Reprinted from

Symposium on Machine Processing of Remotely Sensed Data

June 29 - July 1, 1976

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
76CH1103-1 MPRSD

Experimental Examination of Similarity Measures and Preprocessing Methods Used for Image Registration

M. Svedlow, C. D. McGillem and
P. E. Anuta

Laboratory for Applications of Remote Sensing
Purdue University
W. Lafayette, Indiana

I. ABSTRACT

An experimental analysis of a number of image registration techniques is described. The objective is to provide a better understanding on a comparative basis of some of the methods of registering imagery that have been proposed. In particular, this study encompasses the choice of a similarity measure and the effects of preprocessing the imagery prior to the registration.

II. INTRODUCTION

Image registration is an integral part of the analysis of multitemporal data. Extraction of the information contained in changes of a scene from one time to the next generally requires that the images of the scene at these two times be matched spatially as close as possible to provide a comparison on a point by point basis. Conceptually this registration is easy to understand. However, it is considerably more difficult to find the processor that will accomplish this registration in an optimum fashion.

The purpose of this study is the experimental examination of several approaches to certain aspects of this problem. In particular, the problem considered is that of finding the relative translations with respect to each other of images which are not spatially distorted. Several algorithms for registering images have been proposed or implemented [1,2,3,4,5,6]. However, each was developed and tested independently of the others. This is the situation which provided the impetus for this study. The question to be answered is: which method of image registration should one choose and why?

The research reported in this paper was supported by NASA Contract NAS9-14016.

This study experimentally explores the basic concepts which underlie these algorithms. The first part of the study examines the criterion one uses to measure the similarity between images and thus find the position where the images are registered. The three similarity measures considered are the correlation coefficient, the sum of the absolute differences, and the correlation function. Table 1 contains the formulas for these measures. This particular choice was made since it represented those measures actually used.

Secondly, preprocessing of the imagery prior to the actual registration and its effect on the results is examined. Suggestions that preprocessing might enhance the performance or save operational time prompted this part of the study. Three basic types of preprocessing were picked: taking the magnitude of the gradient of the images (equation 1), thresholding the images at their medians (all values greater than or equal to the median are set equal to one, and all else set equal to zero), and thresholding the magnitude of the gradient of the images at an arbitrary level to be determined experimentally. Again, this particular choice was made to approximate the preprocessing methods that had been proposed or implemented.

Finally an extension of the results of these two sections is made to an algorithm [6] designed for operational image registration which has been implemented.

III. SIMILARITY MEASURES

When one considers the approaches to the registration problem, a preliminary decision that must be made is, what criterion should one use to evaluate the similarity between two images: that is, what similarity measure should be selected.

The similarity measures being considered can be divided into two general classes.

The first provides a measure on an absolute scale. An example of this is the correlation coefficient [Table 1], where the values range between plus and minus one. Not only is the scale limited for the correlation coefficient, but its value on that scale gives an indication of how good the images are linearly related.

The second class indicates the registration position by a maximum or minimum value at the registration location. Two examples of this are the correlation function and the sum of the absolute value of the difference between the two images [Table 1]. In these examples there is no absolute scale, so that the value of this maximum or minimum by itself will not give a good indication of how closely the two images match. The exception to this occurs in the absolute value of the difference case when the two images are identical. A value of zero then implies that the two images match perfectly. However, if one models the difference between the two images as additive noise, one may establish a confidence interval in the absolute value of the difference case by using the resulting minimum value in conjunction with the probability distribution of the noise.[4,5]

The choice one must make with regard to the similarity measures is influenced by several factors. (1) How well do the different methods perform? Is there a way to theoretically predict this performance, and if so, what are the results? Also included in this question is whether there exists some sort of confidence measure so that one may evaluate the results quantitatively. (2) What operations are involved for each of the methods, and what are the comparative times needed? (3) If it has been determined that several methods of registration yield reasonable results with respect to the ability to find the correct registration position, then what are the tradeoffs between the accuracy and the time and number of operations involved? For example, if one method yields the correct registration position in 95% of the attempts, but requires twice the operational time as a method which is able to find the correct location 75% of the time, which method should be used? One criterion that is essential for this decision is whether the occurrence of a false indicated registration position is known to be false when it appears.

