AD-A095 436 RAND CORP SANTA MONICA CA F/G 9/2
HYBRID CORRELATION ALGORITHMS. A BRIDGE BETWEEN FEATURE MATCHIN--ETC(U)
NOV 79 J A RATKOVIC
UNCLASSIFIED RAND/P-6418 NL

1 OF 1
AD A
095436

END
DATE
FILMED
3-81
DTIC

HYBRID CORRELATION ALGORITHMS—A BRIDGE BETWEEN FEATURE MATCHING AND IMAGE CORRELATION

Joseph A. Ratkovic

Nov[redacted]979

P-6418

81 2 10 08

HYBRID CORRELATION ALGORITHMS--A BRIDGE BETWEEN FEATURE MATCHING AND IMAGE CORRELATION

Joseph A. Ratkovic
The Rand Corporation
Santa Monica, CA 90406

ABSTRACT

Up to the present time there have been two basic classes of map matching algorithms--those based on feature matching techniques and those based on image correlation. This paper describes a new class of hybrid correlation algorithms which incorporate features as an integral part of the matching process. These algorithms can be implemented such that it is not necessary to extract features from the sensed image. This paper concludes by showing the domains in which each class of matching algorithm (feature matching, image correlation, and hybrid algorithm) is most appropriate.

INTRODUCTION

The map matching problem has been in search of an "optimal universal" matching algorithm since its inception. Because of difficulty in (1) defining a performance criteria for both accuracy and probability of correct match, and (2) in knowing a priori the distributions associated with all map errors, most researchers have resorted to the use of "ad hoc" algorithms. These have generally been divided into two classes--feature matching and correlation.

The image matching problem, as shown in Fig. 1, is a two-phase problem. In phase 1, the acquisition phase, one is concerned with locating, somewhat grossly, the area in which the match point is centered and avoiding false matches. In phase 2, one is concerned with refining the accuracy with which the match location can be determined. In general, no one algorithm can possibly be suited for solving both the acquisition and accuracy problems, and it is probably necessary to develop algorithms separately for each phase of the problem.

The overall matching problem, shown in Fig. 2, involves four major components: (1) error sources, (2) the scene, (3) preprocessing, and (4) matching algorithms. Before discussing algorithms and describing some algorithm techniques, it is necessary (to provide background for the algorithm discussion which follows) to briefly describe scenes, errors, and preprocessing.

