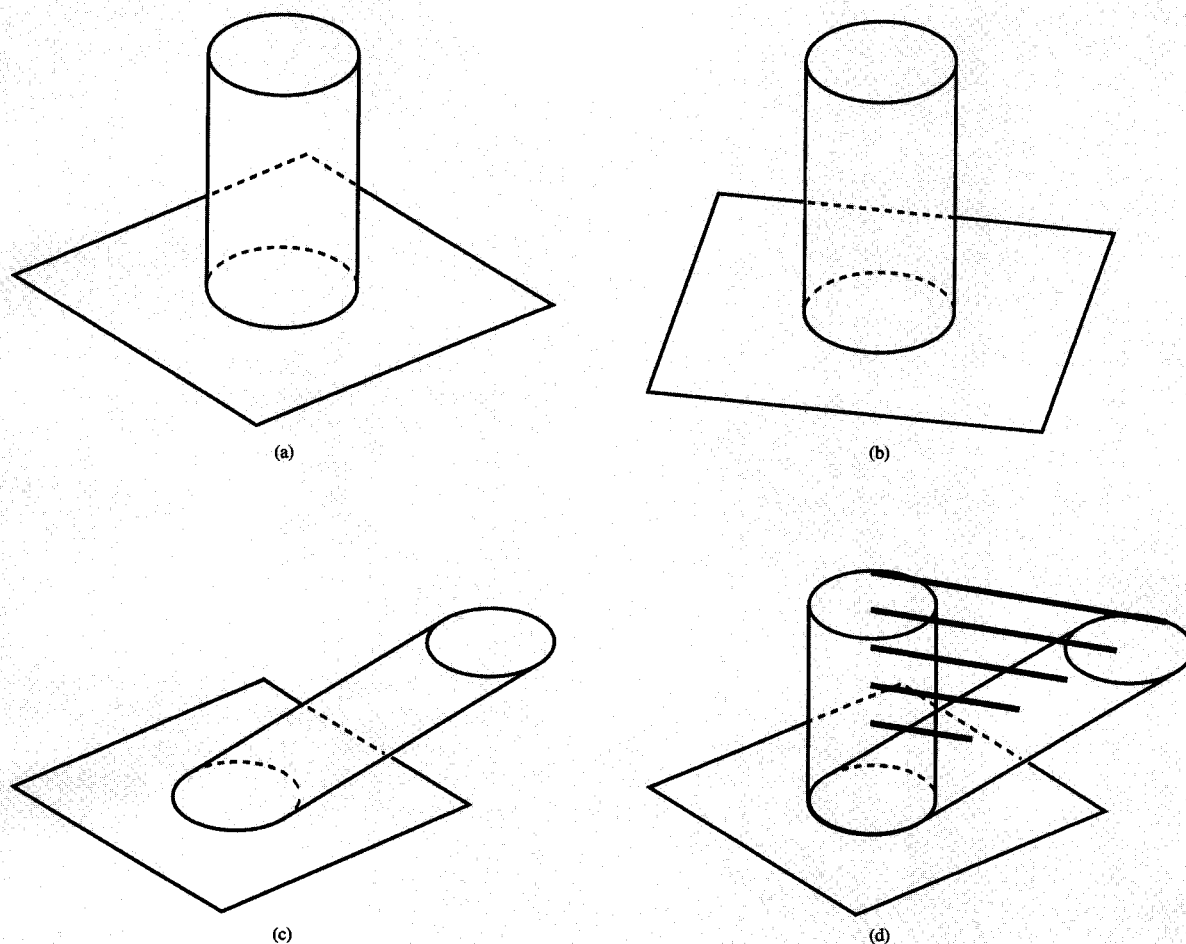


Figure 2

(a) An image of a vertical cylinder. (b) A second image of the cylinder from a different viewing angle. (c) The second image after application of an affine transformation. (d) The transformed image overlaid on the first image, along with the epipolar lines.

(The current implementation uses Hermite polynomials [22], although other interpolating functions could be used instead.) This defines the registration-mapping function $F(x, y)$.

All registration algorithms that are based on match point identification must deal with the problem of false match points, which degrade the registration accuracy. (See [14], for example.) The algorithm described in this paper eliminates false match points in Steps 1 and 2 through the use of least-squares techniques. False match points are eliminated in Step 3 by application of the epipolar and smoothness constraints. **Figure 3** illustrates the effectiveness of the epipolar constraint. A 30×30 -pixel

correlation surface, which has been smoothed and scaled by an exponential function to accentuate the peaks, is shown with the epipolar line overlaid in white. The peak over which this line passes corresponds to the true match point, but there are taller peaks which do not lie on the line. Without the epipolar constraint, the match point identification algorithm would choose one of these taller peaks, thereby yielding a false match point.