For the experimental analysis, test sites were chosen from LANDSAT imagery over Missouri and Kansas. The primary reason for this particular choice was the availability of multitemporal data. The subimages registered were 51 lines by 51 columns in size.

Evaluation of the results is in terms of the percentage of acceptable registrations out of a given number of attempts. The nonacceptable attempts are those where the indicated registration location was known to be false. Such a criterion clear[ly] requires some a priori knowledge of the relative translation between the images in question. For the Missouri imagery three temporally differing sets of data had been previously registered to within a few pixe[ls] via the LARS registration system [1], so that any substantial deviation from this was taken as an unacceptable attempt. For the Kansas data this a priori information was supplied by careful visual checking of the imagery.

The overall acceptability comparisons are listed in Table 2. The results are tabulated for both the original and preprocessed imagery so that one can also cross-reference a particular similarity measure among the different types of imagery registered.

Between the three similarity measures examined, the correlation coefficient consistently yielded the highest percentage of acceptable registrations. So that on a performancewise basis, these results indicate that one should use the correlation coefficient as the similarity measure.

However, what about the tradeoff between operational time required and performance? Is there a measure which reduced the reliability only slightly while accompanied by a large time savings? Refer to the percentage acceptable registrations in Table 2 for the magnitude of the gradient of the imagery. Note that while there was 100% acceptability using the correlation coefficient, there also was a 92% performance with the sum of the absolute difference measure. This result in conjunction with the time savings achieved by using this latter measure (Table 1), indicates that in a time-performance evaluation, it might be more advantageous to use the sum of the absolute difference measure as opposed to the correlation coefficient.

Overall,the best performance was achieved by the correlation coefficient using the magnitude of the gradient of the imagery. Therefore, if percent acceptabili[ty] is of prime importance, this preliminary comparison indicates that preprocessing the imagery via a gradient type processor enhances the ability to find an acceptable registration position. The next section concerning the effects of preprocessing prior to registration pursues this observation in more depth.

IV. PREPROCESSING METHODS

In the search for an optimum processor for image registration it has been proposed that preprocessing of the data prior to the actual overlaying procedure may be a step towards the solution of this problem. There are several underlying reasons for this suggestion. (1) With a given similarity measure, preprocessing may yield a greater reliability of the system's registration performance. (2) The time and operations required may be substantially reduced. An example of this is conversion of the original image into a binary image (data values of only 0 or 1) so that logical operations may be used.

The study undertaken here is an experimental examination of several preprocessing techniques and their effects on image registration. Three basic methods were chosen. The first method utilizes the magnitude of the gradient of the imagery given by,

$$|\text{Gradient of } X_{ij}| = [(X_{i+i,j}-X_{i-1,j})^2 + (X_{i,j+1}-X_{i,j-1})^2]^{1/2}. \quad (1)$$

The second method consists of thresholding the imagery at its median (all values greater than or equal to the median are set equal to one, and all else set equal to zero) and the third method computes the magnitude of the gradient of the imagery and then thresholds it at an appropriate level. Typical images resulting from carrying out these preprocessing operations are shown in Figure 1.

LANDSAT imagery over Hill County, Montana; Tippecanoe County, Indiana; and Kansas were used for the analysis. The ready availability of multitemporal data prompted these particular choices. The actual subimage sizes that were to be registered for these comparisons were 51 lines by 51 columns.

Again, evaluation of the performance is in terms of the percent of acceptable registration attempts. Like the similarity measure comparisons, visual examination or previous registration to within a few pixels provided the a priori information for determining the acceptability of an indicated registration position. Also, in order to provide a common basis for comparison, the correlation coefficient was chosen as the similarity measure for all of the attempted registrations.