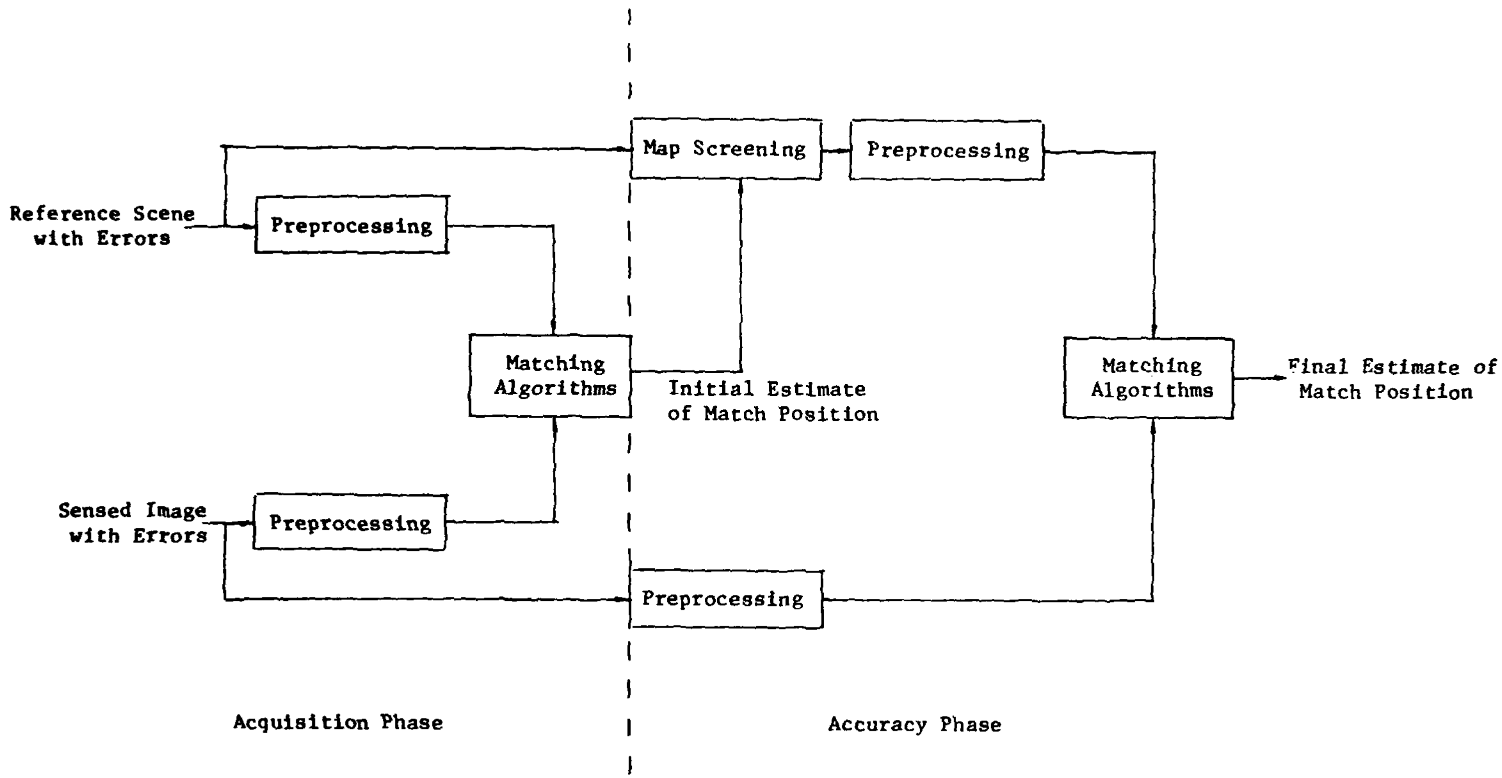

Fig. 1—Acquisition and accuracy phases of map matching process

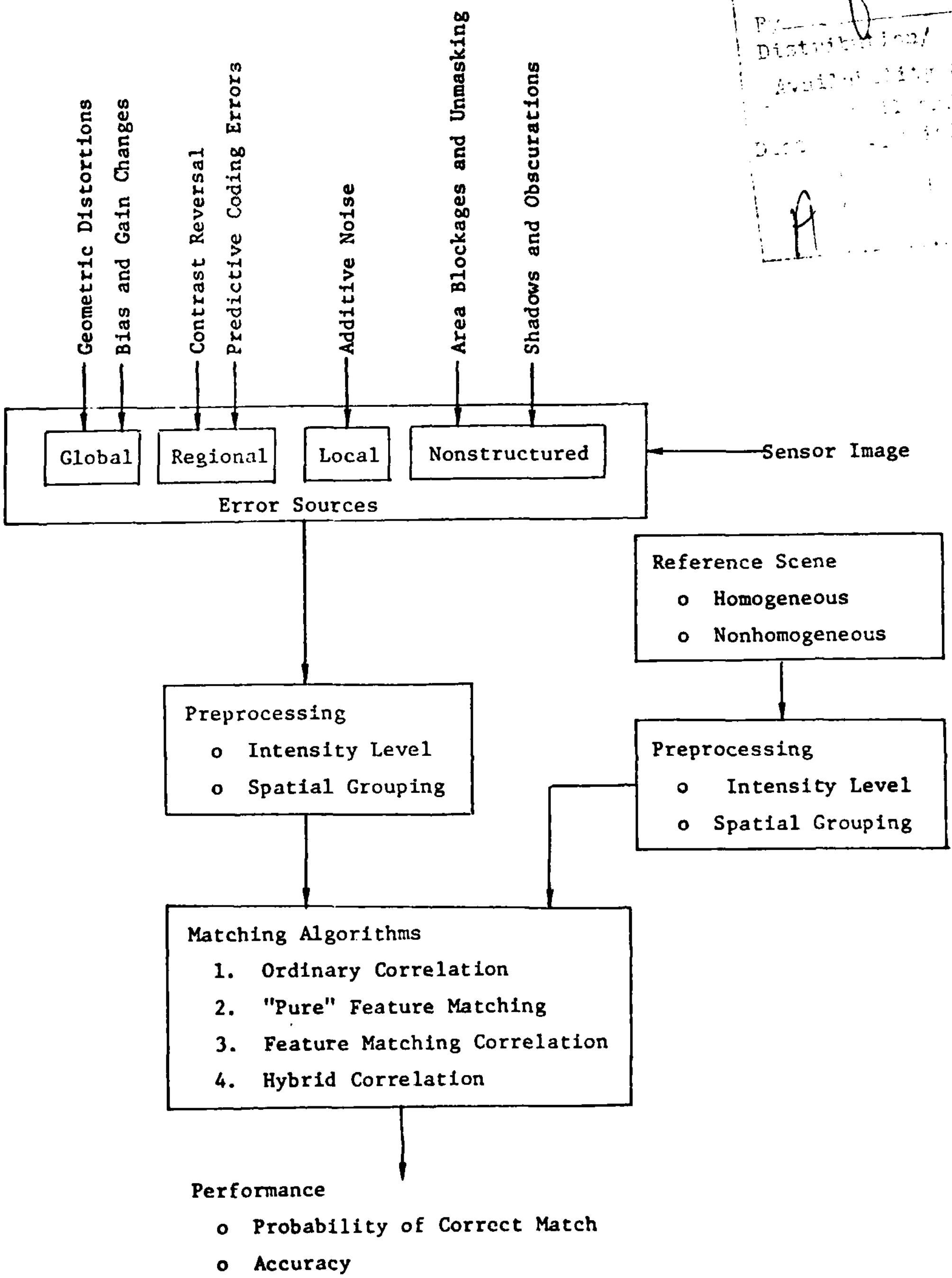

Fig. 2—Generic overview of map matching process

The Scene--Its Composition

The scene is the most complex component of the map matching problem and the most difficult to model. In the discussion that follows we shall examine the question of "scene composition" (relative to both a visual and statistical representation of a scene), and methods for decomposing the scene.

Scenes can be described in the visual domain by the eyeball process as being composed of a set of features. Let us consider as an illustrative example the simple scene shown in Fig. 3. Here, for example, the window feature consists of a set of four panes enclosed by a frame.

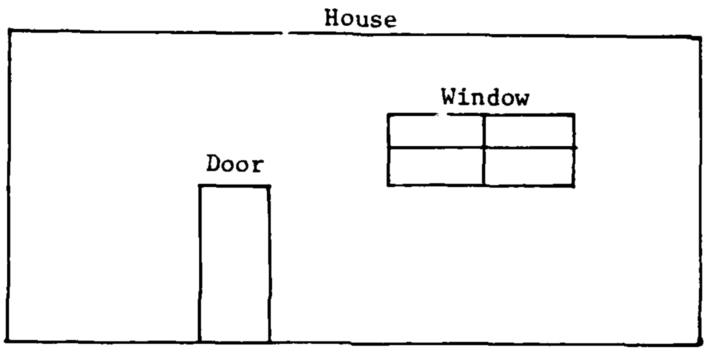

Fig. 3—Example of features consisting of a set of homogeneous regions

In dealing with actual sensor data, picture elements (pixels) are described by a set of intensity values, as indicated in the agricultural scene of Fig. 4. In dealing with intensity values, there are regions in the scene which can be considered analogous to features in the visual domain. These are homogeneous regions within the scene. We shall define a homogeneous region to be a set of spatially connected pixels or elements which possess the statistical property of at least first-order* stationarity and possibly second-order stationarity** and will assume that homogeneous regions are equivalent to features (as a feature can be defined by a single homogeneous region or a set of homogeneous regions).

In Fig. 4 we have identified four homogeneous regions and tagged each pixel (indicated at the bottom portion of the figure) as belonging to one of the four regions. Examining each region we see that the

---

*Mean intensity level constant over the region.

