7.4 INTER-IMAGE MATCHING

A Tutorial Presented to the NASA Working Group on Image Registration and Rectification

Robert H. Wolfe, Jr.
IBM Federal Systems Division
Houston, Texas

Richard D. Juday
NASA Johnson Space Center
Houston, Texas

November 18, 1981

FOREWORD

This paper was invited to provide a review of the portion of the registration process relating to determining relative positions of reference and registrant images at a set of locations within the images and embodiment of those positions in a reference-to-registrant mapping function. There was an unfortunately short time available between request and conference date, and as a result the review is not as thorough as the authors would have liked to have given. We are most familiar with the processing of earth observations from Landsat, and have concentrated our review on the processing of Landsat data by the Master Data Processor (MDP) at the Goddard Space Flight Center (GSFC), the registration processor installed at the Johnson Space Center (JSC), the registration process used in the DAM (detection and mapping) Package in conjunction with a Corps of Engineers project to map surface water, the registration process used by the LACIE processor at GSFC, and the sequential similarity detection algorithm (SSDA) used by IBM/Houston in the evaluation of the LACIE processor imagery. Our review slights the body of information available from military research, in which shape recognition and artificial intelligence play a prominent part. In that field, there is interest in finding objects that have changed position with respect to a fixed background, or monitoring changes in aspect for guidance, and there is often limited interest in spectral development. Our own field of concentration involves images in which considerable spectral development has taken place (observations are made throughout a crop year for agricultural applications), although for the most part the scene's basic geometrical shape is unchanged. To some extent field boundaries change from year to year, but this is usually a small enough effect so as not to interfere with the registration process. In the presence of spectral development, one does not ordinarily correlate between images on the radiance measurement making up the parent image, but instead relies on a derivative feature (in our case, edges) that one hopes will be stable despite spectral development. We defer the discussion of extracted feature selection to other papers being presented at this workshop, noting in passing that we are most familiar with the edge-detection algorithm implemented in both the MDP and the JSC registration processor. This paper further concentrates on automated interimage correlation; the manual techniques involved in the DAM Package and other systems are a fundamentally different technology. First, of course, the points of correspondence

between images are manually determined. Second, the interpretation is normally done on radiance images transformed into visual presentations. And finally, the manual techniques are typically done on a point basis with only contextual involvement of the neighborhood around the point designated as being in correspondence. In the image comparisons discussed in this paper, measures of similarity are made over areal extents of subsets of the parent images, as opposed to points.

Introduction

Interimage matching is the process of determining the geometric transformation required to conform spatially one image to another. In principle, the parameters of that transformation are varied until some measure of "difference" between the two images is minimized or some measure of "sameness" (e.g., cross-correlation) is maximized. The number of such parameters to vary is fairly large (six for merely an affine transformation), and it is customary to attempt an a priori transformation reducing the complexity of the residual transformation or subdivide the image into small enough match zones (control points or patches) that a simple transformation (e.g., pure translation) is applicable, yet large enough to facilitate matching. In the latter case, a more complex mapping function is fit to the results (e.g., translation offsets) in all the patches. The methods reviewed have all chosen one or both of the above options, ranging from an a priori along-line correction for line-dependent effects (the "high-frequency correction") to a full sensor-to-geobase transformation with subsequent subdivision into a grid of match points.

There is, of course, a correct geometric transformation to apply, but it is unknown (otherwise, an empirical image matching process would be unnecessary); thus, some sort of model must be assumed whose parameters can be solved for by correlation of offset-fitting. Commonly, a portion of the geometric model is established a priori based on external data such as preflight measurements. If that portion is incorporated in an a priori transformation, then the demand for fidelity in the overall model becomes an issue of tradeoffs between the a priori geometric correction and the residual correction to be determined by matching. For example, if one know the geometric transformation is affine with an unknown translation, one might let the image matching function solve for the whole affine transformation, risking the introduction of errors into the affine coefficients, or apply the affine transformation a priori and solve only for translation. Thus, in addressing the various parts of image matching in the following sections, we must also consider their interaction with portions of the overall geometric correction process which otherwise might be regarded as beyond the scope of this paper. The utility of an a priori transformation is also manifested in another way. Data gathered from prior registrations can be used to bias the a priori correction. Such "experience" data might be gathered by

analyzing ensembles of prior registrations such as associated acquisitions (different dates, same scene) or associated geometry (e.g., different subscenes, same frame). Such experience data are then included in the data base.

The following sections logically divide, according to the above precepts, into (1) a consideration of correlation techniques, (2) examination of other matching methods, (3) a discussion of determining the translational offset from the correlation, (4) a consideration of the residual geometric model and how to determine it, and (5) a summary of techniques going beyond the assumption implicit in the previous sections that match points have been established to allow correlation for translational offsets only. Throughout these discussions we make reference to the registration processor recently installed at JSC, note experimental results pertaining to that system as appropriate, and note general limitations and areas deserving further study.

Image Correlation

Suppose that an a priori correction has been applied or patches are defined small enough that any residual geometric error worthy of consideration is pure translation. Matching, then, consists of determining the translational offset of one subimage from another "reference" subimage corresponding to the same scene. The most common form of matching is some sort of cross-correlation technique. The "image" referred to is the actual radiance image or a feature-space image such as an edge or gradient image. In principle the cross-correlation of two acquisitions of the same scene should resemble the autocorrelation displaced by the same amount as one image from the other. Since the peak of the autocorrelation function must occur at the origin, the peak of the idealized cross-correlation then measures the displacement. In practice, the acquisitions differ because of instrument, atmosphere, and other environmental noise and because the scene may have changed appearance due to seasonal, weather, or cultural changes. Thus, the cross-correlation only approximates the autocorrelation, and the peak may not be well-defined. Another argument for the cross-correlation peak measure is that it minimizes the sum-square-difference between the two images. This measure of cross-correlation can be normalized (to be between -1 and 1) by two techniques summarized in Figure 1. The "template matching"[1] alternative utilizes only the pixel values from the search area, whereas the "classical" alternative

utilizes both images. The pixel values indicated are after subtraction of the image means. For binary edge images, whose pixel values are either 0 or 1, rigor fails in the following sense. Strictly speaking, the means of the 0's and 1's should be subtracted before summing; however, considerable gain in computation speed is achieved if the coincidences and numbers of 1's are summed directly.

The correlation techniques just described are employed in several production registration systems. The Master Data Processor (MDP)[2,3], the GSFC baseline system for Landsats 2 and 3, basically performs the "classical" cross-correlation on the radiance image. As a functional equivalent, Fourier techniques are used; that is, the reference and search areas are transformed by the standard FFT algorithm, one of the transforms is multiplied by the complex conjugate of the other, and the result is inverse-transformed by FFT. A slight variance is effected in the denominator by subtracting the local mean (mean over portion of search area overlayed by reference) from the search area's pixels, rather than the mean of the whole search area. The latter is used in the numerator in conformance with the classical formula. The GSFC LACIE Registration Processor[4] employed the template matching algorithm on binary edge images, and the same scheme was used in an evaluation of that processor.[5] Both classical and template matching algorithms are provided for binary edge correlation in the JSC Registration Processor.[6,7] The classical option has been chosen for production processing.

