Edge- and Shape-Based Geometric Registration

THOMAS C. HENDERSON, MEMBER, IEEE, ERNST E. TRIENDL, AND RUDOLF WINTER

Abstract—The standard method for geometric registration of images consists of selecting control points in the two images and computing the correlation maximum of small subimages containing the control points. This method does not work well when applied to images taken at different seasons or with different sensors. The use of edge-based registration has been proposed to overcome these difficulties but has so far achieved no better than picture raster element accuracy. This paper presents edge- and shape-guided correlation (or comparison) of control point areas for the analysis of multitemporal and multisource data. The direct correlation of control areas for registration is supplemented by comparison of descriptions of elementary objects, e.g., drawn lines, borders, and edges, whose positions are known with subpixel accuracy. These methods have been implemented as a set of image registration modules within the context of the DIBIAS image processing system.

Keywords—Image registration, edge models, shape models, Hough transform, remote sensing, multitemporal images, subpixel accuracy.

I. Introduction

Digital image processing techniques have been found quite useful in the analysis of remotely sensed imagery and, in particular, Landsat images. Although these techniques are valuable for on-board processing, radiometric corrections, image enhancement, etc., we will outline results only insofar as they are related to geometric registration.

Informally, the problem of geometric registration involves transforming an image so that a simple relation exists between the locations of the picture elements (called pixels) and the actual geographic locations of those pixels, e.g., only a change in scale. Usually the distortions in an image are corrected in two separate steps. First, the systematic errors of sensor and flight path are corrected; that is, all known errors of the imaging systems are accounted for. This can be regarded as a preprocessing step, and in some image registration problems can be omitted entirely.

The remaining distortion of the image is due to unknown changes in position of the sensor. The most commonly used method to model these errors is to use the known locations of control points located on the surface of the Earth (called ground control points). For example, twenty or so ground control points well spread throughout a Landsat image usually suffice to give the required transformation.

Most registration systems allow for both manual selection of control points and automatic matching of a template containing the control point. The matching operation is performed by one of several standard techniques including normalized cross correlation or sequential similarity detection. Some experimental systems also allow for correlating edge responses (see Svedlow [12]); however, the edge correlation is most often treated exactly as gray-level correlation, where the gray levels correspond to the (thresholded) edge response.

A markedly different approach to the analysis of remotely sensed imagery has been proposed by Tennenbaum *et al.* [13] and stems from his work in scene analysis. The basic idea is to replace the reference image with a symbolic reference map containing explicit ground coordinates and elevations for all monitored sites, as well as landmarks (e.g., roads, coastlines, etc.). The geometric correspondence between the map and the sensed image is established by calibrating an analytic camera model. The camera model makes it possible to predict precisely the image coordinates (in the original unrectified image). The aim of the system is to make possible the continuous monitoring of predetermined ground sites, e.g., factories, reservoirs, etc., and to use various kinds of maps and knowledge sources to guide extraction of relevant information from the sensed images.

We propose to use models of edges and shapes to aid in the registration process. Section II defines the registration problem. Although the techniques developed here are applied to remote sensing data, they are also applicable to scene analysis problems. A model of physical edges, their appearance and their detection is described in Section III. A review of edge-based image registration is given in Section IV. Sections V and VI demonstrate the use of edge descriptors and shape models, respectively, in image registration. Conclusions and discussion of the method are presented in Section VII.

II. The Registration Problem

Let $f(i, j)$ and $g(i, j)$ be two digital images, i.e., two integer-valued arrays. Let $C = \{(i, j, k, l)$: the geometric location (i, j) in f corresponds to the geometric location (k, l) in g with i, j, k, l in $R\}$. We say that *T geometrically registers f with g* if

1) $T: R \times R \rightarrow R \times R$, and
2) $T(k, l) = (i, j)$, for every (i, j, k, l) in C.

T is called *the registration function of f with g*. The set C is called the *control point set*, where every four-tuple defines one control point pair: (i, j)-(k, l). Choosing the individual points from their respective images is called the *control-point selection problem* and involves finding suitable areas for matching; assigning two such selected points (one from each image) to a four-tuple is called the *control-point correspondence problem*. The image, f' defined by

Manuscript received September 21, 1982; revised January 1983. This work was supported in part under a grant from Deutsche Forschungsgemeinschaft.

T. C. Henderson is with the Department of Computer Science, University of Utah, Salt Lake City, UT 84112.

E. Triendl and R. Winter are with the Deutsche Forschungs-und Versuchsantalt für Luft-und Raumfahrt, (DFLVR, Oberpfaffenhoffen), 8031 Wessling, Germany.

0196-2892/85/0500-0334$01.00 © 1985 IEEE

$$f'(i,j) = f(T(i,j))$$

is called *the registered image of f with g*. This method of computing f' is called indirect registration since T produces a value at each pixel of f' by going back to a pixel (or neighborhood) in f. When referring back to a pixel in f, if that coordinate location lies between integer-valued pixels in f then some method must be provided for choosing a gray-level value f', e.g., nearest neighbor, bilinear interpolation, or cubic convolution; when the pixel referred to in f lies outside the boundary of f, then the pixel in f' is assigned some constant background value. Direct registration involves computing for each pixel in f the corresponding location in f'. Direct registration, however, has the disadvantages that the transformed points may not be evenly distributed and do not necessarily lie at integer-valued coordinates.