**Mean and variance constant and the autocorrelation independent of position.

REGION 12 USED FOR REFERENCE SCENE

20.	21.	22.	21.	22.	22.	22.	22.	22.	21.	17.	15.	15.	15.	16.	15.	15.	20.	24.	25.
20.	20.	21.	21.	21.	21.	21.	21.	21.	21.	18.	14.	13.	13.	13.	14.	14.	17.	22.	23.
22.	20.	21.	22.	22.	22.	21.	21.	22.	21.	20.	17.	15.	15.	15.	14.	14.	16.	21.	23.
21.	21.	21.	21.	21.	22.	21.	21.	22.	21.	22.	18.	14.	15.	14.	14.	14.	15.	21.	23.
20.	20.	22.	22.	21.	21.	21.	22.	22.	22.	21.	18.	14.	14.	14.	13.	13.	14.	21.	24.
21.	21.	21.	21.	21.	21.	21.	22.	21.	21.	21.	18.	14.	14.	14.	13.	13.	15.	20.	24.
21.	21.	22.	22.	22.	22.	22.	23.	23.	22.	21.	18.	15.	15.	14.	14.	14.	16.	18.	19.
20.	20.	21.	21.	21.	21.	21.	21.	21.	20.	20.	18.	14.	14.	15.	14.	14.	14.	15.	16.
20.	20.	21.	22.	22.	21.	21.	21.	21.	20.	21.	21.	19.	14.	14.	14.	14.	14.	15.	16.
21.	20.	19.	21.	21.	21.	22.	22.	22.	22.	22.	22.	20.	14.	13.	14.	14.	14.	15.	15.
15.	15.	17.	22.	22.	23.	23.	22.	21.	21.	23.	24.	22.	14.	12.	13.	14.	14.	14.	15.
14.	14.	15.	21.	22.	23.	23.	22.	22.	22.	23.	24.	22.	15.	12.	13.	13.	14.	14.	14.
15.	15.	17.	22.	24.	24.	23.	22.	22.	23.	24.	25.	23.	16.	13.	13.	15.	14.	15.	16.
14.	14.	15.	20.	22.	22.	22.	21.	21.	22.	23.	23.	22.	16.	13.	13.	14.	14.	14.	15.
15.	14.	14.	15.	21.	23.	21.	22.	21.	21.	22.	23.	23.	22.	15.	13.	13.	15.	15.	15.
15.	14.	14.	15.	21.	23.	22.	23.	22.	22.	23.	24.	23.	23.	16.	13.	13.	14.	15.	15.
14.	14.	14.	14.	20.	23.	23.	23.	22.	22.	22.	22.	21.	21.	17.	14.	13.	14.	14.	14.
14.	14.	14.	14.	19.	23.	23.	23.	22.	22.	22.	23.	21.	19.	16.	15.	16.	19.	21.	22.
15.	15.	16.	18.	20.	23.	24.	23.	21.	19.	21.	22.	23.	24.	23.	23.	22.	24.	23.	25.
17.	17.	17.	17.	19.	19.	18.	17.	18.	19.	22.	24.	25.	23.	23.	23.	23.	23.	24.	24.

REGION NUMBERS FOR REFERENCE SCENE

2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	4	4	4
2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	4	4
2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	4	4
2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4
2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4
2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4
2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4
2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3
2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	2	2	2
1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Fig. 4—4 km × 4 km agricultural scene

intensity value of a given pixel does not vary significantly from the mean value and that there are distinct boundaries (defined by differences in the mean intensity level) between regions.

Thus far we have shown that scenes are composed of homogeneous regions which may be considered equivalent to features. From a physical standpoint homogeneous regions are areas in which the signature (emissivity for visual and IR, reflectivity for radar, and altitude for terrain contours) is expected to remain fairly uniform, e.g., a grassy field in which all the elements in the region are expected to have the same mean value but this mean value may change as a function of time.

Having established that a scene is composed of homogeneous regions, is there a further subdivision by which we can characterize homogeneous regions? Returning to Fig. 4 we see that there are small variations in the intensity level within a homogeneous region. Some of this variation can be attributed to sensor noise but, neglecting this possibility for the moment, one can consider the variation to be due to some perturbation in the signature of the region. For instance, one can consider the grassy field not to be uniform, but instead to have a few fallen tree trunks and shrubs dispersed within it. If the ground resolution of the sensor is of the same magnitude as the size of the shrubs and tree trunks, then we would expect variations in the intensity level of the grassy region due to these objects, presuming, of course, that the signature of the objects was different from the grass at the wavelength of the sensor. Thus, we can further categorize a homogeneous region in the physical domain by the number of objects which contribute to a signature variation, and in the statistical domain by the number of statistically independent elements* which comprise the region.

The "scene resolution" provides a useful concept in analyzing the statistical variation of a region. We shall define the "scene resolution" as the number of sensor resolution elements or pixels required to make up one independent element in the scene. If there are N pixels within a homogeneous region and N_I independent scene elements ($N_I \leq N$) then the average "scene resolution" for the region would be given by N/N_I. Returning to the grassy field example, if

---

*Statistical independence is different from the property of homogeneity. For instance, one can generate a completely random map from a single distribution which will have the property of homogeneity but will also have all the elements independent. One can imagine a homogeneous region containing a number of independent elements, e.g., a desert area in which the shrub patterns (depending on resolution) constitute the independent elements. It is a difficult procedure to test for and locate independent elements in a scene. Reference 1 describes a short-cut method for estimating this parameter by working backwards from the statistics of the correlation surface and assuming a homogeneous scene with all elements being independent.

the field were completely uniform with no variations in intensity level, then it could be considered to contain only one independent scene element and the scene resolution would be given by the total number of sensor elements in the region, N. In this particular case one could not expect to resolve any features within the region due to the uniformity of the region; thus the scene resolution equals the size of the region (in terms of sensor elements). If, on the other hand, there had been a number of objects (with different signatures) such as tree trunks and shrubs within the grassy region, then we would expect the region to be statistically represented by several independent scene elements. It should be noted also that if the resolution of the sensor were to increase to the point that dimensions of objects within the grassy field were to cover several sensor resolution elements, then these objects would be considered homogeneous regions in themselves. If the resolution were to increase further, then areas within the objects (e.g., moss on the fallen tree trunks) would eventually become homogeneous regions and the process of identifying homogeneous regions could continue ad infinitum.

At this point we see that for a given sensor resolution it is possible to statistically describe a scene as being composed of a set of homogeneous regions with each region being described by a number of statistically independent elements.

Structuring the Errors

There are a number of error sources that affect the performance of the system. It would be desirable to lump these errors into generic categories in discussing system performance rather than treating each error source separately. Such a generic categorization should possess the following properties:

1. The error categories should be mutually exclusive.
2. They should be comprehensive.
3. There should be a positive relationship between the category and a specific preprocessing technique or correlation algorithm to accommodate all errors in that category.

Based on the types of errors that occur in the map matching process and the statistical description of the scene, the following generic categories of errors are proposed:

1. Global Errors--those errors which uniformly affect the intensity level of all scene elements equally. In this category the following errors would generally fit:
   - geometric distortions
   - bias and gain changes

2. Regional Errors--those errors where the change in the intensity levels occurs uniformly only within homogeneous regions or features within the scene. Examples would be:

   - region level shifts (contract reversals)
   - predictive coding errors

3. Local Errors--in this situation the errors are expected to affect each pixel or grouping of pixels (contained within an interpixel correlation length) independently. The primary example of this error source is additive noise.

