A Spatial Clustering Procedure for Multi-Image Data

ROBERT M. HARALICK, MEMBER, IEEE, AND ITS'HAK DINSTEIN, MEMBER, IEEE

Abstract—A spatial clustering procedure applicable to multi-spectral image data is discussed. The procedure takes into account the spatial distribution of the measurements as well as their distribution in measurement space. The procedure calls for the generation and then thresholding of the gradient image, cleaning the thresholded image, labeling the connected regions in the cleaned image, and clustering the labeled regions. An experiment was carried out on ERTS data in order to study the effect of the selection of the gradient image, the threshold, and the cleaning process. Three gradients, three gradient thresholds, and two cleaning parameters yielded 18 gradient–thresholds combinations. The combination that yielded connected homogeneous regions with the smallest variance was Robert's gradient with distance 2, thresholded by its running mean, and a cleaning process that considered a resolution cell to be homogeneous if and only if at least 7 of its nearest neighbors were homogeneous.

I. Introduction

CLUSTER analysis has become in the last 30 years a multi-disciplinary technique of data analysis. Different methods of cluster analysis have been developed and applied in different areas of research. We will mention here only a few examples of the application of clustering in various disciplines (the order is of no significance). A clustering algorithm was used (Gose, [6]) as part of a procedure to *identify breast cancer* using radiographs and xerograms. A program that simulates the taxonomic process for plant classification (Rogers and Tanimoto, [19]) was applied to 300 herbarium specimens of manihot esculenta. Clustering techniques were applied to a collection of 1400 aeronautical documents in an *information retrieval* experiment (Sparck, [25]). Sneath proposed an analysis intended to produce taxonomic *groups of bacteria*. ISODATA clustering procedure (Wolf, [29]) has been used in a system for determining *cloud motions*. Cluster formation was used to diagnose 199 subjects in a *psychiatric* institute (Kaskey, [12]). A hierarchical grouping technique was applied (Ward, [27]) to 25 *test profiles* based on art preference. *Behavioral problems* of deaf children (Haralick, [9]) were analyzed using clustering of variables. Different clustering techniques have been used to produce dendrograms for biological data (Sokal and Sneath, [24]). Dendrographs have also been used to represent mutual relationship among geologic variables (McCammon, [13]). Clustering was used in image processing of remote sensing data (Haralick, [8], Smedes *et al.*, [22]). This list does not intend to cover all the areas in which cluster analysis has been applied. It only illustrates the broad possibilities and the multidisciplinary nature of clustering.

Clustering of any set of data is a subjective process that can be done in a number of different ways depending on the purpose of the classifier (Gilmour, [5]). For example, a zoologist might place a whale and a monkey in the same class whereas a fisherman will prefer the whale and tuna in the same class (Watanabe, [20]). There are some different objectives of cluster analysis. One objective is "...to gain more information about the structure of the data set" (Nagy, [14]). When dealing with large masses of data, cluster analysis might compress the data so that it can be analyzed more easily (Bonner, [2]). Such generalization might result in some loss of information but it may emphasize some other interesting parts of the information as well as increase the efficiency with which large masses of data can be processed (Ward, [27]). Cluster analysis may "inform the researcher where the 'action' in his data set lies" (Haralick, [8]). It is a way of analyzing the details of the data's structure (Ball, [1]).

Many clustering algorithms deal with measurements in ways that do not consider the order by which the measurements were taken. Indeed, for many problems, the order by which the measurements were taken or the spatial distribution of the units under consideration are irrelevant to the clustering process. In some cases, the relation between the order by which the data was collected and the clusters of the measurements depends very much upon the way the data was gathered. Such a case is in the clustering of image data. The resolution of the image data, when properly selected, should yield many measurements for each object of interest. Therefore, resolution cells of the same neighborhood are likely to belong to the same object, except at boundaries. We will refer to clustering procedures that take into account spatial relations between elements as spatial clustering procedures.