When the viewing angles are very close together, or when there is little height variation in the scene, the contribution of the second term $h(x, y)e$ to the registration-mapping function is negligible. This condition is reflected in there being only a few match points

eliminated by the iteration of Step 1, or not enough vectors to perform a reliable least-squares fit in Step 2, rendering Step 3 unnecessary. This results in a much shorter registration time.

Implementation

The registration algorithm has been coded in C on an IBM RISC System/6000® Model 560 workstation. The FFTs were computed using the IBM Engineering and Scientific Subroutine Library (ESSL) [23]. For a pair of 6000×6000 -pixel images, 900 cross-correlations were computed for the least-squares iteration of Step 1, and 3600 were computed in Step 3. The cross-correlation template sizes were 60×60 pixels, the search distance was 30 pixels, and the peak sharpness threshold was 0.5.

The 6000×6000 -pixel images were registered in less than eight minutes when the scenes were planar or when the viewing angles were sufficiently close together. Otherwise, the registration required 18 minutes. These times included four minutes for the image-resampling step. The registration accuracy was estimated to be within 0.5 pixel. This estimate was obtained by shifting the registered images various fractions of a pixel until the change detection images (see the Appendix) were noticeably degraded.

Before the epipolar constraint was implemented, the registration algorithm fitted the match points to two-dimensional Hermite polynomials. To achieve a satisfactory level of registration accuracy, it was necessary to identify a very large number of match points (about 16–25 times as many as required by the later algorithm using the epipolar constraint). This resulted in registration times of the order of hours rather than minutes, even for planar scenes. Implementation of the epipolar constraint has resulted in a dramatic decrease in processing time.

Current and further work

Current work is aimed at implementing the algorithm on a parallel processor computer. The goal is to yield registration times of less than one minute. Most of the time is spent computing the cross-correlation surfaces, which could be computed in parallel [24, 25].

Further work will be aimed at more carefully measuring the registration accuracy, since at least one study suggests that accuracies of 0.2 pixel or less are required for effective change detection for certain types of scenes [13]. The assumption of similar viewing angles will also be relaxed by implementing a correction for perspective distortion and rotation in the correlation process, as suggested in [3] and [14].

Summary

A complete algorithm for the automated registration of remotely sensed imagery has been presented. This

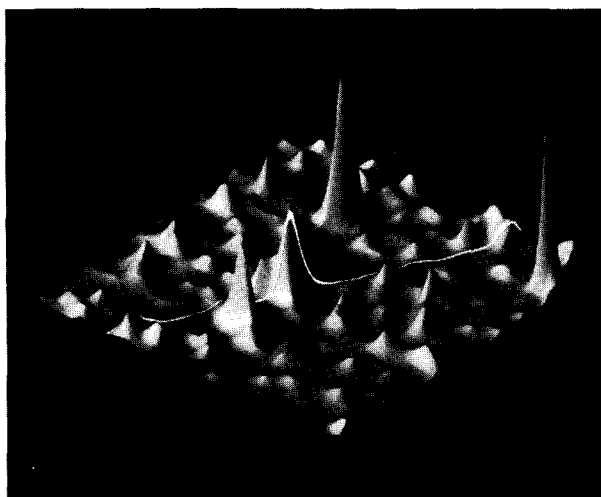


Figure 3

A cross-correlation surface with an epipolar line overlaid on it.

algorithm is relatively fast, reliable, and accurate to the subpixel level, even for large, noisy images. It uses the epipolar constraint to dramatically improve its speed and accuracy. The innovative feature of the algorithm is its ability to derive the data necessary for exploiting the epipolar constraint solely from the image data. Knowledge of viewing orientations and three-dimensional scene point coordinates is not needed.

The modeling equation (3) used by the algorithm greatly clarifies the image-registration problem. It assumes only that the imaging geometry is a parallel projection, a valid assumption for many types of remotely sensed imagery. It presupposes no restrictions on the relative viewing orientations (although the assumption of similar viewing orientations is required by the cross-correlation method for identifying match points), and it can be used with almost any method of match point identification.

Appendix

This appendix briefly describes two applications of registered imagery: change detection and interferometry.