4. Nonstructured Errors--this is a rather catchall category designed to fit those errors whose effect on the scene cannot be described as being global, regional, or local (e.g., a cloud cover over the target area casts a ground shadow which changes the signature in a nonstructured manner).

Although some errors may sometimes fit into more than one category, this generic categorization will normally accommodate all error sources as well as provide a convenient means of establishing guidelines for algorithms and preprocessing selection.

Preprocessing

The preprocessing of sensor imagery consists of either changing the intensity levels through the image or segmenting the scene spatially into groups of pixels. The intensity level preprocessing is designed to compensate for any biases or gain changes in the system; whereas, spatially grouping of elements is designed to accommodate geometric errors.

In general, preprocessing is designed to accommodate global errors that occur in the scene and which, by definition, effect all scene elements equally. Thus global errors such as gain changes and bias errors are handled by normalizing the intensity level and by zero meaning the data, respectively. As discussed previously, geometric errors also are global in nature and reduce the degree of congruence between sensed image and reference image. In order to reduce the effect on system performance, geometric errors always force one to work with smaller map sizes and, depending on the nature of the distortion (in azimuth and elevation), may also force one to shape the window of the sensed image or to search for a rotation or scale error. Thus, to accommodate this type of error, it is necessary at a minimum to spatially group the sensor map elements into a single (or number of) smaller map(s). If distortions are uneven in azimuth and elevation it will also be necessary to spatially group the elements so that the appropriate window shape may be obtained.

MATCHING ALGORITHMS

The matching algorithm is only one part of the overall matching process, as indicated in Fig. 5. To begin with, there are a number of system parameters which can be chosen to lessen or worsen the severity of the errors on system performance. These include the sensor orientation, resolution and wavelength, the reference map preparation, and the flight geometry of the vehicle. There are, as indicated in the figure, separate processes for accommodating each of the error sources. Global errors (e.g., geometric distortions, gain changes, etc.) are accommodated in the preprocessing by either reducing and shaping the map size or by normalizing the intensity level of the sensed image. They can also be accommodated by searching in the matching algorithm for rotation and/or scale factor errors. The scene composition problem involves checking to insure that the reference map contains a sufficient amount of independent information and that there are no "scene redundancy" problems within the reference map boundaries.

The algorithm itself is primarily designed to accommodate regional and local errors with nonstructured errors being more difficult to foresee and accommodate. The basic matching algorithm for accommodating regional and local errors can be categorized as belonging to a feature matching or image correlation class of algorithms. It should be noted that none of these algorithms have been mathematically derived to maximize system performance (probability of correct match or accuracy) and, therefore, must be considered in a sense to be "ad hoc."

It is first necessary for the "feature matching" procedure to extract the features from the scene. The first part of the feature extraction process involves locating the edges or boundaries of features. Thus, the scene can be reduced to a set of lines which are the boundaries of the feature. Next the line intersection points are located. In general, the number of lines emanating from each vertex is retained and used as part of the weighting criteria in the feature matching algorithms.

In image correlation there are two basic types of algorithms utilized--those which emphasize the degree of similarity between scenes such as the product, and those which emphasize differences between scenes such as the difference squared and MAD* algorithm.

The standard correlation process works on the gross characteristics of the scene and all preprocessing is done globally (i.e., the mean level when subtracted out is zero-meaned over the entire scene, and similarly when the scene is normalized by the variance, this is

---

*Mean absolute difference.