The matching policy discussed above is based on minimizing the sum-square-difference. The sequential similarity detection algorithm (SSDA)[8] is based on minimizing the sum-absolute-difference. However, rather than summing over all pixels, the SSDA selects at random a subset of pixels to sum. Summing stops when a present threshold is reached, and the number of pixels needed for that sum is noted. Then, the correlation maximum occurs at the same point as the maximum of the numbers of pixels required (because the sum-absolute-difference is smallest there requiring the most pixels to reach the threshold). The utility of the SSDA lies in its reduction in computation by utilizing only partial sums whose computational rigor is less, the smaller the correlation is. Thus, the bulk of the computation is devoted to correlation samples showing promise with little effort wasted on low values. The SSDA was also used[9] this time on the

radiance images, in the performance evaluation of the GSFC LACIE Registration Processor. The correlation was normalized by dividing the pixels of each image by their standard deviation before correlating.

The design of the SSDA illustrates a primary concern in standard cross-correlation, viz diminishing the number of required computations. In general, methods to address that issue have focused on devoting the bulk of the computations to the region around the appropriate extremum. Such a method, well-suited to binary edge correlation, was used in the GSFC LACIE Registration Processor[4] and for coarse acquisition in the JSC Registration Processor. First, cross-correlation is done only for offsets every fourth row and every eighth column. The mean and standard deviation of the results are computed, and thrice the latter is added to the former to generate a rejection threshold. Next, an approximate correlation is computed at every offset but with only pixel values from every third row and third column included in the sums in Figure 1 (thereby reducing the computation by about a factor of nine). If, at a given displacement, the resulting approximate correlation value exceeds the rejection threshold, the sums are repeated using every pixel value. In binary edge correlation the peak is generally fairly sharp, and only a very small fraction of the correlation samples exceed the threshold, thereby sparing a considerable amount of computation. This method, used for a segment-level correlation (8 x 10 km portion of a Landsat MSS frame), proved very effective in the GSFC and JSC Processors.

The foregoing discussion assumes the geometric difference between the compared patches is pure translation. This assumption may not be valid, which raises the question of how geometric distortion (other than translation) affects the cross-correlation function. The effect has been studied for standard cross-correlation (i.e., for minimizing the sum square difference) where one patch is distorted linearly (rotation, scale change, shear distortion) from the other.[10,11,12] The results indicate that linear distortion effectively blurs the cross-correlation function by applying a running average over the ideal nondistorted counterpart, where the dimensions of the averaging filter are proportional to the degree of distortion. This conceptual averaging does not apply to the noise present, so that the signal-to-noise ration (SNR) is effectively reduced. Also, the area around the peak is effectively flattened somewhat, making the peak search less accurate, and the peak-to-background (or peak-to-sidelobe) ratio is distorted

(these subjects are discussed in a subsequent section). The magnitude of these effects is greater, the greater the patch size, because the geometric distortion becomes larger. Thus, choosing a smaller patch decreases the effects of geometric distortion, but the SNR is also decreased with fewer pixels. Some compromise is generally required, as is discussed in a later section.

The effects of geometric distortion on binary edge correlation could be drastic in certain cases. For example, a scale difference might prevent overlaying edge features on opposite sides of a patch at the same time. Thus, rather than smearing the correlation function, as described above, it could divide into several peaks. This kind of problem has not been identified in typical patch correlation, but it has been suspected of degrading the full sample-segment correlation used in the GSFC LACIE and JSC Processors. One solution to the problem (at least for coarse correlation) might be to resample the images to a coarser grid by use of a low-pass or median filter before edge detection. Then, the geometric distortions are reduced relative to the pixel spacing or, equivalently, effective edge "thickness."

Other Techniques for Image Offset Matching

Several other matching techniques, though not yet implemented in large-scale registration, deserve mention as potentially applicable. The Cluster Reward Algorithm (CRA)[13] conceptually analyzes the bivariate histogram of the two images at each displacement. A measure is established characterizing the definition of pattern in the histogram, e.g., a measure of clustering. At the offset denoting an image match the histogram should show a relatively high image-to-image correlation by exhibiting clusters. The technique has been applied to several sample-segment-size (about 8 x 10 km) scenes with notable success.

The point-matching technique[14,15], though designed primarily for target arrays such as star fields, might find application to binary edge images. The method minimizes a geometric distance measure between points (say, edge pixels in the two images) rather than minimizing radiometric or feature differences. A variation[15] of the method ascribes weights to the point-pairs according to how well they associate in the matching. This technique is iterative with first

a matching, then a comparison of the distance spanned by each point-pair to the estimated common displacement for assigning a weight, then a weighted match, etc.

The use of normalized cross-correlation for minimization of sum-square-differences or sum-absolute-differences described in the previous section can also be viewed as an application of maximum-likelihood classification in which the measurement sets are distinguished by translational offset (analogous to pixel number in spectral classification), the channels of information are the pixels in the search overlay area (analogous to the bands in spectral classification), and the classes are Match and Non-match (analogous to classes, e.g., crop species, in spectral classification). In the spatial case, however, it is known a priori that only one measurement set should be classified as a Match (if the match scene is unique) corresponding to the single correct translational offset. Analogous to spectral classification a Bayes technique has been applied[16] to spatial matching. The probability of misacquisition (match lock-on to the wrong point), expressed as the Bayes risk is minimized by maximizing the a posteriori probability of a correct match (i.e., the probability that there is a match at an offset, given the particular pixel values in the overlay region for that offset). The a posteriori probability is estimated in principle by least squares by assuming it is a linear function of the pixel values. In fact, that probability should exceed 1/2 only at the match offset which provides a direct means for identifying a probable misacquisition (a posteriori probability never exceeds 1/2). The method is conveniently implemented by Fourier transformation techniques. The a posteriori probability, though assumed linear in the pixel values above, can be expressed as a linear combination of any functions of the measurements, as deemed appropriate to the application. The case of a second-order relation has been investigated[17], and trials on synthesized data indicate improvement over the maximum likelihood alternative offered by cross-correlation.

The methods described in this section should be considered as new candidates for interimage matching. They must be integrated with the various feature selection methods (e.g., the Cluster Reward Algorithm makes no sense with binary edge images, while the point pattern matching technique is especially suited to such images). The methods should be tried on a representative set of sensor data and evaluated for applicability to different sensor types, scene types, season, sensor-to-sensor differences, etc.

Offset Determination

The cross-correlation techniques for image matching presumes that the position of the maximum or minimum of the correlations or difference function characterizes the translational offset between the two images. The true offset might be in between pixel centers so that the correlation peak lies between correlation samples. To achieve subpixel accuracy in a given patch, some form of interpolation is needed. Several alternatives have been developed, as we discuss in this section. These techniques lend themselves to estimating the accuracy of peak location and to evaluating certain measures for establishing pass/fail status, and we address this subject also.

As an alternative to peak interpolation for subpixel accuracy, a number of patches can be defined, the peak can be determined to the nearest sample, and the subsequent mapping function fit to the numerous peak offsets can be relied upon to yield subpixel accuracy. This approach is predicated on the assumption that a random position error is introduced by choosing a sample position rather than interpolating. This method was used in matching images produced by the GSFC LACIE Processor for the purpose of evaluating the performance of that processor[5,9]. Secondary peaks were also identified and compared to the primary peak. A measure of decline away from the peak was established in terms of averages of correlation samples in concentric rings extending out from the peak. If the peak was not strong enough relative to the secondaries or the decline was too shallow, that zone was rejected in the mapping function fit.