The acceptable-unacceptable attempts are tabulated in Table 3. Note that the results have been divided into three sections: (1) the cases where the magnitude of the correlation coefficient ($|\rho|$) for the original imagery is greater than or equal to 0.5, (2) the $|\rho|$ for the original imagery is less than 0.5, and (3) the overall results. The underlying reason for this partition is to examine the relative performance for the high correlation cases ($|\rho| \geq 0.5$) and the low correlation cases ($|\rho| < 0.5$) separately, as well as for the overall results.

First, consider the overall results. Preprocessing the imagery via the magnitude of the gradient yielded the highest percent acceptability (with 100%). Also, thresholding the magnitude of the gradient performed very well (97%). The important point to note, aside from the best performance, is on an overall basis preprocessing of the imagery with a gradient type transformation boosted the performance over that utilizing the original imagery. This indicates that preprocessing may indeed provide a more reliable registration processor.

We might now ask, is there any trend to this increased reliability? Are there any image characteristics which seem to contribute to these observations? One answer to these questions is embodied in the partitioning of the overall results into the high and low correlation cases.

Examination of the high correlation instances ($|\rho| \geq 0.5$ for the original data) shows that all of the imagery types performed exceedingly well with 96% acceptability for thresholding the data at its median and 100% for the rest. This indicates that when the original imagery is highly correlated, any of the preprocessing methods works equally well. In this case one does not gain any advantage performancewise by preprocessing the imagery prior to registration.

The most striking result came with the low correlation cases ($|\rho| < 0.5$ for the original data). For these cases one obtained a marked advantage over using the original imagery by preprocessing the data via a gradient type processor. Use of the magnitude of the gradient of the imagery provided a 100% acceptability compared with the 65% performance for the original data. Thresholding the magnitude of the gradient also indicated a distinct increase in reliability. These results suggest that one may achieve a substantial increase in the reliability of the registration processor when the original imagery is not highly correlated by preprocessing the imagery prior to registration.

Earlier it was mentioned that a priori information was used to determine the acceptability of indicated registration positions. For imagery that had not been previously registered this took the form of visual examination for an individual test site. Such a procedure is quite time consuming and does not lend itself readily to an automatic mode of operation. However, while attempting the registrations at the selected test sites it was found that relative spatial information could be used for several test sites located in the same general area, or the same test site over several different times. For example, if several different test sites indicated the same relative translation for registration, while the registration position of another test site within the same general area indicated a substantially different translation, then this latter registration attempt would be unacceptable. Similar reasoning follows for several time pair registration attempts for a single test site.

Another observation which may be made directly from Table 3 also suggests a way by which a partial acceptability decision might be made automatically. This approach is in terms of an absolute scale confidence measure. Since the value of the correlation coefficient (ρ) indicates the linearity of the relationship between two images, possibly a range of values for ρ exists which could be used to determine acceptability. This is suggested in the first line of Table 3 where the results when the magnitude of the correlation coefficient is greater than or equal to 0.5 for the original imagery are listed. For the original data there is a 100% acceptability for this range of ρ. This suggests that one may be able to use the value of the correlation ceefficient for determining the acceptability of an indicated registration position.

V. EXTENSION OF THE RESULTS IN SECTIONS III AND IV TO A REGISTRATION ALGORITHM THAT HAS BEEN IMPLEMENTED

The observations made in sections III and IV suggested that a very reliable registration processor could be developed by first preprocessing the images with a gradient type processor and then using the correlation coefficient as the similarity measure. Independent of this experimental study, but at approximately the same time, an algorithm for registering imagery was developed at the Computer Sciences Corporation [6]. This algorithm embodies the two basic observations made above: a gradient type preprocessor is used, and the similarity measure closely approximates the correlation coefficient. The availability of this algorithm made it possible to experimentally observe the extension of the results obtained in the similarity measures and preprocessing comparisons to an algorithm designed for operational image registration.