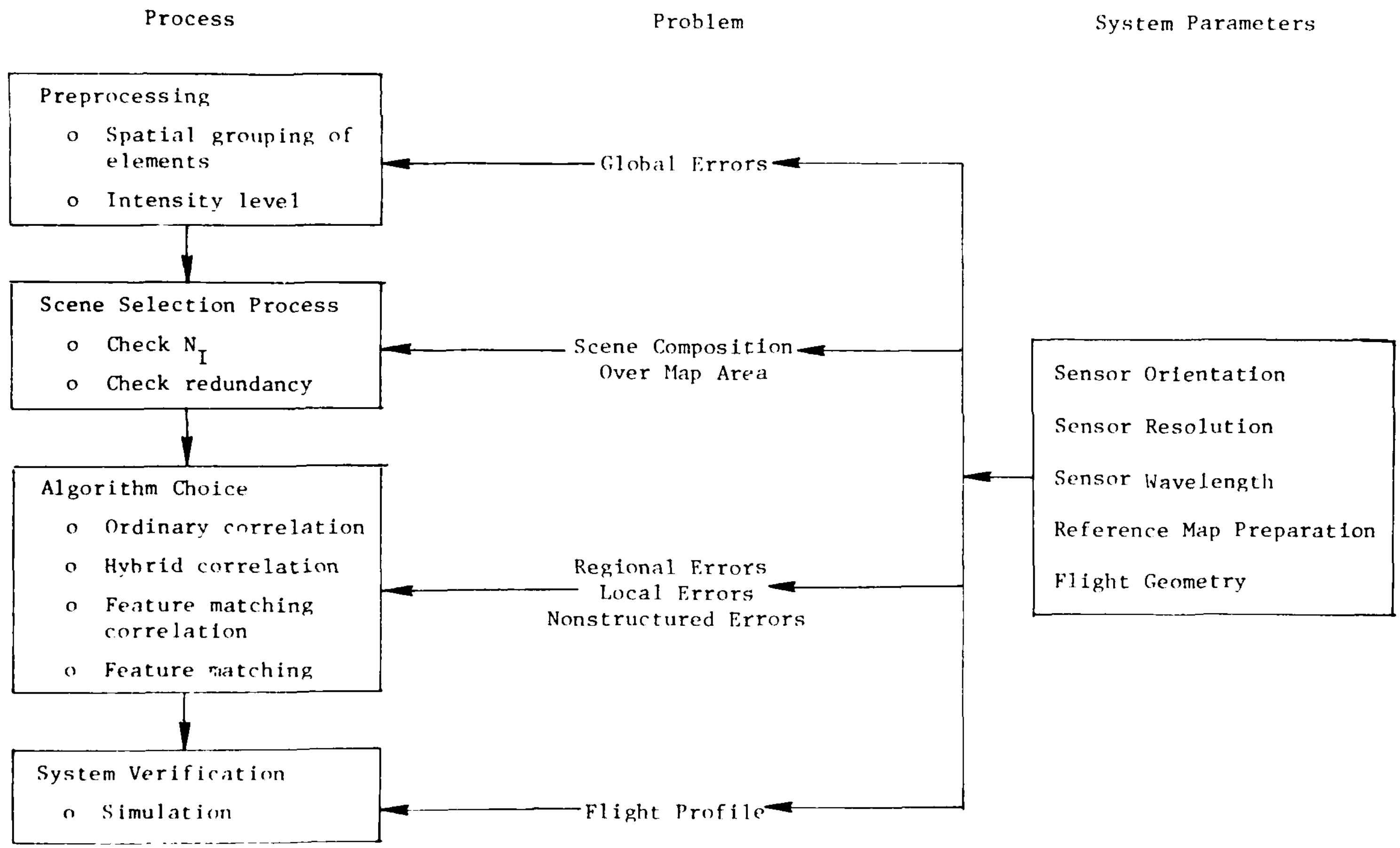

Fig. 5 - Acquisition system design

done over the entire scene). In a sense the usual correlation process is designed to work on a homogeneous scene. There are two basic variations to the standard or usual correlation algorithm which are more specifically tailored to nonhomogeneous scenes and the errors associated with them. It should be noted that these variations, in the absence of nonhomogeneity in the scene, reduce to the usual correlation process. We shall denote these variations that deal with scene nonhomogeneities as (1) feature matching, and (2) hybrid algorithms.

One could introduce a feature matching algorithm into the correlation process by breaking up separately the sensor and reference maps into homogeneous subareas. Each of these maps would then consist of a set of homogeneous regions and all processing (rather than being on a global scale) would be performed separately on each homogeneous subregion. Thus, when maps are zero-meaned and normalized, the local mean and variance in each subregion is computed and used to perform the normalization.

After processing both the reference and sensor map on the basis of homogeneous regions, a standard correlation algorithm can be used to determine the position of match between the two maps. The major generic difference between this feature matching correlation algorithm and the "pure" feature matching algorithm (employing pattern recognition techniques) is the weighting given to homogeneous regions. In "pure" pattern recognition algorithms, edges are first extracted and used to identify line intersection points. These line intersection points or vertices then form the primary basis for matching two scenes. In a sense (since edges can be considered the boundaries of homogeneous regions, and vertices are formed by the intersection of edges) a pure feature, or pattern matching algorithm weight all homogeneous regions equally, whereas in the feature matching correlation algorithm, each homogeneous region would receive a weighting proportional to its size (measured in terms of the number of independent elements contained within). In summary then "pure feature matching algorithms can be viewed as being different from feature matching correlation in that different weights are assigned to the various homogeneous regions.

There is another adaptation of the standard correlation algorithm what has been developed at Rand which one can implement to accommodate homogeneous regions. We shall refer to this as a hybrid algorithm which processes only the reference scene into homogeneous regions. The principal idea here is that every position of comparison between the two images is assumed to be the correct one. Thus at each displacement position or comparison point the sensor scene is segmented identically as its counterpart reference map.* At the position at which the two maps correctly match the sensor scene will then be segmented almost perfectly, enhancing the match, and at all other positions the sensor map segmentation will essentially look like noise.

---

*For each displacement position the matching process consists of correlating each homogeneous region of the reference map and segmented sensor image separately, and combining additively the correlation in each individual region.

The objective of this correlation method is to avoid the errors associated with extracting homogeneous regions or features from the sensor image, and the additional processing requirements placed on the system. If the image is noisy, normal edge operators have difficulty in performing their feature extraction task and, as a compromise, the hybrid approach, which strictly is not as good as a "pure" feature matching or correlation feature matching algorithm, does possess significant advantages over the standard correlation approach at accommodating certain types of regional errors such as contrast reversals.

In Fig. 6 we show an example of this hybrid processing scheme. We have in the figure identified each reference pixel with a homogeneous region. Thus each reference pixel has both a region identification and an intensity associated with it. The template for the sensor map processing is shown for two map displacement positions. As indicated in the figure, the sensor map is segmented into homogeneous regions at each of these displacement positions in a manner identical to that of the reference map elements occupying the same spatial position. The sensor map elements are then processed by homogeneous regions (i.e., the mean intensity level subtracted out and possibly normalized by the intensity variation in the region) with the total correlation between sensed images and reference map being the sum of the correlation in each region at each displacement position. Thus we have identified four generic types of image matching methods:

1. Standard correlation algorithm
2. "Pure" feature matching algorithm
3. Feature matching correlation algorithm
4. Hybrid algorithm

The first two methods are the two basic approaches to image matching, while the latter two methods are variations of the standard correlation process designed specifically to accommodate nonhomogeneous scenes and the nonglobal errors associated with them.

SIMULATION RESULTS

Let us examine the effects of regional and local errors on the performance of matching systems for various classes of algorithms. First, let us examine the accuracy of the system measured in terms of the sharpness of the correlation peak. The general broadening of the correlation peak around the match point is caused primarily by the nonhomogeneous nature of the scene. Thus if we could process out the nonhomogeneous regions in the scene by a feature matching or hybrid algorithm we could expect a general sharpening of the correlation peak around the match point.

To illustrate these points we will decompose several Earth Resource Satellite (ERTS) maps into homogeneous regions and perform an autocorrelation between a sensor and reference map using the standard