A straightforward means of interpolating for the peak consists of fitting a surface function to the cross-correlation samples in the vicinity of the peak and evaluating the peak location analytically. The MDP and JSC Registration Processor[6,7] use a bivariate polynomial (i.e., containing terms $x^M y^N$, $MN \leq$ polynomial order). The MDP uses a fourth-order polynomial on a 5 x 5 neighborhood around the peak and also evaluates the curvature at the peak as a measure of the breadth of the correlation peak. The smaller the curvature is, the larger is the uncertainty in locating the peak. The MDP uses the minimum curvature on the surface and the height of the peak as rejection criteria.

The JSC Registration Processor allows any order polynomial up to and including fourth order, but the latter has been used mainly, with the lower orders being studied[18] for applicability. The fit neighborhood size can also be varied[18], but a 5 x 5 has been used in production. The peak is located by the use of the MDP algorithm, and the curvature matrix is used, in lieu of the curvature itself, in a variational approach to express the peak uncertainty covariance in terms of the covariance of uncertainty in the polynomial fit coefficients. Since the bivariate polynomial is fit by least squares, the coefficient uncertainty covariance can be estimated from the fit residuals. Often, relatively few points are fit, compared to the number of coefficients solved for (25 points for a 5 x 5 neighborhood, as compared to 15 coefficients for a fourth-order polynomial), and little confidence can be placed in the uncertainty estimate. To circumvent this problem, the user may supply an a priori fit variance of the fit uncertainty rather than determining it from the residuals. The a priori variance can be established by tests on representative data, whereby the a priori variance can be adjusted to make the estimated peak uncertainty agree with the observed peak displacement errors. Preliminary tests[18] established a value for early production in the JSC Registration Processor. A peak-to-background ratio is also established by subtracting from the correlation peak value the mean of samples away from the peak and then dividing by the standard deviation of those samples. The subtraction was included to compensate for the fact that the edge image means are not subtracted out in binary edge correlation.

Rather than establishing the order of the bivariate polynomial a priori, successive orders could be tried and the RMS residuals compared. When that error falls significantly and then starts tapering off, whatever order has been used at that point is adopted as the fit. As an attempt to sidestep the issue that a good model of the cross-correlation is not really at hand, this approach follows the rationale that a good representation lies somewhere between a simple model and a fit, with no remaining error, to every point, as a sufficiently high-order polynomial would do. Choosing an intermediate order, as outlined, attempts to balance a lack of understanding of the general form of the cross-correlation function with the knowledge that errors do exist in the data which makes forcing a perfect fit unwarranted. This problem of model uncertainties is encountered again in the section dealing with the mapping function.

Functions other than a bivariate polynomial can be fit to the cross-correlation samples. A bivariate gaussian function is an obvious choice; however, its use has not come to our attention. One registration application[19] chose an elliptical cone. The orientation of the cone was also a solution variable. Although surface fitting has been the commonest method of peak interpolation, it is by no means universally accepted. A principal criticism with that method involves the sensitivity of the intersample peak location to the particular form of the fitting function.

As an alternative to surface fitting, the peak can be assumed to lie at the centroid of the correlation neighborhood[20]. That is, the peak location is computed as the weighted sum of neighboring sample locations, with the correlation values serving as weights. Analogous to the peak uncertainty or surface curvature mentioned earlier, a corresponding measure of the breadth of the peak can be computed as the moment of gyration[21] or simply as the second moment of the correlation distribution (just as the centroid is the first moment). Unfortunately, little more can be said about this method at this time. It would be interesting to compare the centroid and surface fit methods for a representative set of correlation sample arrays.

Another method of peak location takes advantage of cross-correlation by Fourier techniques. The offset in the spatial domain corresponds to a non-vanishing phase in the frequency domain. In fact, if the cross-correlation function were symmetric about its peak, the phase function would directly specify the offset (since phase introduced by non-symmetry would vanish). One approach[22] transforms the phase portion of the correlation transform back to the spatial domain wherein the peak is located. The spatial result was noted to resemble closely a delta function so that subpixel peak location seems viable. The peak location was facilitated by using the inverse transform operation directly to compute a finer grid of spatial samples around the indicated peak than would normally be obtained from Fast Fourier Transform (FFT) algorithms. If indeed the transformed phase function is consistently nearly a translated delta function, then in the frequency domain the phase should be proportional to the spatial displacement. Thus, the phase vs frequency points can be fit by a straight line forced to pass through the origin, and the slope gives the translational displacement. However, somehow the possibility of non-symmetry may need to be accounted for. Preliminary results[18] have shown that binary edge correlations, for example, do exhibit nonsymmetry.

This section closes by returning to the issue of estimating peak location uncertainty or sharpness. The schemes described above can be categorized basically into curvature at peak, second moment of correlation samples around peak, and rate of decline away from peak. Any of these sharpness measures can be used to compare with a rejection threshold or to construct a weight for use in the mapping function fit. In the latter case the weight would be lower, the broader the peak lobe is. The weight is given directly by a peak uncertainty estimate (as the inverse of the uncertainty covariance). As mentioned earlier, a fixed a priori surface fit variance should probably always be used to estimate the peak uncertainty when fitting bivariate polynomials so that effectively its variance is directly related to the curvature. The sharpness measures can also be computed for the autocorrelation function to give an idea of the intrinsic clarity of the image. This application has been implemented as well in the JSC Registration Processor.

Residual Mapping Model

If the approach of the previous sections has been adopted, viz. local translational mismatches are determined at an array of control points, some means is needed to distribute the corrections for those mismatches over the regions in between the control points. Toward that end, a geometric transformation function is to be formulated in terms of the measured displacements. The nature of that function depends upon one's confidence in the measured mismatches versus one's understanding of the geometrical or physical processes causing the mismatch. If the control point translations are expected to contain errors, and the nature of the geometric distortions is well understood, then some sort of error minimization or maximum likelihood estimation technique is utilized. On the other hand, if the data are highly trusted and the model is unknown, some form of "rubber sheeting" technique is employed (analogous to curve-smoothing in one dimension). These alternatives are addressed below. The case that the model is understood is discussed first, with a breakdown into a geometry-based model (e.g., platform attitude error, trajectory error, etc.) and a sensor-based model (e.g., scan irregularities, band-to-band offsets, etc.). Then, the rubber sheet approach is considered. If uncertainty is associated with both the measurements and the model the attack is not so clear. This situation is considered briefly along with the situation that some measure of confidence in the model is available and the control point data are being used to improve it. It is well established that cross-correlation will occasionally result in an erroneous translational offset, and if such an error goes undetected it can ruin the mapping fit. The section ends with a description of several methods for identifying and handling such outliers.