The *registration problem* is then to: 1) determine C (control point correspondence), 2) determine T (equation determination), and 3) compute f' (resampling). This is geometric registration in its most general form and is called *image to image registration*. The image g, however, plays no role other than to facilitate control-point selection. When no image g is required, the control points are specified merely with respect to some secondary coordinate system; this is called *absolute image registration*. Thus geometric correction is an instance of this. Finally, a model of edges or shapes (e.g., coastlines) can be used to register an image (or images). This involves edge or shape extraction and matching. This type of registration process is called model-guided image registration. See Henderson *et al.* [8], [9] for more detailed reports.

In the next sections, we will examine the use of edge and shape models in solving the control point correspondence problem. It should be pointed out, however, that on any given machine, special care should be taken in choosing the appropriate regression analysis and resampling algorithms. Otherwise, inefficiency in these areas will outweigh any benefits made in the control point correspondence.

III. Edge Models

The simplest model-guided methods are those which attempt to characterize edges or lines in an image. The resulting edge features can be matched directly or used to describe higher-level shapes. Originally, the motivation for taking such an approach was that even though extra processing was required for edge analysis, there were fewer edge elements in an image and consequently less matching computation. The major advantage now, however, is that such methods allow registration of diverse types of images (e.g., maps and aerial photos) which is usually not possible with gray-level methods.

Many remote-sensing applications, such as the production of land-use maps, call for locating edges in remotely sensed imagery at a higher resolution than the scanning raster. (Tennenbaum *et al.* [13] give a good description of using map knowledge to locate boundaries to better than image resolution accuracy.) This edge extraction step requires a model of the relation between the edge which exists in reality and the appearance of the edge in the image. Thus an edge model must account not only for the appearance of an edge in an image, but must also give some relation between the appearance of the edge in an image and the (subpixel) location of the edge in an infinite-resolution image plane.

Once a physical edge model has been chosen, i.e., assumptions have been made which describe the appearance of an edge in an image, a representational model must be chosen for image edges. In particular, edge detection is seen as a local operation, and each pixel has an edge descriptor associated with it. The edge detector produces at each pixel:

1) The exact location of the ideal edge, which is done using the physical edge model and the location is in terms of an infinite resolution image plane;
2) the orientation of the edge; and
3) the edge quality, i.e., the likelihood of an edge with the given location and orientation.

Edges arise in nature when two homogeneous regions are juxtaposed. A region is homogeneous because of the uniformity of some feature across the region, e.g., intensity, color, texture etc. The usual assumption is that an ideal step edge exists, that is, each region has a constant feature value, the regions have different feature values and the boundary between the two regions is a straight line. Therefore, the quality of a physical edge can be characterized by the: 1) uniformity of each region, 2) difference in the feature value of the two regions, and 3) transition ramp between the two regions. Of course, ideal edges are rarely found in nature, and in special cases, it may be preferable to assume a more complicated edge model, e.g., a land–water interface might be well represented by low variance on one side of the edge, and high variance with a shifted mean on the other side of the edge.

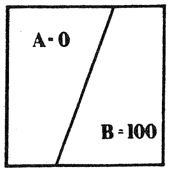
Given the physical edge model, it is possible to design algorithms to detect the appearance of edges in digital images. (See Davis [4] for a review of edge detectors.) Many of these edge detectors, however, are unsuitable for producing the required edge descriptor in that an exact subpixel edge location cannot be determined. The optimizing edge detector proposed by Triendl [14] is a local edge operator that does provide a complete edge description.

The optimizing edge operator is basically a Hueckel operator with the main differences that it: 1) works in the spatial domain, 2) uses a square aperture, and 3) incorporates the sensor point-spread function.

Sampling and digitization produce a transition ramp between the two uniform regions; that is, pixels have been produced which represent the sum of radiation intensities around the pixel center weighted by the point-spread function of the scanning device (e.g., for Landsat images this is a circular spot with a diameter of 75 m). The value of a pixel near an edge is a function of the distance of the edge from the pixel and can be computed from the point spread function.

An edge appearance model is used to produce a digitization of a given edge. Fig. 1 shows an ideal edge and one digitization of it. Correlation of the edge model with a subarray of the image gives a measure of the probability of an edge at angle a and distance r from the center of this array. Maximization of the correlation between the edge appearance model and the subarray gives the required position and edge quality.

The angle of the vector gradient is used as the initial value



0	0	70	100
0	20	90	100
0	40	100	100
0	70	100	100

Fig. 1. An ideal edge and a digitization of it.

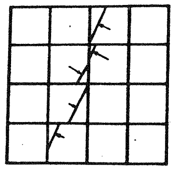


Fig. 2. Computed edge descriptions at each pixel.

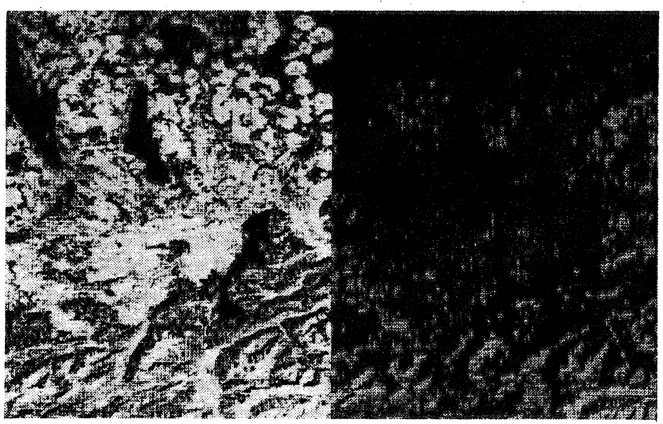


Fig. 3. Two-channel Landsat image used.