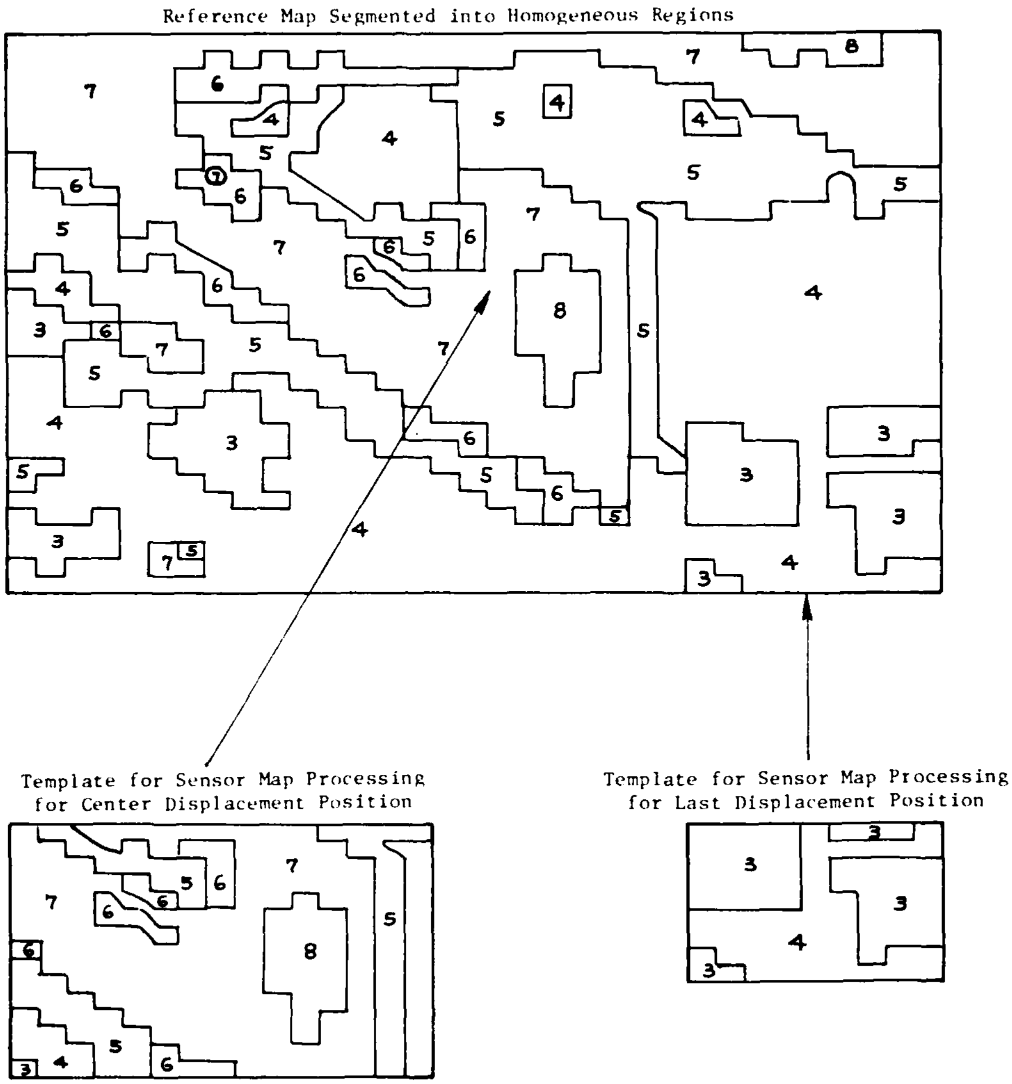

Fig. 6—Illustration of hybrid matching process

product algorithm, a feature matching algorithm, and a hybrid correlation matching algorithm which have been described previously. The feature matching algorithm essentially removes the effect of homogeneous regions since all homogeneous regions are zero meaned and normalized separately. The hybrid algorithm, on the other hand, takes out some but not all of the effects of the scene nonhomogeneity. Figure 7 shows the effect of using these three different algorithms upon the correlation surface for four different ERTS scenes. The normal autocorrelation process produces a spread out correlation peak, while the feature matching algorithm (homogenizing both the reference and sensor scene) produces the sharpest correlation peak, being limited