• Change detection

Change detection indicates regions of the scene which have changed between the imaging times. It is used for assessing and monitoring such processes as urban development [26] and changes in vegetation [9, 27, 28]. There are a number of techniques for detecting the differences between two registered images, the simplest of which subtracts the images pixel by pixel [12]. Normalized cross-correlation is another technique which is a natural measure of change for images registered by the algorithm

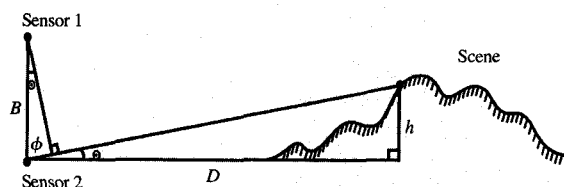


Figure 4

Topographic mapping by means of interferometry.

described in this paper. At each pixel the two-dimensional normalized cross-correlation coefficient of small templates from the two images is computed. The values of this coefficient are scaled to pixel intensity values to produce a change-detection image which can be overlaid on the registered images. High pixel intensities represent regions of little or no change, while low intensities represent regions that have changed.

• Interferometry

Topographic mapping constructs a three-dimensional model of the scene using the disparity information contained in a pair of registered images. The traditional method of topographic mapping is stereo imaging, which derives the three-dimensional information from the displacements of match points along the epipolar lines. A more accurate method of topographic mapping is *interferometry* [29–35], which derives the three-dimensional information from the phase components of complex-valued imaging data. The phase components measure the distances from the sensor to the points of the scene modulo the wavelength λ of the imaging signal. The differences in phase between pairs of corresponding points in the registered images create a pattern of interference fringes which indicates the topographic heights. The similar triangles of **Figure 4** show that the relationship between the height h of a point in the scene and the phase difference ϕ is given by the equation

$$\frac{h}{D} = \frac{\phi}{\sqrt{B^2 - \phi^2}}, \quad (\text{A1})$$

or $h = (D/\sqrt{B^2 - \phi^2})\phi \approx (D/B)\phi$, where B is the baseline distance and D is the distance to the scene. (It is assumed that D is much larger than B , so the phase difference ϕ is approximately the quantity shown in the figure.) The topographic heights h are thus directly proportional to the measured phase differences ϕ

modulo λ . [The orientations of the imaging sensors can be effectively rotated to the orientation shown in the figure by subtracting the dominant spatial frequency of the interference fringes. This frequency can be computed by finding the peak of the two-dimensional power spectral density of the fringes. In addition, the right-hand side of Equation (A1) must be multiplied by a suitable scale factor, and the phase differences must be resampled.]

Since the phase differences ϕ (and the heights h) are only known modulo λ , they must be “unwrapped.” Several techniques for two-dimensional phase unwrapping have been described in the literature [30, 32, 36], the most robust of which uses a least-squares approach to determine the function $f(x, y)$ which minimizes

$$\iint \left[\left(\frac{\partial f}{\partial x} - \frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} - \frac{\partial \phi}{\partial y} \right)^2 \right] dx dy. \quad (\text{A2})$$

The principle is that the partial derivatives of the wrapped phases $\phi(x, y)$ and the unwrapped phases $f(x, y)$ should agree. It is demonstrated in [37] that the solution to a discretization of (A2) is given by solving a discretization of Poisson’s equation,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \rho(x, y), \quad (\text{A3})$$

with Neumann boundary conditions. This equation can be solved by the iterative technique of simultaneous overrelaxation (SOR) described in Section 17.5 of [38]. For images larger than 1000×1000 pixels, a more practical method is based on the use of the cosine transform and is described in Section 17.4 of [38]. A more recent technique is based on the use of FFTs and is described in [39].

Acknowledgments

The author would like to acknowledge helpful discussions with Daniel Bondy, Dwight Divine, Charles Guarino, William Rice, and Jerome Shipman. David Cyganski and Richard Vaz of the Worcester Polytechnic Institute suggested the use of the epipolar constraint. Mr. Shipman offered helpful comments for improving the manuscript.

RISC System/6000 is a registered trademark of International Business Machines Corporation.