A simulation of this registration algorithm was developed and implemented over the same test sites used for the similarity measure and preprocessing analysis. For a meaningful comparison the size of the test sites and the acceptability-unacceptability criteria remained the same. A complete description of this simulation and the results obtained are contained in reference [7].

The results of this evaluation are in close agreement with the previous findings. The overall tabulation showed that 190 out of 192 registration attempts were successful. This 99% success rate is on par with that indicated by the observations described in sections III and IV.

Conclusion

From the experimental evaluation of multitemporal image registration techniques applied to agriculture areas the following conclusion may be drawn. A combination of preprocessing by a gradient operator followed by calculation of the correlation coefficient gives the most reliable registration process. A substantial reduction in computation time with only a slight reduction in registration performance can be obtained by using the sum of the absolute differences as the similarity measure in place of the correlation coefficient.

VI. BIBLIOGRAPHY

1. Anuta, P. E., "Computer-Aided Analysis Techniques for Remote Sensing Data Interpretation," LARS Information Note 100675, LARS, Purdue University, W. Lafayette, IN, 1975.

2. Anuta, P. E., "Guide to Use of the Fast Fourier Transform Algorithm for Two Dimensional Imagery Correlation," LARS Information Note 121069, LARS, Purdue University, December 1969.

3. Anuta, P. E., "Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques," IEEE Trans. Geoscience Electronics, October 1970, p. 353-368.

4. Barnea, D. I., Silverman, H. F., "A Class of Algorithms for Fast Digital Image Registration," IEEE Trans. on Computers, February 1972, p. 179-186.

5. Barnea, D. I., Silverman, H. F., "The Class of Sequential Similarity Detection Algorithms (SSDS's) for Fast Digital Image Registration," IBM Research Paper, May 10, 1971. Copies may be requested from IBM

Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY, 10598.

6. Nack, M. L., "Temporal Registration of Multispectral Digital Satellite Images Using their Edge Images," AAS/AIAA Astrodynamics Specialist Conference, Nassau, Bahamas, July 28-30, 1975, paper # AAS75-104.

7. Svedlow, M., McGillem, C. D., Anuta, P. E., "Experimental Examination of Similarity Measures and Preprocessing Methods Used for Image Registration," LARS Information Note 052175, LARS, Purdue University, W. Lafayette, Indiana.

Table 1. Equations for Correlation Coefficient, Correlation Function, and Absolute Value of Difference Function

A. Correlation Coefficient, $\rho_{\ell k}$:

$$\rho_{\ell k} = \frac{N^2\ \overline{xy}_{\ell k} - \bar{x}\ \bar{y}_{\ell k}}{\{(N^2\ \overline{x^2} - \bar{x}^2)(N^2\ \overline{y^2_{\ell k}} - \bar{y}^2_{\ell k})\}^{\frac{1}{2}}}\ ; \quad \text{correlation coefficient at shift } (\ell,k)$$

where,

$$\overline{xy}_{\ell k} = \sum_{i=1}^{N} \sum_{j=1}^{N} x_{ij}\ y_{i+\ell,\ j+k}$$

$$\bar{x} = \sum_{i=1}^{N} \sum_{j=1}^{N} x_{ij}$$

$$\bar{y}_{\ell k} = \sum_{i=1}^{N} \sum_{j=1}^{N} y_{i+\ell,\ j+k}$$

$$\overline{x^2} = \sum_{i=1}^{N} \sum_{j=1}^{N} x^2_{ij}$$

$$\overline{y^2_{\ell k}} = \sum_{i=1}^{N} \sum_{j=1}^{N} y^2_{i+\ell,\ j+k}$$

B. Correlation Function, $r_{\ell k}$:

$$r_{\ell k} = \overline{xy}_{\ell k} = \sum_{i=1}^{N} \sum_{j=1}^{N} x_{ij}\ y_{i+\ell,\ j+k}; \quad \text{correlation function at shift } (\ell,k)$$