Although image data is a good candidate for spatial clustering procedures, much of the clustering which has been done with multi-image data has not been spatial. This certainly is true for the measurement space iterative clustering techniques used on image data and described by Haralick [7], Darling and Juris (1970), Haralick and Dinstein [8], as well as the ISODATA or K-means clustering techniques used by Wacker and Landgrebe [26] and Smedes *et al.* [22] and many others in the remote sensing area.

The artificial intelligence community has been an active user of spatial information from scene data. Much work

Manuscript received April 18, 1974; revised August 20, 1974.
R. M. Haralick is with the Department of Electrical Engineering and University of Kansas Remote Sensing Laboratory, Lawrence, Kans. 66045.
I. Dinstein was with the Department of Electrical Engineering and University of Kansas Remote Sensing Laboratory, Lawrence, Kans., as a Ph.D. candidate. He is now with the Communications Satellite Corporation, Clarksburg, Md.

has gone into the definition of homogeneous regions and edge detection. Brice and Fennema [3] describe a procedure for partitioning an image into a large set of primitive regions each of which is typically a connected component having the same grey tone. Then a merging algorithm is applied to group together those regions of similar tone. Gradient and derivative computation algorithms have been used to find edges (Roberts, [17]; Prewitt, [16]; Hueckel, [10]; Rosenfeld and Thurston, [21]). In this paper, we discuss extensions of these kinds of techniques to multi-image data sets. Such data sets occur naturally with multispectral scanners where the same scene is viewed in many wavelengths from the infrared through the ultraviolet as well as with aerial photography taken over the same area at different times.

II. On Some Spatial Clustering Procedures for Multi-Image Data

Let $Z_r = \{1,2,\cdots,N_r\}$ be a row index set and let $Z_c = \{1,2,\cdots,N_c\}$ be a column index set. We call the set of ordered pairs $Z_r \times Z_c$ the set of *resolution cells*, or a *spatial domain*. Let $G_i = \{1,2,\cdots,N_i\}$ be a set of N_i grey levels, and let $G = G_1 \times G_2 \times \cdots \times G_K$ be the Cartesian product of the sets of grey levels $G_1,G_2,\cdots,G_K$. A *digital multi-image* is a function $I: Z_r \times Z_c \rightarrow G$. The function I assigns a K-tuple of grey levels to every resolution cell in $Z_r \times Z_c$. Multispectral scanner data and multi-data aerial image data take such a form. Using this notation, we now review three clustering procedures that take into account the spatial distribution of the grey level K-tuples.

Haralick [7] introduced the following spatial clustering algorithm. Let $S = \langle S_n \rangle_n^Q = 1$ be a set of spatially connected subsequences of I. On each subsequence S_i, define the empirically observed probability $P_i(g)$ as the portion of resolution cells in S_i which have K-tuple g. Define a function $W: G \rightarrow (0,1)$ by $W_g = \max_i P_i(g)$, $i = 1,\cdots,Q$. $W(g)$ is the highest relative frequency which the K-tuple g has in the subsequences S_i, $i = 1,2,\cdots,Q$. The range of the function W consists of N_g numbers, one for each K-tuple of G. Define the sequence $B = \langle g_i \mid g_i \in G \rangle$ to be constructed such that $W(g_i) \geq W(g_j)$ for $i \leq j$. The elements of the sequence B are ordered according to the respective descending order of their images through the function W. The first element in the sequence B is that K-tuple which has a maximum observed probability $P_i(g_i)$ over all $i = 1,2,\cdots,Q$ and all $g_i \in G$. The resolution cells containing that K-tuple (the first element in B) are considered to be the most important center set. The set of resolution cells in I containing the K-tuple which is the second element in B is considered to be the second most important center set, and so on. Once the collection of center sets is found, clusters are "grown" around those centers. Two parameters govern the clustering process. One is the maximum number of expected clusters and the second is a probability cutoff parameter.