References

1. B. K. P. Horn, *Robot Vision*, MIT Press, Cambridge, MA, 1986, Ch. 13.
2. K. T. Constantines and G. D. Shiflett, “Increasing Correlation Matcher Space-Bandwidth Product via a Mosaic Field-of-View,” *Applications of Digital Image Processing XIII*, SPIE Vol. 1349, International Society for Optical Engineering, Bellingham, WA, 1990, pp. 526–536.
3. J. C. Elsey, “A Technique for Image Registration and Motion Detection,” *Conference Record of the Twenty-*

- third Asilomar Conference on Signals, Systems and Computers, Vol. 2 (IEEE Cat. No. 89-CH2836-5), 1989, pp. 829-833.
4. G. Q. Maguire, Jr., M. E. Noz, H. Rusinek, J. Jaeger, E. L. Kramer, J. J. Sanger, and G. Smith, "Graphics Applied to Medical Image Registration," *IEEE Comput. Graph. Appl.* **11**, No. 2, 20-28 (1991).
5. I. J. Cox, J. B. Kruskal, and D. A. Wallach, "Predicting and Estimating the Accuracy of a Subpixel Registration Algorithm," *IEEE Trans. Pattern Anal. Machine Intell.* **12**, No. 8, 721-734 (1990).
6. Maria Teresa Pareschi and Ralph Bernstein, "Modeling and Image Processing for Visualization of Volcanic Mapping," *IBM J. Res. Develop.* **33**, No. 4, 406-416 (1989).
7. F. Barberi, R. Bernstein, M. T. Pareschi, and R. Santacroce, "Displaying Morphological and Lithological Maps: A Numerically Intensive Computing and Visualization Application," *IBM J. Res. Develop.* **35**, No. 1/2, 78-87 (1991).
8. J. Ton and A. K. Jain, "Registering Landsat Images by Point Matching," *IEEE Trans. Geosci. Remote Sensing* **27**, No. 5, 642 (1989).
9. A. Singh, "Digital Change Detection Techniques Using Remotely-Sensed Data," *Int. J. Remote Sensing* **10**, No. 6, 989 (1989).
10. U. R. Dhond and J. K. Aggarwal, "Structure from Stereo—A Review," *IEEE Trans. Syst. Man Cybernet.* **19**, No. 6, 1489-1510 (1989).
11. Y. C. Hsieh, F. Perlant, and D. M. McKeown, "Recovering 3D Information from Complex Aerial Imagery," *Proceedings of the 10th International Conference on Pattern Recognition*, Vol. 1, IEEE, Piscataway, NJ, 1990, p. 145.
12. P. Gong, E. F. Ledrew, and J. R. Miller, "Registration-Noise Reduction in Difference Images for Change Detection," *Int. J. Remote Sensing* **13**, No. 4, 773-779 (1992).
13. J. R. G. Townshend, C. O. Justice, C. Gurney, and J. McManus, "The Impact of Misregistration on Change Detection," *IEEE Trans. Geosci. Remote Sensing* **30**, No. 5, 1054-1060 (1992).
14. Y. C. Shah, R. Chapman, and R. B. Mahani, "A New Technique to Extract Range Information from Stereo Images," *IEEE Trans. Pattern Anal. Machine Intell.* **11**, No. 7, 768-773 (1989).
15. D. Sherman and S. Peleg, "Stereo by Incremental Matching of Contours," *IEEE Trans. Pattern Anal. Machine Intell.* **12**, No. 11, 1102-1106 (1990).
16. I. O. G. Davies and M. Pedley, "Image Registration Algorithms," *IEE Colloquium on Applications of Motion Compensation (Digest No. 128)*, London, UK, 1990, pp. 2/1-4.
17. A. Goshtasby, "Image Registration by Local Approximation Methods," *Image Vision Comput.* **6**, No. 4, 255-261 (1988).
18. J. Flusser, "Adaptive Method for Image Registration," *Pattern Recogn.* **25**, No. 1, 45-54 (1992).
19. A. D. Mandalia, R. Sudhakar, K. Ganesan, and F. Hamano, "Low-Level and High-Level Correlation for Image Registration," *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, IEEE, Piscataway, NJ, 1990, pp. 206-208.
20. M. Sone, M. Sakauchi, and M. Onoe, "A Note on Image Registration Based on Geometric Correspondence Between Two Levels of Extended Pyramid Structure," *Syst. Comput. Jpn.* **20**, No. 12, 1-13 (1989).
21. J. P. Frisby, *Seeing*, Oxford University Press, New York, 1980, Ch. 7.
22. W. K. Giloi, *Interactive Computer Graphics*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1978, pp. 123-148.
23. *IBM Engineering and Scientific Subroutine Library Guide and Reference*, Order No. SC23-0184, available through IBM branch offices.