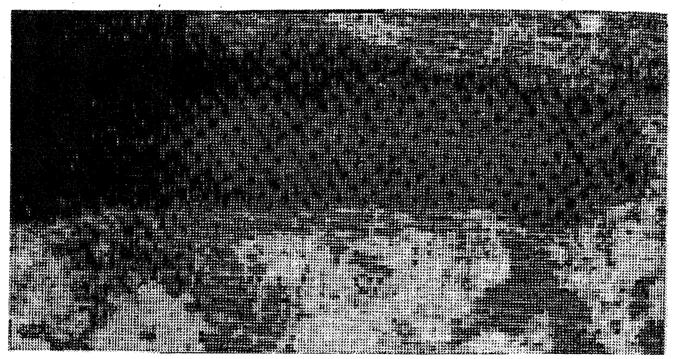


Fig. 4. Aerial image used.

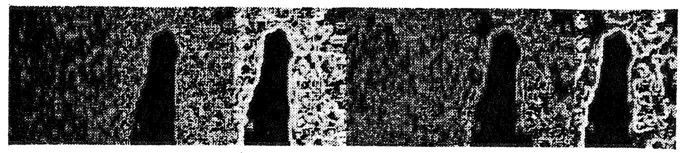


Fig. 5. Edge description image of Landsat image.

for the optimization of the edge model. In this way an edge triple (α, r, ρ) is produced at each pixel, where α is the edge orientation, r is the shortest distance from the center of the pixel to the edge, and π is the edge likelihood. Fig. 2 shows the edge response for Fig. 1.

Since the edge triples describe the location of the edges with subpixel accuracy, it is possible to reconstruct the edge image at a finer resolution and produce much smoother results (see Triendl [14]). The same idea can also be applied to produce texture edges once the texture areas have been characterized by some texture parameter (see Triendl and Henderson [15]).

This edge detection method has been implemented in the image processing system at DFVLR. Given a k-channel image as input, the edge detector produces a $3k$-channel output image, three channels per input channel. These three channels contain the (α, r, ρ) descriptor for each pixel in the channel. Figs. 3 and 4 show the Landsat image and the Aerial image, respectively, which will be used to illustrate the edge description method and the registration system. Fig. 5 shows the result of applying the edge operator to the Landsat image of Fig. 3. This figure shows that the edge quality channel (channels 3 and 6) provides a good visual description of the edges in an image. Landsat was taken March 17, 1973 (MSS 5, no. 1237-09392) at 47° 22′ north latitude and 10° 47′ east longitude. The aerial photo was taken February 28, 1978 from an elevation of 7372 ft.

Usually, however, one does not care to see the edge triple images, but rather an edge appearance image. Moreover, since the edge descriptors are at subpixel accuracy, the edge appearance image can show the edges at any resolution. Fig. 6 shows the results of displaying Landsat with a resolution of 2 and a resolution of 5.

This also allows for crossing edges at a single spatial location to be displayed since each pixel maps into a k by k window, and the pixels in this window can be set to display several edges.

Edge responses can be grouped across channels and spatially. First, the basic edge-grouping operation is performed across channels. Each spatial location in the image has (possibly) several edge descriptors associated with it. These edge descriptors are grouped in the following way: 1) Let (α, r, ρ) and (α', r', ρ') be the edge descriptors from two channels. 2) If the difference in angle between α and α' is small enough, and if $|r - r'|$ is small enough, and if ρ and ρ' are high enough, then (α, r, ρ) and (α', r', ρ') are grouped in such a way as to weight more heavily the more likely edge. Up to two distinct edge descriptors are produced for each spatial location; i.e., up to two edges from different channels are preserved. The output is always a six-channel image with the descriptor triple of the most likely edge in channels 1 to 3.

Once the edges have been channel grouped, they can then be grouped spatially. This step results in more compact and continuous edges. At each spatial location of the image, an n by n window of edge descriptors is grouped (with possibly two edge descriptors allowed at each spatial location) to produce an edge response at the pixel in question. The edge descriptor of each pixel in the window is first mapped into the coordinate system of the central pixel.

Once the transformation of coordinates has been affected, the grouping operation proceeds similarly to the channel grouping. One extra condition, however, for an edge to be produced at a pixel is that there exist some minimum number m of edges in the window which contribute to that edge. Fig. 7 shows the results of the channel and spatial grouping for the