A spatial approach to imagery clustering was described by Nagy [15]. Nagy's clustering procedure is based on a simple "chain algorithm" proposed by Bonner [2]. In the first step, points are assembled into row strips. This is based upon the assumption that "spatially adjacent vectors tend to belong to the same type of ground cover except at field boundaries." Examining the resolution cells along the scan lines, similar resolution cells were assigned to strips. Each strip was terminated only when the addition of the next resolution cell would have increased the internal scatter of the strip above a given threshold. At that point, the formed strip was assigned to a cluster (or designated to start a new cluster), and the formation of a new strip began. The assignment of a strip to a cluster was done by comparing the strip to the cluster centers. The search for a cluster was done in a decreasing order of cluster populations, to save computation time and to eliminate a small group of abnormal components.

The following is a formalization of this clustering procedure. Let I, $Z_r \times Z_c$, and G be defined as before. Define a function $SS: Z_r \times Z_c \rightarrow S$ which partitions the spatial domain $Z_r \times Z_c$ into a set S of spatially connected subsets, $S = \{S_1,S_2,\cdots,S_Q\}$. The subsets S_i, $1 \leq i \leq Q$ are the "strips" that were mentioned previously. The spatial connectivity of those strips is explained in the definition of the function SS, which is given here in a sequential manner.

1) $SS(1,1) = S_1$.

2) $SS(m,n) = S_i$ if and only if $SS(m, n-1) = S_i$ for $n \neq 1$ or, $SS(m-1, N_c) = S_i$ for $n = 1$, and

$$\sum_{(k,l) \in S_i \cup \{(m,n)\}} [I(k,l) - I(S_i \cup \{m,n\})]^2 \leq T_1$$

where N_c is the last column in $Z_r \times Z_c$, $I(A)$ is the mean of the grey levels of the resolution cells belonging to A, and T_1 is a specified threshold.

Let $C = \{C_1,C_2,\cdots,C_k\}$ be a set of cluster codes. Define a function $CC: S \rightarrow C$ that assigns strips to clusters in the following manner:

a) $CC(S_1) = C_1$;

b) $CC(S_2) = \begin{cases} C_1, & \text{if } d[I(S_2),I(S_1)] \leq T_2 \\ C_2, & \text{otherwise; and} \end{cases}$

c) $CC(S_j) = \begin{cases} C_i, & \text{if } d[I(S_i),I(C_i)] \leq T_2 \text{ and } C_i \text{ is the largest cluster for which it holds} \\ C_k, & \text{if the above condition does not hold and } C_k \text{ is empty.} \end{cases}$

This clustering procedure was applied to multispectral data collected during the Imperial Valley study conducted by NASA in 1969. Flight lines of 5000 feet and 10 000 feet by the University of Michigan aircraft were selected. This simple and economical clustering procedure was demonstrated to be a feasible alternative to the conventional terrain classification methods.

Jayroe [11] introduced a three-step spatial clustering procedure for multi-images. In the first step, a boundary map is prepared by thresholding of gradient images. The two gradient images used are obtained by computing the Euclidian distance between nearest neighbors in the horizontal and vertical directions. In the second step, clusters are formed by scanning the boundary map with a fixed-size square of resolution cells. When the square hits a region in which there are no boundary cells, that region is assigned to cluster 1. The square is then moved farther, and if no boundary cells are encountered, the area within the square is assigned to cluster 1. The scanning continues until all possible cells are assigned to that cluster. Next, the square is moved until it hits a new region with no boundary cells, and the process repeats itself. In the third step, clusters are merged according to their spectra measurements.

Robertson *et al.* [18] describe a procedure for successively partitioning a multi-image into rectangular blocks. However, the final clustered images produced by the algorithm yield images which suffer from excessive blockiness.

III. A Spatial Clustering Procedure Based on Gradient Images

The spatial clustering procedure discussed here is based on Dinstein [4] and takes into account the distribution of the measurements in the measurement space as well as their distribution in the spatial domain of the image. The procedure consists of 1) computation of a gradient image, 2) thresholding the gradient image, 3) cleaning the thresholded image, 4) labeling connected regions in the cleaned image, and 5) clustering the labeled connected regions. We now define functions and operators for those operations, using the notation as defined in the previous section.