Geometric errors are generally modeled as time varying satellite attitude and altitude errors. The MDP[3] assumes for MSS processing that the Landsat yaw, pitch, and roll can be modeled as third-order functions of time over the dimensions of a double frame (about 340 km of downrange). Similarly, the altitude variations are modeled as linear in time. The resulting 14 coefficients are solved for by weighted least squares utilizing both control point offsets and the less-precise attitude/altitude data available from Landsat's attitude measurement system (AMS) and ground tracking. The weights are set up a priori based on past experience in control point location accuracies and in AMS and

ground tracking accuracies. Hence, the control point weights are considerably larger than the others so that if many control point offsets are utilized the solution is essentially based on them. However, if few or none are available the AMS and ground tracking data do contribute appropriately to the solution. To handle the fact that the AMS data are poor for estimating the constant portion of the model (biases), a scheme is implemented to allow devoting the control point data to estimating the attitude biases, while the other measurements are utilized for rates and higher-order coefficients. Nominally, this scheme is employed only if a very few control points are available. Also, the attitude biases from a good solution (one with control points) are carried over to the next frame if that frame has insufficient control points (nominally, less than two). A weight associated with that carry-over is propagated in a manner that makes it decay with time, in order to model the growth in uncertainty of the biases as the satellite moves further away from the estimation point. This whole approach is predicated on the measurement errors being normally-distributed, an assumption which breaks down if one or several of the control point offsets are erroneous. To overcome that problem, the MDP has been tested with the threshold number of control points for triggering the bias estimation scheme raised to 15 from the nominal value of unity.[23] Hopefully, such a large number of control points will contain enough good offsets to overwhelm any erroneous ones. Another approach to eliminating the outliers is discussed later.

A TRW study[24] has adopted the same model as the MDP except for assuming a constant altitude deviation. The coefficients were evaluated by Kalman filtering. ERIM's geometric correction process also models the geometric errors as yaw, pitch, and roll, with pitch and roll assumed linear and yaw assumed constant over the area of interest.[25]

The GSFC LACIE processor models the geometric error (after the more-complicated a priori transformation) as pure translation over the dimensions of its 8 x 10-km sample segments. Performance evaluation[5,9] of that system indicated that, although specifications were met, there was a residual error that appeared to be affine. Thus, the JSC Registration Processor has adopted an affine model. A weighted least squares solution is made in which the weights are the reciprocals of the peak uncertainty standard deviations estimated from the correlation surface fit. At user option, default weights can be substituted if the correllation

surface fit fails in its estimation process, or weights can be eliminated altogether. A preliminary evaluation[26] of the JSC Processor has indicated that an affine model is adequate; that is, observed residual distortion does not show any systematic effect associated with an incomplete geometrical description. The success of the affine model lies in the fact that the JSC Registration Processor is only removing residual distortion left by the MDP. The a priori geometric correction has taken care of map projections, sensor geometry, and sensor line effects as characterized by the MDP data (a priori modeling constants and attitude/altitude solution).

Sensor effects can also be modeled and solved for by use of the acquired scenery. Alternatively, they can be assumed static, determined from a priori data, such as laboratory or flight tests, and then used in the a priori geometric correction. Landsat MSS effects include scan nonuniformity, band-to-band offset, "staircasing" caused by simultaneity of six scan lines, and line-by-line offsets due to sampling delays and changes in scan speed. The MDP, GSFC LACIE Processor, and JSC Registration Processor all assume a static model for those effects. All effects are assumed constant, except for the scan speed and scan nonuniformity. The latter is modeled as a third-order polynomial in position with the line (i.e., sample number). The coefficients were determined prior to implementation in the registration systems. The scan speed is assumed to have a fixed relation to the number of samples actually obtained in a scan line (line length majority). Although this model seems fairly good in normal circumstances, it breaks down for Landsat 3 when the MSS sample initiation fails (the "late line start") because the correct line length majority is not available. A model has been installed in the GSFC processing line to estimate the line length majority from the number of samples in the partial line. The resulting line length majority is used by both the MDP and the JSC Registration Processor. An inaccurate line length majority is manifested as a line jitter over portions of the image, and indeed a jitter is sometimes apparent as irregular field boundaries in some Landsat 3 agricultural scenes (especially at lower lattitudes where field boundaries tend to be parallel to lines and columns of pixels). An error of one in the line length majority will result in a misregistration of up to one pixel or about 58 m in MSS imagery. The possibility of variations in scan nonuniformity can also not be ruled out as the cause. These anomalies merit further investigation. Techniques for line-to-line registration to straighten field boundaries should be

considered, if for no other purpose than clarifying operational scan characteristics.

The sensor geometrical model for the Landsat RBV is considerably simplified by the presence of reseau marks on the face of the Vidicon tube. Also, the satellite attitude model is simplified by the fact that a whole frame is gathered at one instance, hence, for one value of roll, pitch, and yaw.

The rubber sheet approach to geometrical modeling assumes the form of the model is poorly understood. Whatever is the correct model, it should be possible to express it as a bivariate polynomial (e.g., assume a Taylor series). Thus, a common rubber sheet technique fits a polynomial to the array of control point offsets. The JSC Registration Processor and the GSFC Digital Image Rectification System[27] can utilize polynomials, the former to fourth order and the latter to second order. The principal concern is generally what order to use, although omission of certain terms may also be considered (such as sample-dependent terms if distortions seem to be from line to line). A sufficiently high order will fit all control points with dubious results in between. If the points are known rigorously to contain no error, then fitting all the points is reasonable, and the problem becomes one of selecting an interpolating function with suitable behavior between the points. Generally, the control points are assumed to have errors so that fitting them rigorously at the risk of inter-point error is not particularly suitable. A compromise is often achieved by trying a progression of successively higher-order polynomials and tracking the decrease of the error, expressed as the RMS of fit residuals. Assuming the idea that lower order is better for inter-point behavior, the fit is chosen for which the error has dropped substantially from lowest order but for which little error reduction is apparent at higher orders.

Another rubber sheet approach effectively spatial-filters the "sampled" offsets at the control points, thereby distributing the offsets over areas between the control points. This approach parallels the reconstruction of a two-dimensional analog signal from its samples by cubic or sinc convolution. The filter might be derived from maximum likelihood considerations such as a Kalman filter. This approach might be reduced to any desirable scale by reducing the size of control points while increasing their density so that a fairly

high-frequency matching process could be designed. When sensor resolution becomes good enough, atmospheric turbulence will be discernible as high-frequency distortion for which such a high-frequency geometry-matching filter would be appropriate.

Sometimes, the model may be partially known before control point information is acquired; that is, a priori values for the model parameters are available with an associated covariance characterizing the level of confidence. Then, the control point offsets provide additional information which can be used to improve the estimates of the model parameter. A Bayes approach is utilized. Simply stated, a least squares solution is performed on the control point offsets, and a weighted average of the results with the a priori estimate is computed. The weights are the inverses of the respective covariances normalized by matrix premultiplication by the inverse of the sum of the covariance inverses. This approach has been implemented in the JSC Registration Processor to account for the fact that it follows the MDP. Since the MDP has already geometrically corrected the data, the a priori values for the residual mapping function are zero. The covariance of the MDP's attitude/altitude solution is provided, and, after transformation to the covariance of a local (8 x 10-km sample segment) affine geometrical model, it serves as the a priori covariance. The JSC Processor performs its own least squares solution and utilizes its and the a priori covariances to average its solution with zero (the MDP's value). The appropriate expressions are summarized in Figure 2. Effectively, the result down-weights the JSC Processor's solution according to the size of the MDP covariance, with greater attenuation the smaller the MDP covariance. It must be borne in mind that all this theory is predicated on the MDP's and JSC Processor's error sources being normally-distributed. A crude compensation (shrinkage) is provided, as shown in Figure 2, in case they are not normally-distributed.