C. Sum of Absolute Values of Differences, $a_{\ell k}$:

$$a_{\ell k} = \sum_{i=1}^{N} \sum_{j=1}^{N} \left| x_{ij} - y_{i+\ell,\ j+k} \right|\ ; \quad \text{sum of absolute values of differences at shift } (\ell,k)$$

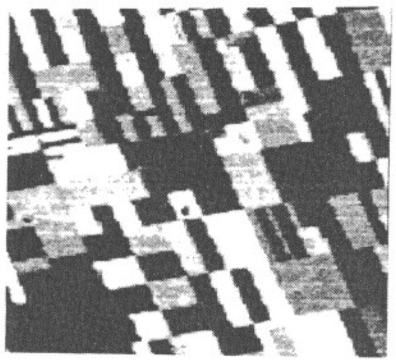
Table 2. Percent (Number) of Acceptable Registration Attempts

Similarity Measure	Total Number of Attempts	Original Imagery	Magnitude of the Gradient	Thresholding at the Median	Thresholding the Magnitude of the Gradient
		90	66	66	30
Correlation Coefficient		90% (81)	100% (66)	65% (43)	90% (27)
Correlation Function		38% (34)	74% (49)	55% (36)	87% (26)
Sum of Absolute Values of Differences		69% (62)	92% (61)	62% (41)	87% (26)

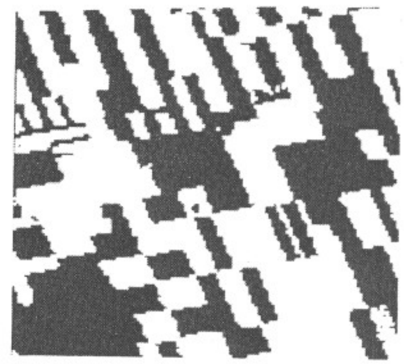
Table 3. Percent (Number) of Acceptable Registration Attempts

	Original Imagery		Magnitude of the Gradient		Thresholding at the Median		Thresholding the Magnitude of the Gradient	
	Acceptable Attempts	Total # Attempts	Acceptable Attempts	Total # Attempts	Acceptable Attempts	Total # Attempts	Acceptable Attempts	Total # Attempts
$\lvert\rho\rvert \geq 0.5$ for Original Imagery	100% (75)	75	100% (75)	75	96% (72)	75	100% (64)	64
$\lvert\rho\rvert < 0.5$ for Original Imagery	65% (37)	57	100% (57)	57	61% (35)	57	89% (25)	28
Overall	85% (112)	132	100% (132)	132	81% (107)	132	97% (89)	92

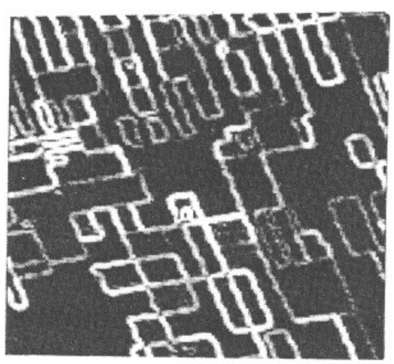
Figure 1. Preprocessed Images



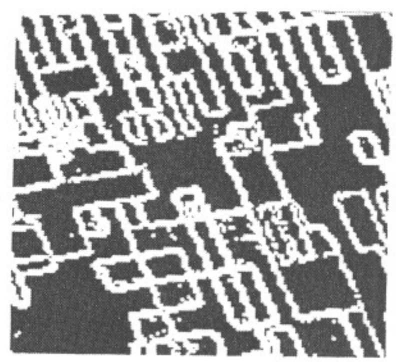
a. Original



c. Thresholded at Median



b. Magnitude of Gradient



d. Thresholded Magnitude
of Gradient