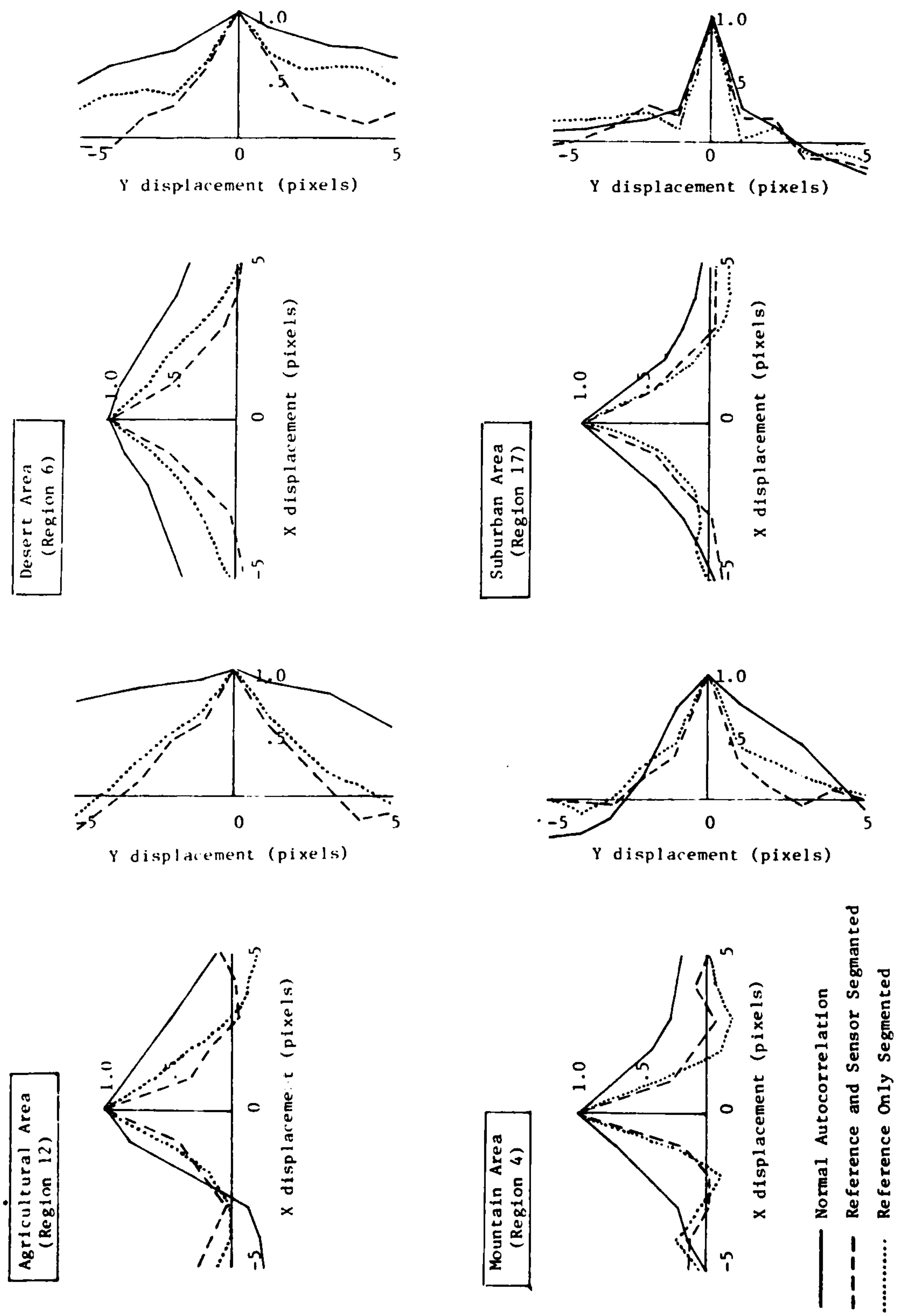

Fig. 7—Effects of homogeneous segmentation

only by the interpixel correlation. The hybrid algorithm produces a correlation surface between the two indicating that it does remove some but not all of the effects of scene nonhomogeneity. The remaining width of the correlation surface is due to interpixel correlations between nonindependent map elements contained within the homogeneous regions. To summarize, the slope of the correlation surface is dominated by the size and shape of the homogeneous regions composing the scene. Thus by utilizing feature matching or hybrid algorithms it is possible to filter out these low spatial frequency components and sharpen the correlation peak. The interpixel correlation and intensity variations between pixels, represented by the number and size of independent elements within the region, are only significant to the correlation process for completely homogeneous scenes (which are rare) and for scenes which have been homogeneously processed. Conversely, by homogeneously segmenting the scene, sharper correlation peaks can be produced whose widths are limited only by the interpixel correlation or the size of the independent elements.

The choice of matching algorithms for acquisition (P_c being major performance measure) will depend on the nature and magnitude of the regional and local errors. Some analysis has been performed in relating nonstructured errors to changes in system performance. In general the algorithm choice is not strongly dependent on the nature of nonstructured errors. Nonstructured errors are best accommodated in the mission planning phase of the operation. By proper route planning obscuration and masking errors may be avoided, and by timing and weather planning it may also be possible to reduce the diurnal and weather effects which can cause nonstructured errors. Thus the occurrence of nonstructured errors can be reduced by careful mission planning. Generally any residual nonstructured errors cannot be adequately modeled and thus one can only hope that they do not seriously degrade system performance.

The algorithm choice, then, in the extreme case of local errors only, tends toward ordinary correlation, whereas, in the other extreme (regional errors only) the algorithm tends toward pure feature matching. As one is generally never confronted by an either-or-situation, except in the case of Terrain Contour Mappping (where there are primarily local errors), it is necessary to weigh the relative magnitude of local and regional errors present in deciding upon the choice of algorithm.

Let us first consider the differences between the various categories of correlation algorithms when only local errors (additive noise) are present. To examine the effect, we took several 10 × 10 element sensor maps from the center of 20 × 20 reference scenes in various parts of an ERTS map. To these sensor scenes we added white Gaussian destributed noise such that the S/N ratio was 0.5. The simulation consisted of creating 25 different noisy sensor images and matching the reference and sensed imagery for different categories of algorithms (feature matching correlation, hybrid, and ordinary correlation algorithms) using the product algorithm. Table 1 shows the percent of successful matches (P_{SIM}) for each category of algorithm. The feature matching algorithm scored perfectly each time and is not shown in the table.

Table 1

MONTE CARLO SIMULATION RESULTS
Reference Map: 20 × 20
Sensor Map: 10 × 10

Terrain Type	Region	Type of Algorithm	Simulation Results (Product)
Mountain	2	Ordinary Correlation	0.96
Mountain	2	Hybrid	0.68
Suburbs	17	Ordinary Correlation	1.00
Suburbs	17	Hybrid	0.80
Desert	10	Ordinary Correlation	1.00
Desert	10	Hybrid	1.00
Desert	6	Ordinary Correlation	0.96
Desert	6	Hybrid	0.68
Agricultural	12	Ordinary Correlation	0.76
Agricultural	12	Hybrid	0.36

The homogeneous regions within the reference map boundary were defined manually. The homogeneous regions or features in the sensor image were also defined manually for the feature matching correlation algorithm. In the real world these regions must be extracted automatically so that the results for the feature matching correlation algorithm are, in a sense, an optimum case. In the real world, homogeneous regions are generally extracted through the use of edge operators. These systems generally do not perform well in the presence of local errors. Simulation results achieved for real-world scenes using pure feature matching approaches generally indicate results closer to or worse than those achieved by the hybrid algorithm are obtainable when automated edge finding feature extraction techniques are used.

To determine the change in system performance measured in terms of probability of correct match (P_c) due to regional errors interacting with the three different categories and types of algorithms described previously, we ran an experiment to test the effects of such errors. In an attempt to place regional errors into the correlation process we decided to see the effect of changing the mean values of the "intensity" levels in the homogeneous regions of the scene. For this experiment a sensor map (20 × 20) was chosen with a larger number of homogeneous regions (mountain area, region 4) and the mean level of each homogeneous region was changed by a random amount. The magnitude of the level change was drawn from a zero mean Gaussian distribution with three different standard deviations chosen to be 25, 50, and 100 percent of the dynamic range of intensity values in the scene. Two different algorithms (the normalized product and the difference-squared with the mean intensity value subtracted out) and three different processing schemes (both sensor and reference maps homogeneously segmented, only the reference map segmented (hybrid) and no segmentation) were utilized. Additionally a small amount of noise was added to each pixel in the scene. The results are shown in Table 2. Shown in this table are the percent of successful correlations (out of 25), P_{SIM}, for each run using the different algorithm categories and types. Since we are using the "perfect" feature matching correlation algorithm we would not expect any change in performance with change in level and the results so indicate. On the other hand, there is a definite degradation in P_{SIM} for the ordinary correlation cases for all types of algorithm, with increasing changes among homogeneous levels in the scene. The hybrid algorithms, while generally having performance somewhat below that of the "perfect" feature matching algorithms, essentially do not degrade with increasing regional error.

Table 2

SIMULATION RESULTS WITH LEVEL CHANGES BETWEEN HOMOGENEOUS REGIONS
Mountain Area--Region 4
(20 × 20 Sensor Map, 40 × 40 Reference Map)

Process	Algorithm	Magnitude of Level Change: 25 Percent P_{SIM}	50 Percent P_{SIM}	100 Percent P_{SIM}
Ordinary Correlation	Normalized Product	0.92	0.88	0.52
Hybrid	Normalized Product	0.72	0.72	0.68
Perfect Feature Matching	Normalized Product	1.0	1.0	1.0
Ordinary Correlation	Difference Squared (zero-meaned)	0.88	0.68	0.48
Hybrid	Difference Squared (zero-meaned)	0.96[a]	1.0	1.0
Perfect Feature Matching	Difference Squared (zero-meaned)	1.0	1.0	1.0

[a]The lower value relative to higher magnitude level changes is attributed to a statistical variation in only using 25 samples.