Unfortunately, control point correlations tend to either work reasonably well, or they are very bad. Thus, unless false fix detection is very good, erroneous control point offsets will be interspersed among the good ones. One straightforward approach to eliminating the outliers simply compares the residuals after the least squares fit to a preset threshold, or a threshold scaled by the estimated covariance. This technique was used in the GSFC LACIE Processor evaluation[5,9]. Control points with residuals greater than three standard deviations

were discarded, and the least squares solution was repeated. The new residuals were again tested, and the procedure was repeated until no more failures were noted.

Another approach performs a least squares solution with one control point omitted. Similarly, another solution is made with the first control point returned but a second omitted, and the process is repeated to yield as many solutions as there are control points. The post-solution residual for each omitted control point may then be examined for rejection, or a weighted combination of all the solutions may be made, with the weights related to the post-solution residuals.

Finally, a robust estimation[28] technique can be utilized in which the control point error is assumed to come, not from a normal distribution, but from a distribution with long tails which characterize the possibility of "outliers." Just as least squares is a maximum likelihood estimator for a normal distribution, maximum likelihood estimators can be formed for other longer-tailed distributions. A number of such solution techniques exist.[28] They down-weight measurements with large errors, while sacrificing a small amount of efficiency (i.e., not being quite as good as least squares) when the measurement errors really are normally-distributed. To our knowledge, none of these techniques have been explored for image matching.

Sizing and Placement of Control Points

The method of inter-image comparisons reviewed here involve the distributing of a number of patches within each image (the reference image and the registrant image). The patches are chosen small enough that the inter-image shifts within them may be considered to be purely translational; that is, scale, skew, and other interimage distortions have negligible effect when compared to the translation, viewed on the scale of the patch. Translation we will regard as zeroth order, amounting to a constant bias in the functions that relate coordinate values in one image to coordinate values in the other. Similarly, scale, rotation, and skew are first order distortions, and keystoning, etc., are higher order. In the zeroth order, coordinates $\underline{x}$ in the reference image and $\underline{y}$ in the registrant image are related

$$(x_i - y_i) = b_i \quad ,$$

while to first order they are related

$$(x_i - y_i) = \sum_j a_{ij} y_j + b_i \quad .$$

So, it is seen that the above description of the order corresponds to the order of the polynomial in the equations relating the coordinates. The essence of registration is to specify the mapping between the reference and registrant images, which is done in either "rubber-sheeting" form or by modeling. The latter uses rigid descriptions of the sensor geometry to permit an economy in describing the interimage coordinate relationship. So, rather than describing the coordinate relationship in the form above, coefficients are often drawn that relate the coordinates through platform ephemerides and known sensor geometries. This section interrelates with the previous one via economic and accuracy considerations; an optimum mapping method will require the fewest control patches to establish the desired degree of accuracy, since there is a high cost per correlation patch.

Number, location, and size of the patches used for control points is a very important consideration in the design of a registration system, from both economic and accuracy standpoints. A large part of the registration computation load is proportional to the number of the patches and to the second power of the patch dimension, so it behooves one to minimize the number and individual size of the patches. At the same time the interaction of the spatial distribution of the patches with the remapping model will affect the accuracy of the registration; so the location of the population of patches is a consideration to be combined with their number. One patch gives a single estimation of the translational correspondence between the images in the vicinity of the patch. If it be assumed that errors are not systematic in the estimation of that translational correspondence, more patches (in a neighborhood) are better in that the errors will tend to cancel. However, although there is no conceptual difficulty with having patches overlap, there is a limit to the number of functionally independent patches that can be drawn from any image. That number puts an upper limit to the tendency for growth to larger number of patches in order to have statistical cancellation of errors.

There is a tradeoff in the determination of the optimum size of a correlation patch that is dependent on the accuracy with which the a priori corrections can account for the encountered misregistration. Consider first a purely translational misregistration. As the size of the patch gets larger, there is more picture structure within it; although the value of the correlation peak is not affected, at translational locations off the peak there will be (generally) lower cross-correlation values due to the larger amount of structure within the window. The result is an enhancement of the peak-to-background ratio (the value of the peak correlation divided by the correlation in the general vicinity), enabling increased sensitivity in detecting a peak and perhaps increased accuracy in determining the location of the peak.

Now, however, consider the ramifications of first-order departures from the condition in which there are only translational offsets between the images being registered (i.e., suppose there are skew and scale differences as well as translation). In the position in which the true centers of the patches being correlated are coincident (that is, the position one would want to find as a result of the cross-correlation process), picture structure towards the edges of the patches will be relatively displaced by scale and skew differences. In contrast to the translation-only situation, the correlation in this case is reduced as the size of the image is increased. But the presence of effects other than translational offsets does not mandate the use of the smallest possible patch--one pixel!--because the correlation is at its noisiest there. For a binary extracted feature on which correlation is being done, the correlation takes on only the values zero or one for the one-pixel patch, and it would be impossible to locate a peak. So one determines an optimum patch size by considering a reasonable envelope for the distortions (beyond translational) that one expects, and works out the spatial range over which they would cause significant displacements for the translationally-correct position.

Where appropriate by necessity of increased registration accuracy, a reentrant technique is possible. In that technique, patches of the original images are extracted and run through the correlation process. The resulting remapping for the whole image is then used to reextract the correlation the correlation patches, which can now be at least first-order corrected for the modeled distortions. The patches are extracted from the input imagery by

resampling onto a grid derived from the first registration process, and consequently larger patches, with potential for increased accuracy, may be drawn. The cost of such a reentrant technique is not necessarily to double the computation time, inasmuch as a coarser registration job is enabled for the first pass.

As mentioned earlier, the strategy for location of the correlation patches is important. In the case of the MDP, there is a scattering of patches amounting to approximately 0.1% of the scanned area (say 10 patches of 32 x 32 pixels within a frame of 3596 x 2983 pixels). In order for this small a correlated area to control the overall registration accuracy to an acceptable level, the patches must be located efficiently so as to exercise reasonable control over the entire image. Further, the modeling of the scanner behavior between control points becomes of paramount importance. In the other extreme, the JSC Registration Processor actually uses something over 100% of the area of the registrant image, when it is taken into account that the patches overlap slightly (about 7 pixels/line).

The registration of two images logically breaks into two philosophically different kinds of operations--the determination of the mapping of the coordinate grid of the reference image into the coordinates of the registrant image onto that remapped grid of the reference image. The mapping process is guided by the selection of patches within which interimage comparisons give the local estimate of the mapping (typically just the translational characteristics are determined on an area small enough that translation is the dominant effect over the size of the patch). Since considerable computation effort goes into the interimage comparisons done on the patches, and since the accuracy of the overall registration depends linearly on the accuracy with which a single patch is compared between images, and also since the model, usually driven by the translational offsets determined at the patches, will be more tightly and accurately determined with a "sufficient" number and distribution of patches, it is very important that the following qualities are achieved in the location of the control patches.