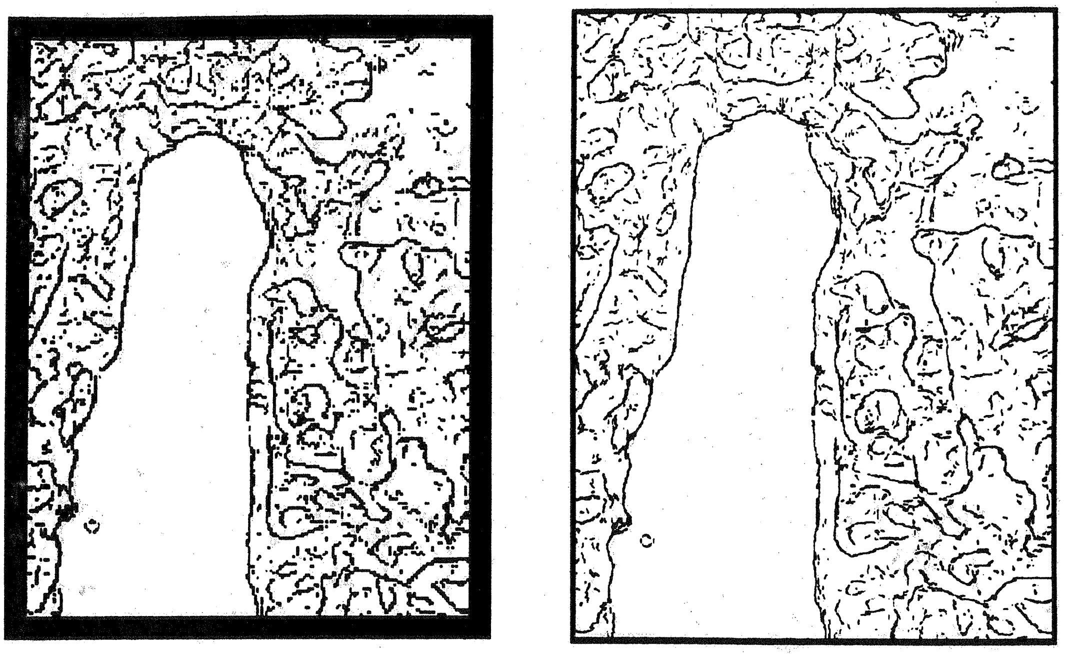


Fig. 6. Edge appearance of Landsat at resolution 2 and 5.

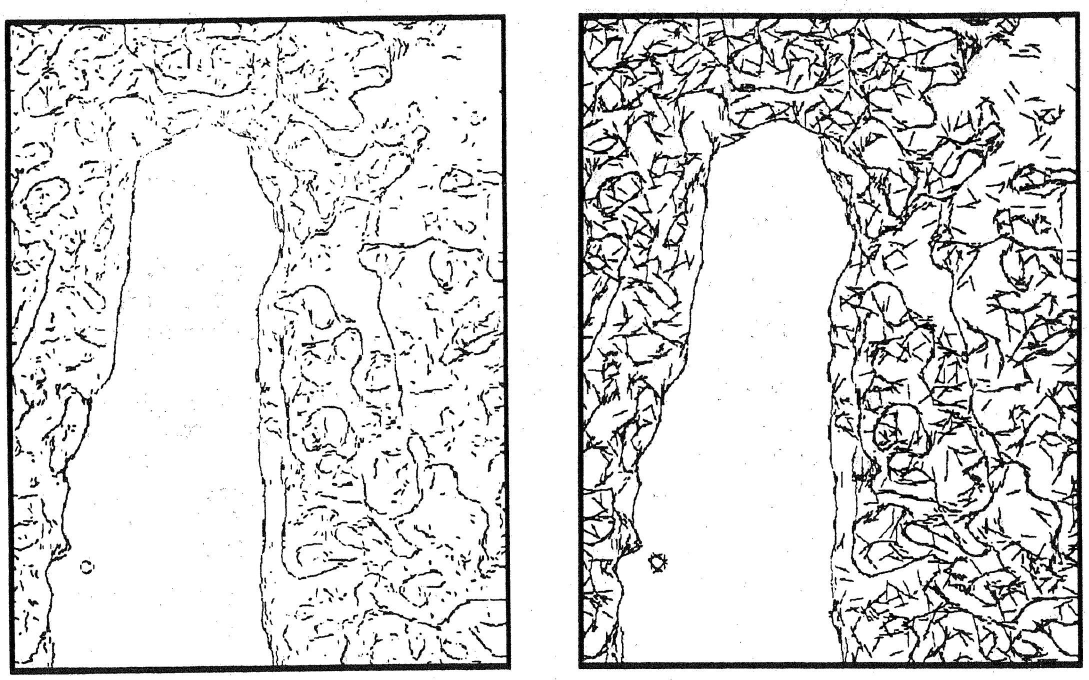


Fig. 7. Channel and spatial grouping of Landsat.

edges in Fig. 5. A group of modules have been developed for producing these edge descriptions. The usual order of application of the edge description modules is: 1) EDGING which produces the edge description image from a gray-level image; 2) EDGGRP which does channel grouping; 3) EDGSPG which performs spatial grouping; and finally, 4) EDGOUT which produces an edge appearance image.

IV. Review of Edge-Based Image Registration

Edge descriptors such as those produced by the edge operator described in the previous section can be used as the fundamental unit of description of an image-analysis system. Dudani and Luk [5] describe a scene-analysis system based on the description produced by the Hueckel edge detector. The usefulness of edge features in these systems is strongly related to the world model one works with, and the edge representations provide the link between the actual physical signal of the image and symbolic processing.

Image registration can be achieved by locating edges in images and then matching edges (see Andrus *et al.* [1], Price and Reddy [11], Wong [16], and Wong and Hall [17]). Andrus investigated the use of binary boundary maps for image registration. A binary edge response is produced (grouping all channel responses to one channel), and then correlation coef-

TABLE I
PERCENTAGE OF ACCEPTABLE REGISTRATION ATTEMPTS (FROM SVEDLOW)

Similarity Measure	Original Image	Magnitude of the Gradient	Thresholding Magnitude of Gradient
Correlation Coefficient	90%	100%	90%
Correlation Function	38%	74%	87%
Sum of Absolute Values of Differences	69%	92%	87%

ficients are computed for the binary response images. This allows much computational savings since the correlation coefficient can be computed using additions instead of multiplications. The motivation for their work was change detection, i.e., multitemporal analysis.