SUMMARY AND CONCLUSIONS

This paper described the image matching process as a two-phase process, with the first phase being concerned with the acquisition of the correct match area, and the second stage being concerned with accurately locating the match point. The major rationale for the failure of the system to acquire is described as being due to a combination of noise plus interscene redundancy (e.g., checkerboard), this latter problem being extremely difficult to model. Accuracy was shown to depend on two components of the scene structure--the size and magnitude of homogeneous regions in the scene and the interpixel correlation (expressed in terms of an independent scene element)--and the amount of geometric distortion present.

It has been shown that accuracy can be improved by utilizing a hybrid or feature matching algorithm which segments the scene into homogeneous regions. This segmentation significantly sharpens the correlation. The residual spread in the correlation peak can be attributed to interpixel correlation.

The acquisition problem, described in Fig. 5, consists of determining the preprocessing requirements, developing a scene selection criteria, choosing an algorithm, and verifying the system via a simulation. As indicated in this figure, the first problem that must be accommodated is global errors. These errors are generally accommodated by either normalizing the intensity level or by spatially grouping the scene elements so as to reduce the susceptibility of the matching process to geometric distortion.

The scene selection process requires that two criteria be met. The first is that a sufficient amount of independent information must be contained in the map. Although not discussed, a number of methods have been proposed to measure the independent information contained within the scene. The correlation length appears to be a poor measure because of the ambiguity associated with the term. The number of "independent scene elements" appears to be a good measure to utilize for correlation processes, while the "number of vertices" appears appropriate for pure feature matching processes. The second scene selection process of importance is the avoidance of interscene redundancy (e.g., checkerboard patterns). The height of secondary correlation peaks using ordinary correlation does not appear to be as good a measure of scene redundancy as the height of secondary peaks using the hybrid algorithm. This hybrid class of algorithm assumes that at each displacement position the sensor image is segmented into homogeneous regions in an identical manner to the portion of the reference map against which it is being compared. Thus, this class of algorithm emphasizes the spatial structure of the scene and the few simulation results acquired to date indicate that secondary peaks on the autocorrelation surface associated with the hybrid algorithm are places where false matches are likely to occur due to an interscene redundancy.

Finally, in the acquisition process, an algorithm must be chosen from the generic class of ordinary correlation, hybrid correlation, feature matching correlation, and feature matching such that it can accommodate the amount of regional, local, and nonstructured errors that are anticipated. If only local errors are anticipated (e.g., TERCOM navigation system) then ordinary correlation algorithms are appropriate, whereas, if regional errors dominate, a feature matching or hybrid algorithm is demanded. Most real-world scenes have both regional and local errors superimposed. If the magnitude of the variation in the mean intensity levels between homogeneous regions in the area (that can be accounted for in the signature prediction) exceeds in value 50 percent of the intensity level difference between regions, then it appears that one is forced to use a feature matching algorithm, with the hybrid algorithm looking as an attractive alternative to avoid the near real-time feature extraction process in the sensed image, while at the same time being able to deal with regional errors.

REFERENCES

1. Ratkovic, J. A., et al., *Estimation Techniques and Other Work in Image Correlation*, The Rand Corporation, R-2211-AF, September 1977.

2. Gupta, J. N., and P. A. Wintz, *Closed Boundary Finding Feature Selection and Classification Approach to Multi-Image Modeling*, Laboratory for Applications of Remote Sensing, Note No. 062773, Purdue University, W. Lafayette, Indiana, 1973.

3. Farag, R. F. H., "An Information Theoretic Approach to Image Partitioning," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-8, No. 11, November 1978.

4. Kettig, R. L., and D. A. Landgrebe, *Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects*, Laboratory for Applications of Remote Sensing, Purdue University, W. Lafayette, Indiana, 1975.

5. Robertson, T. F., et al., *Multispectral Image Partitioning*, Laboratory for Applications of Remote Sensing, Purdue University, W. Lafayette, Indiana, 1973.

6. Davis, L. S., "A Survey of Edge Detection Techniques," *Computer Graphics and Image Processing*, Vol. 4, pp. 248-270, September 1975.

7. Brice, C. R., and C. L. Fennema, "Scene Analysis Using Regions," *Artificial Intelligence*, Vol. 1, pp. 205-226, 1970.

8. Garnett, J. M., III, and S. S. Yaw, "Nonparametric Estimation of the Bayes Error of Feature Extractors Using Ordered Nearest Neighbor Sets," *IEEE Transactions, Computers*, Vol. C-26, pp. 46-54, January 1977.

9. Pratt, W. R., *Digital Image Processing*, Wiley & Sons, 1978.

10. Rosenfeld, A., and A. C. Kak, *Digital Picture Processing*, Academic Press, 1978.

11. Wacker, A. G., *A Cluster Approach to Finding Spatial Boundaries in Multispectral Imagery*, Laboratory for Agricultural Remote Sensing, Purdue University, W. Lafayette, Indiana, 1970.

12. Bailey, H. H., et al., *Image Correlation, Part I, Simulation and Analysis*, The Rand Corporation, R-2057/1-PR, November 1976.

13. Lahart, M. J., "Local Segmentation of Noisy Images," *Optical Engineering*, Vol. 18, No. 1, pp. 76-78, January-February 1979.

14. Merchant, J., *Address Modification, Image Technology Program*, Vol. I (Overview and Theory), Vol. II (Experimental Results), Honeywell Electro-Optics Center, Lexington, MA, September 1978.

15. Gerson, G. et al., *Image Sensor Measurements Program: Volume 1 Multiple Subarea Bi-Level Correlation Scene Matching System*, Contract F-30602-77-C-0049, Hughes Research Laboratories, Malibu, California, June 1979.

16. Hall, E., and R. C. Gonzalez, *Scene Content Analysis, Measurement, Refinement, and Verification*, Contract F4701-77-C-0072, DARPA/SAMSO, University of Tennessee, December 1978.

17. Kin, R. L., *Pattern Matcher Development Study*, DARPA Contract DAAK40-77-C-0017, Rockwell International, Missile Systems Division, Anaheim, California, June 1978.

18. Ratkovic, J. A., *Structuring the Components of the Image Matching Problem*, The Rand Corporation, N-1216-AF (to be published).

19. Ratkovic, J. A., *Performance Considerations for Image Matching Systems*, The Rand Corporation, N-1217-AF (to be published).

DATE
FILMED

8