1. The comparison method allows an accuracy of interimage comparison at least as good as the registration accuracy desired for the whole process. Otherwise one must use a superfluity of control patches and hope that they

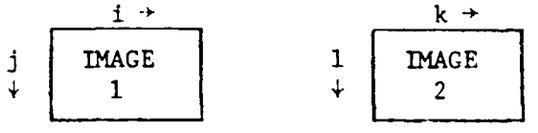
contribute independent measurements, with error cancellation, to the remapping model.

2. The model being driven by the control patches is sufficiently highly detailed to account for the entire behavior of the relative geometry in between control points.

3. The locations of the control points are chosen to interact sensitively with the modeling geometry. For example, if a sinusoidal component to the interimage geometry is hypothesized, one would not want control points (the centers of the patches) placed at uniform one-period separations within the image. If they were, aliasing would have the result of a sinusoidal relative geometry look the same as a translational offset.

This section discusses some aspects of location of control points (taken as the centers of their control patches).

Let the registration of Image 1 to Image 2 be defined as the mapping of the coordinates (i,j) in Image 1 to the coordinates (k,l) in Image 2 (i.e., we will do only the geometric portion of the problem).



Suppose the "true" relationship is given by F_1 and F_2

$$i = F_1(k,l) \qquad j = F_2(k,l);$$

let them be estimated as a result of the interimage correlation and modeling by

$$f_1 = \hat{F}_1, \qquad f_2 = \hat{F}_2 .$$

Then the root-mean-square registration error (RMSE) is

$$\text{RMSE} = \sqrt{\frac{1}{\text{area}} \iint_{\text{image}} [(F_1-f_1)^2 + (F_2-f_2)^2]\,dA}$$

where $\iint_{\text{image}}$ is the integral over the area of the image, and various assumptions are included regarding orthogonality of the coordinate axes, sameness of scale on those axes, etc. Further, suppose the mapping f_1 and f_2 to be derived according to some sensor/platform model, or according to other limitations on the complexity such as a mapping of no higher order than affine, and that the mapping f_1 and f_2 is determined by a finite number of control points within the images. If the images are not initially pathologically misregistered, we may reasonably express the mapping by beginning with the simplest of mappings,

$$f_1(k,l) = k \ , \qquad f_2(k,l) = l$$

and then include the deviations g_i,

$$f_1(k,l) = k + g_1(k,l), \qquad f_2 = l + g_2(k,l).$$

This slightly complicates the expression for RMSE but facilitates the analysis of sensitivity with respect to control point location. Translational offsets are written

$$\begin{aligned} g_1(k,l) &= a_o \\ g_2(k,l) &= b_o \end{aligned} \quad ;$$

affine mappings have the general form

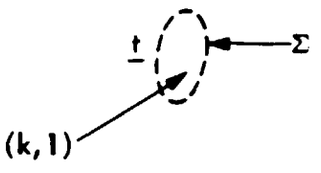
$$\begin{aligned} g_1(k,l) &= a_o + a_1 k + a_2 l \\ g_2(k,l) &= b_o + b_1 k + b_2 l \end{aligned} \quad ;$$

keystoning is modelable by the inclusion of a cross term,

$$\begin{aligned} g_1(k,l) &= a_o + a_1 k + a_2 l + a_3 kl \\ g_2(k,l) &= b_o + b_1 k + b_2 l + b_3 kl \end{aligned} \quad ,$$

etc.

We consider a control point established between the images. Let $\underline{t}(k_1 l)$ be the derived estimate of the "true" offset between the images, and let Σ be the estimated covariance of the estimate $\underline{t}$.



By a method unspecified here (so as to preserve generality), the set of control points {km, lm, $\underline{t}$m, Σm} is then used to generate the mapping functions g_1 and g_2. Let us make some simplifying assumptions:

1. The method of obtaining g_1 and g_2 will work with any number of control points (e.g., no control points gives g_1 and g_2 identically zero, one control point allows a translational solution, etc.).

2. The method of obtaining g_1 and g_2 is tractable to linearization.

3. A priori estimates of Σ are available (they might be in the form

$$\Sigma_{ap} = \text{diag}(\sigma_1^2, \sigma_2^2) \text{ or } \Sigma_{ap} = \sigma^2 I \ .$$

4. The a priori estimates of Σ are independent of location.

5. The a priori estimates of g_1 and g_2 are zero. (Otherwise the image could be adjusted according to the a priori estimates so that this would be true.)

We are now in position to find the sensitivity of the mapping functions g_1 and g_2 to the location of a single control point, that sensitivity being a function of location of control point in the reference image and of location in the registrant image. The sensitivity can be converted into an rms value by integration over the registrant image, as will be seen, and thus an ideal location of the first control point is specified.

Defining $\underline{y} = \begin{pmatrix} i \\ j \end{pmatrix}$, $\underline{x} = \begin{pmatrix} k \\ l \end{pmatrix}$, we have used the set $\{\underline{x}, \underline{t}, \Sigma\}$ to produce the (vector) function $\underline{g}$ such that

$$\underline{y} = \underline{g}(x) + \underline{k}.$$

The fact that the variation in, say, the first coordinate of $\underline{y}$ is given by

$$\delta y_1 = \frac{\partial g_1}{\partial x_1} \delta x_1 + \frac{\partial g_1}{\partial x_2} \delta x_2$$

leads us to the vector variation in $\underline{y}$

$$\delta \underline{y} = \underline{\underline{J}} \, \delta \underline{x} \ ,$$

in which $\underline{\underline{J}}$ is the Jacobian matrix

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{pmatrix} \equiv \underline{\underline{J}} \ .$$

The covariance in $\underline{y}$ is calculated from Σ_x, the covariance in $\underline{x}$, as follows.

$$(\delta\underline{y})(\delta\underline{y})^T = (\underline{\underline{J}}\,\delta\underline{x})(\underline{\underline{J}}\,\delta\underline{x})^T$$
$$= \underline{\underline{J}}\,\delta\underline{x}\,\delta\underline{x}^T\,\underline{\underline{J}}^T \ .$$

Taking expectation values, and assuming that the expectation of $\delta\underline{x}\,\delta\underline{x}^T$ is as estimated from the considerations of the cross-correlation surface for the point,

$$\Sigma_y = E\{(\delta\underline{y})(\delta\underline{y})^T\} = \underline{\underline{J}}\,E\{(\delta\underline{x})(\delta\underline{x})^T\}\,\underline{\underline{J}}^T$$
$$\Sigma_y = \underline{\underline{J}}\,\Sigma_x\,\underline{\underline{J}}^T$$

which will be a non-constant function of $\underline{y}$ inasmuch as the partials forming $\underline{\underline{J}}$ are not constant. The trace of Σ_y gives, as a function of $\underline{y}$, the squared registration uncertainty. We can average over the image and take the square root.

$$\text{RMSE} = \sqrt{\frac{1}{\text{area}} \iint_{\text{image}} \text{tr}\,\Sigma_y \, dA} \ .$$

Alternatively the trace can be taken after integration (the operations communite):

$$\text{RMSE} = \sqrt{\frac{1}{\text{area}} \text{tr}\left(\iint_{\text{image}} \Sigma_y \, dA\right)} \ .$$

Note that the RMSE so calculated is a function of $\underline{x}$, the location of the first control point. If a priori values for Σ_x and $\underline{t}\,(\underline{x})$ are used (the latter being zero), we have an a priori estimate of the sensitivity, measured in terms of coordinates in the image, as a function of location of the first control point. Fixing the location of the first control point at the location that minimizes RMSE, then an exactly similar procedure leads to the location of the best place to position the second control point. The RMSE should monotonically decline as the number of optimally-placed control points increases, and the system designer can quit adding control points when the anticipated accuracy has gotten to requirement level.