Another area in which edge-based registration has been shown useful is in the registration of images produced by different sensors. Wong provides an example of this approach for registering optical and radar imagery. In both types of imagery, certain types of edges appear consistently, e.g., roads, land-use interfaces and man-made structures. The intensities (or gray-level values) of the images, however, differ radically, whereas the edges remain recognizable even after some degree of degradation in resolution, intensity, and geometry.

Wong preprocesses the optical and radar images so as to achieve as close as possible the same size scale in the images and converts the intensities of the radar image to match as closely as possible those of the optical image. Edges are produced in both images, where the main requirements outlined for producing edges are given as: 1) keep salient edges of objects to be matched, 2) eliminate spurious edges, and 3) tolerate minor geometric misregistration. Using the edge detector of Frei and Chen [6], Wong has produced algorithms to eliminate background edges, threshold weak edges, and thicken remaining edges. Using these procedures, he shows that edges are useful for image registration. With the appearance of several similarity measures, e.g., correlation coefficient, sequential similarity detection algorithm, etc., and the use of edge features for image registration, Svedlow *et al.* [12] made a comparison of these methods. They chose three similar measures: 1) the correlation coefficient, 2) the correlation function, and 3) the sum of the absolute value of the differences. In addition to the original gray-level images, three preprocessing edge operations were investigated: 1) a gradient operator, 2) a threshold operator, and 3) a combination of these.

Tests of registration accuracy were done on Landsat data of Missouri and Kansas using separate spectral bands and windows of 51 by 51 pixels. Evaluation of registration results were based on data from previous registration of the images using a sophisticated registration process (e.g., ground truth was determined) and visual inspection. Each registration attempt was classified as either "successful" or "unsuccessful" based on whether or not the result was within a few pixels of the "correct" result. Table I gives the results of their work.

Thus the best performance was achieved by the correlation coefficient using the magnitude of the gradient of the images. Their conclusion is that preprocessing the images via a gradient operator enhances the ability to find an acceptable registration position. It should be pointed out that the thresholded gradient responses from different channels were grouped by an "or" operation.

The major flaw in all these edge-based registration approaches is that only the appearance of the edge or the appearance of the likelihood of the edge is compared between images. A more complete approach is to actually compare the edges themselves at a symbolic level. This problem has been studied in a more general setting by Price and Reddy [11]. In particular, they are interested in change detection in a general-purpose analysis system. Instead of comparing intensity values in a signal-based approach, they propose symbolic registration through the use of segments. A segment is a symbolic representation of some object in the image (e.g., edges, shapes, etc.); these segments are characterized by shape, position, and arrangement in the image. Since these segments are useful anyway for further analysis (e.g., the detection of changes in an image), they may also be useful for registration purposes. The usual registration assumptions are: 1) only translational misalignment, and 2) major portions of the image unchanged. Their symbolic approach relaxes these assumptions; however, the image is assumed segmented by some automatic method. Moreover, registration is understood to be at the level of objects; that is, as scene analysis and not so much as an accurate geometric registration of the images. These methods are less susceptible to mistakes for rotation differences and relative position changes. Finally, the similarity measure they use is defined as a sum of weighted feature differences.

V. EDGE DESCRIPTORS FOR IMAGE REGISTRATION

Edge descriptors, i.e., (α, r, ρ) triples, which describe the orientation, displacement, and edge quality with respect to a pixel can be used in a variety of ways to register images. The most common method is to produce a gradient image for each image to be registered and then cross correlate these images as gray-level images. A more powerful approach is to exploit the complete edge description in computing the correlation of the edge images. Finally, edge descriptors can be used as the basic elements of a shape-analysis system. Once a shape has been located in an image, the known information about the shape can be used to determine the geographic coordinates of the image. All of these methods are described here and are available with the image registration system of DFVLR.

Control-point selection is achieved through the use of the Comtal visual monitor display. Three control point areas were chosen in Aerial. The aerial photo is considered the template image, and the three control point areas are considered as templates to be found in the Landsat image. The three control point areas of Aerial are shown in Fig. 8. Next, the approximately corresponding points are chosen in the Landsat images (when visually chosen this way, misalignment is of the order of 10 pixels).

When the control points in the Landsat scene are chosen only approximately, then the system can be used to find the best correlation response location for the template image

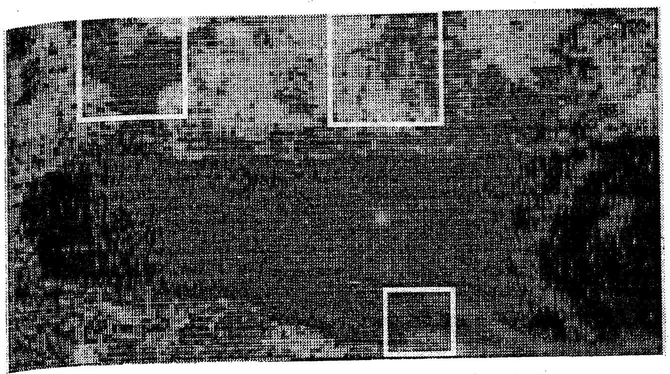


Fig. 8. Three control point areas in Aerial.