24. B. M. Madhusudan, "Parallel Implementation of Template Matching on PARAM," *Applications of Transputers 3, Proceedings of the Third International Conference on Applications of Transputers*, Glasgow, UK, August 28-30, 1991, pp. 705-708.
25. L. J. Siegel, H. J. Siegel, and A. E. Feather, "Parallel Processing Approaches to Image Correlation," *IEEE Trans. Computers* **C-31**, No. 3, 208-217 (1982).
26. N. A. Quarmby, "Monitoring Urban Land Cover Changes at the Urban Fringe from SPOT HRV Imagery in Southeast England," *Int. J. Remote Sensing* **10**, No. 6, 953-963 (1989).
27. J. Schnurr, "Building a Monitoring System Based on Satellite Data to Detect Vegetation and Land Use Changes in a Subtropical Region of Mexico," *Digest—International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 1, IEEE, Piscataway, NJ, 1988, pp. 587-588.
28. S. L. Soria, F. G. Alonso, and J. M. C. Gozalo, "Change Detection in the Doñana National Park," *Proceedings of the 10th Annual International Geoscience and Remote Sensing Symposium—IGARSS '90*, IEEE, Piscataway, NJ, 1990, p. 935.
29. H. A. Zebker and R. M. Goldstein, "Topographic Mapping from Interferometric Synthetic Aperture Radar Observations," *J. Geophys. Res.* **91**, No. B5, 4993-4999 (1986).
30. R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite Radar Interferometry: Two-Dimensional Phase Unwrapping," *Radio Sci.* **23**, No. 4, 713-720 (1988).
31. F. K. Li and R. M. Goldstein, "Studies of Multibaseline Spaceborne Interferometric Synthetic Aperture Radars," *IEEE Trans. Geosci. Remote Sensing* **28**, No. 1, 88-97 (1990).
32. C. Prati, F. Rocca, A. M. Guarnieri, and E. Damonti, "Seismic Migration for SAR Focusing: Interferometrical Applications," *IEEE Trans. Geosci. Remote Sensing* **28**, No. 4, 627-639 (1990).
33. H. A. Zebker and J. Villasenor, "Decorrelation in Interferometric Radar Echoes," *IEEE Trans. Geosci. Remote Sensing* **30**, No. 5, 950-959 (1992).
34. H. A. Zebker, S. N. Madsen, J. Martin, K. B. Wheeler, T. Miller, Y. Lou, G. Alberti, S. Vetrilla, and A. Cucci, "The TOPSAR Interferometric Radar Topographic Mapping Instrument," *IEEE Trans. Geosci. Remote Sensing* **30**, No. 5, 933-940 (1992).
35. C. Covault, "European Satellite Advances Earthquake Forecast Effort," *Aviation Week & Space Technol.* **137**, No. 24/25, 50-52 (1992).
36. N. H. Ching, D. Rosenfeld, and M. Braun, "Two-Dimensional Phase Unwrapping Using a Minimum Spanning Tree Algorithm," *IEEE Trans. Image Processing* **1**, No. 3, 355-365 (1992).
37. D. C. Ghiglia and L. A. Romero, "Direct Phase Estimation from Phase Differences Using Fast Elliptic Partial Differential Equation Solvers," *Opt. Lett.* **14**, No. 20, 1107-1109 (1989).
38. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, New York, 1986.
39. M. D. Pritt and J. S. Shipman, "Least Squares Two-Dimensional Phase Unwrapping Using FFTs," to be published in *IEEE Trans. Geosci. Remote Sensing*.

Received February 19, 1993; accepted for publication August 6, 1993

Mark D. Pritt *IBM Federal Systems Company, 800 N. Frederick Avenue, Gaithersburg, Maryland 20879 (PRITTM at WMAVM7, prittm@wmavm7.vnet.ibm.com).* Dr. Pritt received B.A. and M.A. degrees in mathematics from The Johns Hopkins University in 1985, and M.S. and Ph.D. degrees in mathematics from Yale University in 1987 and 1989, respectively. His doctoral work consisted of the identification and study of a certain class of objects in the field of algebraic topology called minimal simplicial spaces. Since joining the IBM Federal Systems Company in 1989, he has carried out research and development work in image registration, three-dimensional computer graphics, and interactive computer graphics applied to satellite attitude problems. Dr. Pritt is currently an Advisory Engineer/Scientist.