It is true that the placement of n control points by this technique will (probably) not produce as good accuracy as n points placed optimally as an ensemble. The algorithm for determining that optimal ensemble position, however, is a substantially more complicated one. The following relaxation technique is such an algorithm. The n points are placed according to some scheme in the image, and RMSE is calculated in the general manner as described above. The position of the first point is permitted to vary about its original location, and RMSE is observed; the location is altered until a minimum of RMSE against first-point-location is found. Then, the second is allowed to vary, and its minimum location found. Similarly through the set of n control points, and the procedure is iterated for all the control points again until each point is at a minimum. But, even this relaxation technique guarantees only a relative minimum, for small deviations of control point location. The sequential technique is likely to require more points to achieve a given level of anticipated accuracy than the minimum number meeting the requirements with the relaxation technique, but the sequential technique is, as mentioned earlier, more tractable.

This problem, optimizing locations of control points, is analogous to the statistical problem of selecting a set x_i at which to evaluate the associated set y_i, the combination to drive a regression between x and y. It has proven more practical to select an oversupply of x_i's and drop members one at a time according to a utility criterion, until the required performance would fail if any further members were dropped. A practical scheme for location of control points could proceed similarly.

All too often, a desired control point location, even with the allowance of small deviations around it, will not contain picture structure to permit sufficiently accurate cross-correlation. The strategy for control point location should be robust against failure to find proper structure at a subset of the ideal control point locations. An iterative scheme is envisioned, in which that failure would send one back to the image to obtain another set of control points.

These techniques seem clearly impractical for one-shot application. The amount of time spent in the computations could better have been spent in getting on with the registration problem by casting an overkill of control points into the image, and just proceeding. The application for such techniques is in a

production environment, where the fewest control points with the most advantageous effect is desired. Using the remapping model and the a priori estimates of the covariance, a general optimal strategy for control point location is drawn. The covariance estimation is done either entirely without prejudice ($\Sigma_{ap} = \sigma^2 I$), or if a large number of images are to be registered against a single base image (as in the authors' case, multitemporal agricultural images of certain ground areas), that base image can be investigated for the estimations of Σ.

To the authors' knowledge, this sort of optimal placing of control points has not been done. The MDP uses hand-picked control points with control in any given frame of Landsat data coming from not only that frame but from adjacent frames as well. In the LACIE Processor[4], there was essentially only a single point, the entire image; translational registration alone was done, with nearest-neighbor resampling following translational correlation done on whole-segment basis. An experiment[19] associated with the LACIE Processor utilized five control points placed at the corners and center of the image. In the DAM Package[30], manual control points are selected with admonishments to spread them out well in the image being registered. In the JSC Registration Processor[6,7], the control points are placed on a uniform grid, the patch size and spacing relating so as to cause overlap. In the "automatic" registration system[31] for the LACIE ground truth segments, control points are initially placed on a uniform grid, with small deviations allowed in order to find a suitable correlation peak. In all the processors with which the authors are familiar, it has been the practice to follow one's instincts rather than to code in a sophisticated patch location algorithm. The success of those processors indicates the efficacy of the less-elaborate methods, but it remains possible to make algorithmic improvements to permit sufficient accuracy with minimum computation load.

Non-Control Point Methods

The previous sections dealt with the step-wise manner of first matching images within local patches or control points and then fitting the full mapping function to the resultant array of point-shifts. Although the complexity of straightforward multidimensional correlation (say, in translation, rotation, scale change, etc.) generally makes it prohibitive, a fortuitous property of affine distortions under Fourier transformation has given rise to such a method[32]

which may show promise. It is shown that even with affine distortions (i.e., rotation, scale change, and skew or shear distortion, in addition to pure translation) the translation maps into a phase factor in the frequency domain, and the modulus of the transform function exhibits a similar, affine distortion. This fact, coupled with the fact that typical agricultural scenery shows considerable spatial structure, facilitates the use of Fourier transforms for matching in an easier way than with the images themselves. The structure of agricultural scenes generally casts the major portion of the spectral energy (i.e., transform modulus) along two nearly-perpendicular axes, corresponding to the presence of rectangular field and road structure. In that case the axes of the two transformed images can be sought out by searching for lines of maxima, and the matching need only occur over the axes, rather than over the whole domain. The rotation and skew are determined by rotating corresponding axes into coincidence, and scale changes are determined by matching energy distributions along the axes. Once the linear portion of the affine transformation has been determined (wholly from the modulus), the translation is determined from the phase.

The method was used successfully[32] to determine rotation and skew distortion in airborne scanner data. It is not certain that modulus-matching along the axes for scale change determination is sensitive enough to give a good scale solution because the modulus probably changes far more slowly along an axis than traverse to it. This question should be addressed and the method tried on other data types and scene types before further conclusions are drawn. At present, the method does seem applicable to agricultural images in which only translation, rotation, and skew are present.

Summary and Conclusions

A number of image correlation, match offset determination, control point placement, and geometrical modeling techniques are in use today, but there is still considerable ground to cover. There appears to have been little effort in interchanging methods, e.g., trying different offset determination techniques with each matching technique, etc. Perhaps the greatest potential for correlation improvements lies in increased computational efficiency and speed. Several new matching techniques appear to show promise, but they need to be tested on a wider range of image data. There is room for improvement in the robustness of

offset determination schemes. The various schemes need to be compared when used with representative data. Improved techniques for identifying false fixes without rejecting good fixes are seriously needed.

Methods of modeling sensor anomalies such as late line start perturbations need further investigation. Is the current MSS scan nonuniformity model adequate, or does it change with time? More attention to the degrading effects of geometric distortion on cross-correlation is needed by evaluating tradeoffs between patch size and iterative registration. Control point placement techniques need to be compared in terms of overall performance; particularly, the adaptive placement scheme described earlier needs to be tested. Robust estimation techniques should be investigated for their applicability to registration.