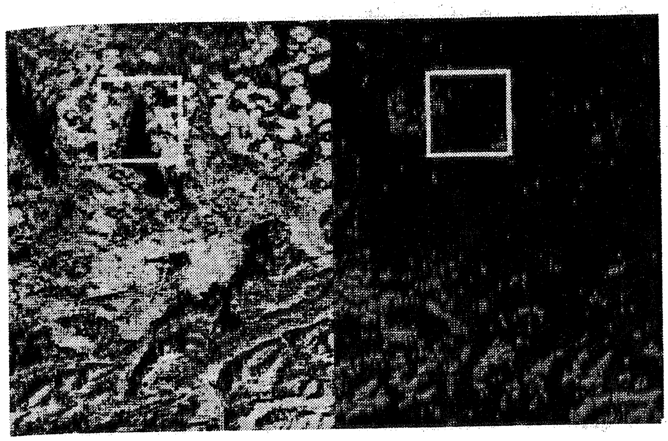


Fig. 9. Landsat search area.

control-point areas within a search area centered at the control points chosen for the Landsat images. Fig. 9 shows the Landsat search area.

VI. Shape-Based Image Registration

When the images to be registered do not provide sufficient information to use conventional methods, e.g., the scale of the images differs (or is unknown) or the image is too inaccurate, for example, hand drawn, then shape information can be extracted and used to make the correspondence between images. Many methods exist for defining and detecting shapes (see Pavlidis [10] for a review).

The position-invariant Hough transform has been incorporated for use since it is fast and easy to program. A generalization of the Hough transform technique (see Ballard [2] and Davis [3]) is used to recognize shapes in satellite images. The transform technique is equivalent to certain conventional template matching procedures, but is, on the average, 10 to 20 times faster. The technique was originally used for object tracking, but has been installed as a shape detector.

The following description of the position-invariant Hough shape transform is after Davis. Let $S = (X_i, Y_i)$, $i = 1$ to n, be the locations of edge elements comprising a shape. Let $R = (X_0, Y_0)$ be any point; note, however, that a central point such as the centroid of S will be computationally more efficient. The Hough representation of S using R, called $H(S, R)$, is the list of vectors (dX_i, dY_i), $i = 1$ to n, where $dX_i = X_0 - X_i$ and $dY_i = Y_0 - Y_i$.

TABLE II
INTERPOLATION RESULTS FOR GRADIENT RESPONSE

Template Area	Distorted Image	Reference Image	Maximum Correlation Location
1	LANDSAT	AERIAL	(266.15,314.01)
2	LANDSAT	AERIAL	(210.91,368.64)
3	LANDSAT	AERIAL	(263.64,366.71)

TABLE III
INTERPOLATION RESULTS FOR EDGE RESPONSE

Template Area	Distorted Image	Reference Image	Maximum Correlation Location
1	LANDSAT	AERIAL	(265.91,313.88)
2	LANDSAT	AERIAL	(210.32,368.20)
3	LANDSAT	AERIAL	(263.67,366.86)

Given an image f, which contains an instance of the shape S, an array h is used to compute the transform of f with respect to $H(S, R)$. Points in h with high values correspond to likely locations for R in f.

The transform is applied to $E(f)$, the edge image of f. Each edge element, $e(i, j)$, in $E(f)$ is potentially in S. Although the edge descriptor (α, r, ρ) can be used to some extent to limit the subset of S to which $e(i, j)$ could correspond, there is no way to be sure to which element of S, if any, $e(i, j)$ corresponds. Since this is the case, every edge element $e(i, j)$ is compared to each vector in $H(S, R)$ to compute a possible location for R, and that location is incremented in h. The following algorithm is used to compute h:

for every $e(l, m)$ at (X_i, Y_i) in $E(f)$ do
  for every (dX_j, dY_j) in $H(S, R)$ do

$$h(X_i + dX_j, Y_i + dY_j) := h(X_i + dX_j, Y_i + dY_j) + 1.$$

This algorithm is computationally simple, but requires a large array h. A good feature is that no explicit object segmentation must be performed since the edge descriptors form the elements of the model.

The results of the normalized cross correlation followed by Lagrangian interpolation are given in Table II. Using the gray-level correlation results as reference values, the Euclidean distance of the gradient results from the gray-level results can be determined as shown in Table III.

Edge descriptors can also be used to correlate images. In this case, however, the gray-level correlation cannot be applied directly, and an alternative method has been developed. Let $w(i, j)$ be the window of the search area to be correlated with $t(i, j)$ the template of size m by n. The correlation coefficient at a given spatial location is given by:

1) The coefficient is $P1 \times A + P2 \times B + P3 \times C$ (where the P's weigh the contribution of each part of the edge description).
2) The correlation coefficient of the whole window is given as the average value of all the coefficients in the window.

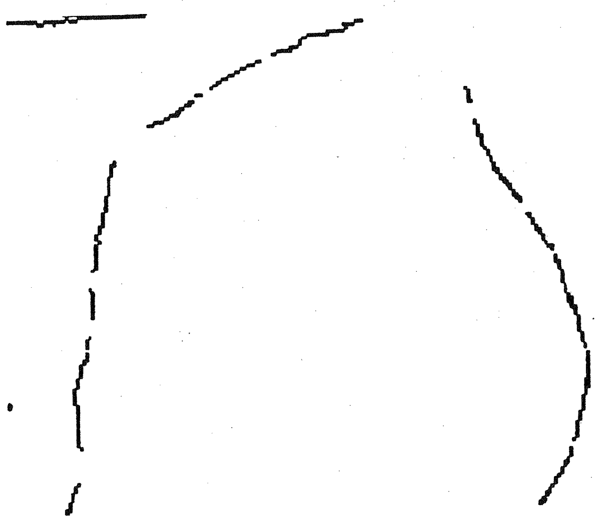


Fig. 10. Edges in shape and model of Starnberger See.

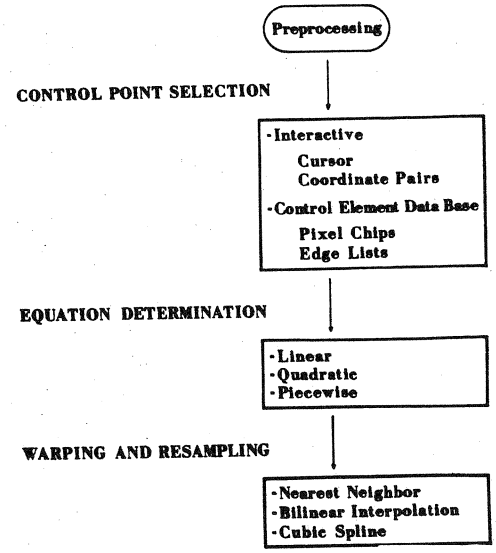


Fig. 11. Image registration system.

TABLE IV
SHAPE MATCHING RESULTS

Image	Shape	Location	Error (in pixels)
LANDSAT	STARN	(31,40)	3.6
AERIAL	STARN	(16,34)	0.0

Table III gives the results of using this measure of similarity and weighing values of $P1 = 0.75$, $P2 = 0.20$, and $P3 = 0.05$. The results match well with the gradient response with a difference of 0.27, 0.81, and 0.15 at the 3 control points. Thus the edge response correlation provides as accurate a registration as the gradient method, but is applicable to a wider class of images, and moreover, provides the basis for a shape–guided approach.

The shape model described in the preceding section has been implemented using the edge triples from the edge operator as shape elements.

Given an edge-triple input image, a reference point and an edge quality threshold, a control point file which contains the displacement vectors of the shape is produced. Each element of the file corresponds to the location of an edge descriptor which was above threshold edge quality. Thus the best edges can be used to define a shape. Fig. 10 shows the edge elements used to define the north end of the Starnberger Lake (near Munich, Germany). This shape was defined using Aerial as the input image.

The system computes the Hough shape transform for a given image and shape. Table IV shows the results of the Hough shape transform to locate the Starnberger Lake in Aerial and Landsat.

VII. CONCLUSIONS

A comprehensive model-guided geometric registration and correction system has been developed and is currently running for remote sensing applications, e.g., see Gossmann and Haberaecker [7]. Control point selection involves choosing pairs of corresponding points in the two images to be registered. This can be done interactively or by means of a control point data base which contains templates (e.g., pixel arrays or shape models) which can be correlated with the image. If the gray levels of two images are too radically different to be correlated, then the edge response arrays of the windows can be correlated. Finally if even the edge responses are too radically different, then special shapes, e.g., coastlines, can be extracted and matched.

The optimizing edge detector has been analyzed for suitability within a registration system as a basis for edge modeling. The edge triples (α, r, ρ) produced by this method, however, provide the primitive elements for a shape model. An algorithm has been developed to compare the edge triples. Such a comparison is necessary if subpixel accuracy in image registration is to be achieved. Subpixel accuracy is the result of the edge model.

The edge detector has been implemented in such a way as to provide three spectral output channels for every input channel. These output channels correspond to the edge triple; thus ordinary edge registration can be performed by cross correlating the likelihood channels of two images. The most accuracy is, however, obtained by using all three output channels. The normalized cross correlation is available for both gray-level arrays (i.e., one channel) and edged-descriptor arrays (i.e., three channels).

A first step has been made toward a general shape-matching capability in that the Hough shape transform model has been incorporated for the image registration system to define and detect shapes composed of edges elements as provided by the edge operator. More work needs to be done to assess the suitability of this shape analysis method for registration purposes, especially since the Hough transfer for Landsat images requires a large amount of memory; however, this problem can be overcome by the use of a *kd*-tree accumulator. Other techniques need to be explored.

The registration system described earlier is embedded within the digitale interactive bild auswertung system (DIBIAS) at DFVLR. In accordance with the goals outlined earlier, a comprehensive model-guided geometric registration and correction system has been developed for remote-sensing applications (see Fig. 11). Our work has centered on the control point selection

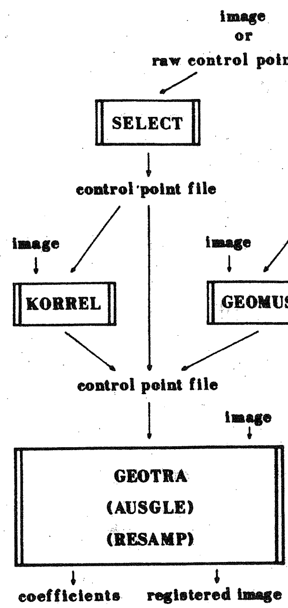


Fig. 12. Control-point selection alternatives.