TEMPLATE MATCHING:

$$\frac{\sum_{\text{WINDOW}} (\text{REFERENCE PIXELS}) \times (\text{SEARCH PIXELS})}{\sum_{\text{WINDOW}} (\text{SEARCH PIXELS})^2}$$

- FOR EDGE IMAGES:

$$\frac{\text{NO. OF COINCIDENCES}}{\text{NO. OF SEARCH EDGES}} \approx P(\text{SEARCH PIXEL=EDGE} \mid \text{REF. PIXEL=EDGE})$$

"CLASSICAL" CORRELATION:

$$\frac{\sum_{\text{WINDOW}} (\text{REFERENCE PIXELS}) \times (\text{SEARCH PIXELS})}{\left[\sum_{\text{WINDOW}} (\text{REFERENCE PIXELS})^2 \times \sum_{\text{WINDOW}} (\text{SEARCH PIXELS})^2\right]^{1/2}}$$

- FOR EDGE IMAGES:

$$\frac{\text{NO. OF COINCIDENCES}}{\sqrt{\text{NO. OF REF. EDGES}}\ \sqrt{\text{NO. SEARCH EDGES}}}$$

$$= \sqrt{P(\text{SEARCH=EDGE} \mid \text{REF=EDGE}) P(\text{REF=EDGE} \mid \text{SEARCH=EDGE})}$$

Figure A. Normalization Options For Registration Cross-Correlation

A-PRIORI (MDP) MAPPING FUNCTION:

MAPPING = NULL

COVARIANCE $\underline{C}_M$ = MDP COV. OF REG. + MDP COV. OF REF.

RESIDUAL MAPPING FIT:

MAPPING ≈ AFFINE $\quad \underline{X}_R = (\underline{A}^T \underline{W} \underline{A})^{-1} \underline{A}^T \underline{W} \underline{Y}$

COVARIANCE $\underline{C}_R = (\underline{A}^T \underline{W} \underline{A})^{-1}$

COMBINATION OF FIT AND A PRIORI

$$\underline{X} = (\underline{C}_M^{-1} + \underline{C}_R^{-1})^{-1} \underline{C}_R^{-1} \underline{X}_R$$

$$\underline{C} = (\underline{C}_M^{-1} + \underline{C}_R^{-1})^{-1}$$

CAN INSERT ADDITIONAL SHRINKAGE α

$$\underline{X} = (\alpha \underline{C}_M^{-1} + \underline{C}_R^{-1})^{-1} \quad \underline{C}_R^{-1} \underline{X}_R$$

$$\underline{C} = (\alpha \underline{C}_M^{-1} + \underline{C}_R^{-1})^{-1}$$

Figure B. Weighted Incorporation of Residual Mapping Fit

References

1. W. K. Pratt, Digital Image Processing, Wiley, New York, 1978.

2. R. Bernstein, "Digital Image Processing of Earth Observation Sensor Data," IBM J. Res. Develop., 20, 40-56 (1976).

3. "MSS Data Processing Description," IBM Federal Systems Division report, Contract NAS5-22999, November 1978.

4. G. J. Grebowsky, "LACIE Registration Processing," in The LACIE Symposium, Proceedings of the Technical Sessions, 87-97, July 1979.

5. T. Kaneko, "Image Registration Accuracy Evaluation by an Edge Method," IBM memorandum to NASA JSC, Contract NAS9-14350, June 4, 1976.

6. "ERSYS Registration Subsystem Level C Requirements," AgRISTARS Report SR-I1-00203, Contract NAS9-14350, JSC No. 17620, September 24, 1981.

7. "ERSYS Registration Subsystem Detailed Design Specification," AgRISTARS Report SR-I1-04154, Contract NAS9-14350, JSC No. 16946, September 25, 1981.

8. D. I. Barnea, and H. F. Silverman, "A Class of Algorithms for Fast Digital Image Registration," IEEE Trans. Computers, C-21, 179-186 (1972 - reprinted in Digital Image Processing for Remote Sensing, edited by R. Bernstein, IEEE Press, Wiley, New York, 1978, pp. 138-145).

9. T. Kaneko, "Evaluation of Landsat Image Registration Accuracy," Photogrammetric Engineering and Remote Sensing, 42, 1285-1299 (1976).

10. H. Mostafavi and F. W. Smith, "Image Correlation with Geometric Distortion Part I: Acquisition Performance," IEEE Trans. Aerospace and Electronic Systems, AES-14, 487-493 (1978).

11. H. Mostafavi and F. W. Smith, "Image Correlation with Geometric Distortion Part II: Effect on Local Accuracy," IEEE Trans. Aerospace and Electronic Systems, AES-14, 494-500 (1978).

12. T. L. Steding and F. W. Smith, "Optimum Filters for Image Registration," IEEE Trans. Aerospace and Electronic Systems, AES-15, 849-860 (1979).

13. C. D. Kuglin and W. G. Eppler, "Map-Matching Techniques for use with Multispectral/Multitemporal Data," in Image Processing for Missile Guidance, Proc. SPIE, 238, 146-155 (1980).

14. D. J. Kahl, A. Rosenfeld, and A. Danker, "Some Experiments in Point Pattern Matching," IEEE Trans. Syst. Man. Cybernet., SMC-10, pp. 105-116 (1980).

15. S. Ranade and A. Rosenfeld, "Point Pattern Matching by Relaxation," Pattern Recognition, 12, 269-275 (1980).

16. T. C. Minter, Jr., "Minimum Bayes Risk Image Correlation," in Image Processing for Missile Guidance, Proc. SPIE, 238, 200-208 (1980).

17. T. C. Minter, Jr., private communication, November 1981.

18. A. G. Wacker, R. H. Wolfe, Jr., and R. D. Juday, to be published.

19. E. W. Cordan, Jr., and B. W. Patz, "An Image Registration Algorithm Using Sampled Binary Correlation," in Proc. 1979 Machine Processing of Remotely Sensed Data Symposium, 202-212 (1979).

20. Canadian Centre for Remote Sensing (A. G. Wacker, private communication) April 1981.

21. N. Chu, private communication, 1980.

22. J. J. Pearson, D. C. Hines, Jr., and S. Golosman, "Video-Rate Image Correlation Processor," in Application of Digital Image Processing, Proc. SPIE, 119, 197-205 (1977).

23. G. J. Grebowsky, private communication, November 1981.

24. S. S. Rifman, A. T. Monuki, and C. P. Shortwell, "Multi-Sensor Landsat MSS Registration," in Proc. Thirteenth International Symposium on Remote Sensing, 245-258 (1979).

25. R. Dye, presentation to NASA JSC, 1979.

26. R. H. Wolfe, Jr., presentation to NASA JSC, July 1981.

27. P. Van Wie and M. Stein, "A Landsat Digital Image Rectification System," in Proc. Symposium on Machine Processing of Remotely Sensed Data, 4A-18 - 4A-26 (1976).

28. R. V. Hogg, "Statistical Robustness: One View of Its Use in Applications Today," The American Statistician, 33, 108-115 (1979).

29. M. L. Nack, "Rectification and Registration of Digital Images and the Effect of Cloud Detection," Proc. Fourth Annual Machine Processing of Remotely Sensed Data Symposium, 12-23 (1977).

30. E. H. Schlosser, "Detection and Mapping Package, Vol. 3: Control Network Establishment," Johnson Space Center document JSC-11379, June 1976.

31. "'As-Built' Design Specification of the Automatic Registration System for the Cartographic Technology Laboratory," LEMSCO Report 15904, Contract NAS9-15800, JSC No. 17017, December 1980.

32. R. A. Emmert and C. D. McGillam, "Multitemporal Geometric Distortion Correction Utilizing the Affine Transformation," in Proc. Conference on Machine Processing of Remotely Sensed Data, 1B-24 - 1B-32 (1973 - reprinted in Digital Image Processing for Remote Sensing, edited by R. Bernstein, IEEE Press, Wiley, New York, 1978, pp. 153-161).