Fig. 13. Registered Landsat image.

problem; user alternatives provided by the system are shown in Fig. 12. The user can interactively designate control point pairs (module Select) or prepare a database of shapes (using the module Makemu). The geometric correction is performed by the module Goetra which simply takes the control point pairs to produce the registered image. The control point file can come directly from Select, or after shape detection by the module Geomus. The registered image for our running example is shown in Fig. 13.

Future work includes the automation of the correlation of data-base elements (gray-level windows, edge descriptor windows or shapes) with images based on approximate knowledge of the location of the sensor. Also, the relation of sensor type (reflectance, infrared, thermal, etc.) to the control point selection problem needs investigation.

References

[1] J. Andrus, C. Campbell, and R. Jayroe, "Digital image registration method using boundary maps," *IEEE Trans. Comput.*, vol. C-19, pp. 935–940, Sept. 1975.

[2] D. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," Tech. Rep. TR-55, University of Rochester, Comp. Sci., Oct. 1977.

[3] L. S. Davis and S. Yam, "A generalized Hough-like transformation for shape recognition," Tech. Rep TR-134, University of Texas, Feb. 1980.

[4] L. S. Davis, "A survey of edge detection techniques," *Comput. Graph. Image Process.*, vol. 4, pp. 248–270, 1975.

[5] S. Dudani and A. Luk, "Locating straight-line edge segments on outdoor scenes," *Pattern Recognition*, vol. 10, pp. 145–157, 1978.

[6] W. Frei and C. Chen, "Fast boundary detection: A generalization and a new algorithm," *IEEE Trans. Comput.*, vol. C-26, pp. 988–998, Oct. 1977.

[7] H. Gossmann and P. Haberaecker, "Image processing of HCMM-satellite thermal images for superposition with other satellite imagery and topographic and thematic maps," *Tellus–Newsletter*, vol. 12, Jan. 1980.

[8] T. Henderson, E. Triendl, and R. Winter, "Model-guided geometric registration," Tech. Rep. NE-NT-D-50-80, *Deutsche Forschungs-und Versuchsanstalf fuer Luft-und Raumfahrt*, Sept. 1980.

[9] ——, "Edge-based image registration," in *Proc. 2nd Scandinavian Conf. Image Analysis*, pp. 106–111, June 1981.

[10] T. Pavlidis, "A review of algorithms for shape analysis," Tech. Rep. 218, Princeton University, Sept. 1976.

[11] K. Price and R. Reddy, "Matching segments of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, no. 1, pp. 110–116, 1979.

[12] M. Svedlow, C. McGillem, and P. Anuta, "Image registration: Similarity measure and preprocessing method comparisons," *IEEE Trans. Aerospace Electron Syst.*, vol. AES-14, no. 1, pp. 141–149, 1978.

[13] J. Tennenbaum, M. Fischler, and H. Wolf, "A scene-analysis approach to remote sensing," Tech. Rep. 173, SRI, Oct. 1978.

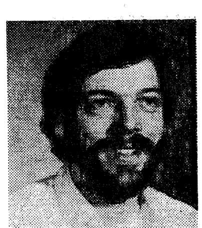
[14] E. Triendl, "How to get the edge into the map," in *Proc. Internat. Joint Conf. Pattern Recognit.*, pp. 946–950, 1978.

[15] E. Triendl and T. Henderson. "A model for texture edges," in *Proc. Internat. Conf. Pattern Recognition*, Dec. 1980.

[16] R. Wong, "Sequential scene matching using edges," *IEEE Trans. Aerospace Engineer. Sci.*, vol. AES-14, no. 1, pp. 128–140, Jan. 1978.

[17] R. Wong and E. Hall. "Performance comparison of scene matching techniques," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. PAMI-1, no. 3, pp. 325–330, July 1979.

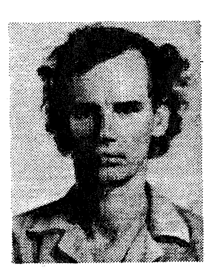
*



Thomas C. Henderson (M'82) received the Ph.D. degree in computer science from the University of Texas, Austin, in 1979.

Starting in 1979, he spent one year at the German Aerospace Research Establishment (DFVLR), Wessling, Germany, and the following year at the National Institute of Research in Computer Science and Control Theory (INRIA), Roquencourt, France. Since 1982, he has been an Assistant Professor with the Computer Science Department, University of Utah, Salt Lake City. He has published papers in the areas of physiological psychology, image processing, pattern recognition, and remote sensing, and is currently interested in computer vision and robotics.

*



Ernst E. Triendl received the Ph.D. degree in physics from the University of Innsbruck, Austria, in 1966.

Since 1967, he has been with the German Aerospace Research Establishment (DFVLR), Wessling, Germany. Under his supervision, the Digital Interactive Bavarian Image Analysis System (DIBIAS) has been developed. He is the author of numerous papers in the fields of image processing, remote sensing, and pattern recognition. His current interest is vision systems.

Dr. Triendl is a member of the German Workgroup for Pattern Recognition and the German Computer Society.

*

Rudolf Winter received the Ph.D. degree in physics from the University of Gottingen in 1973.

After a post-doctoral fellowship at the Scripps Institute of Oceanography, La Jolla, CA, he joined the German Aerospace Research Establishment, Wessling, Germany. He is now heading the Remote Sensing Section there. He has been an active researcher in geomagnetic signal processing, digital image processing, and remote sensing. He has written several articles on magnetotellurics, spline interpolation for digital images, and image processing applied to satellite imagery.

Dr. Winter is a member of the German Geophysical Society and the Remote Sensing Society.