Define a function $GI: Z_r \times Z_c \rightarrow R$ to be the gradient image of I. GI assigns a real number to each resolution cell in $Z_r \times Z_c$. This real number is relative to the changes in grey levels of resolution cells in a neighborhood of that resolution cell. Therefore, we assume that when a resolution cell of the GI image contains a high number, it indicates that the resolution cell is close to a category boundary. It indicates that there is a significant change in the grey tones in its neighborhood. On the other hand, a low value assigned to a resolution cell in the GI image indicates that the resolution cell's neighborhood is homogeneous. It indicates that all the measurement vectors assigned to the resolution cells in its neighborhood are close to its measurement vector. The objective of the proposed spatial clustering is to detect such homogeneous neighborhoods. The next step for achieving this objective is to threshold the GI image in order to distinguish between homogeneous and nonhomogeneous areas.

Define a function $H: Z_r \times Z_c \rightarrow \{1,0\}$ as follows:

$$H(i,j) = \begin{cases} 1, & \text{iff } GI(i,j) \leq \theta \\ 0, & \text{otherwise} \end{cases}$$

where θ is a specified threshold. In our experiments, we have used a dynamic threshold defined as follows. Let

$$S_i = \{(k,l) \in Z_r \times Z_c \mid i - L \leq k \leq i + L\}$$

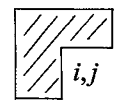
where L is a specified integer. The threshold θ_i for the ith line of the gradient image is the mean of the gradient values of resolution cells in S_i, as seen by

$$\theta_i = \frac{1}{\#(S_i)} \sum_{(k,l) \in S_i} GI(k,l).$$

Before we define the cleaning operator, we define three types of neighborhoods as follows. Let

$$N(i,j) = \{(m,n) \in Z_r \times Z_c \mid m = i - 1 \text{ and } |n - j| \leq 1 \text{ or } m = i \text{ and } n = j - 1\}.$$

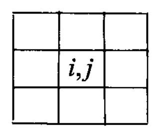
$N(i,j)$ consists of the three nearest neighbors above (i,j) and the one to the left of (i,j).



The Neighborhood $N(i,j)$.

Let

$$N^*(i,j) = \{(m,n) \in Z_r \times Z_c \mid |m - i| \leq 1 \text{ and } |n - j| \leq 1\}.$$



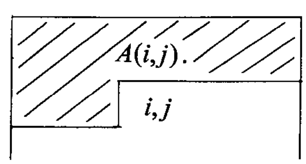
The Neighborhood $N^*(i,j)$.

$N^*(i,j)$ consists of the eight nearest neighbors of (i,j).

Let

$$A(i,j) = \{(m,n) \in Z_r \times Z_c \mid m < i \text{ or } m = i \text{ and } n < j\}.$$

$A(i,j)$ consists of all the resolution cells above and to the left of (i,j)



The Neighborhood $A(i,j)$.

Define a cleaning operator (CL) by

$$\mathrm{CL}(i,j) = \begin{cases} 1, & \text{if } \#\{(k,l) \in N^*(i,j) \mid H(k,l) = 1\} \geq \theta_2 \\ 0, & \text{otherwise} \end{cases}$$

where θ_2 is a specified threshold, and $\#$ denotes the number of elements in the set.

The purpose of the cleaning operator is to eliminate fuzzy boundaries. We want the cleaned regions to be most representative of their clusters. Therefore, a resolution cell is considered as homogeneous if and only if at least θ resolution cells in its nearest neighborhood are homogeneous.

Now we define an operator which labels the connected homogeneous regions in the cleaned image. A region is said to be connected if and only if between any two of its resolution cells, there exists a sequence of resolution cells belonging to the region, such that each two consecutive cells in the sequence are nearest neighbors. This operator is similar to a labeling operator proposed by Rosenfeld [20].

Define the function C in a sequential manner, assigning the resolution cells line by line.

1) If CL$(1,1) = 0$, set $C(1,1) = 0$. If CL$(1,1) = 1$, set $C(1,1) = 1$.

2) Suppose all the resolution cells in $A(i,j)$ have been assigned, then

a) if CL$(i,j) = 0$, set $C(i,j) = 0$, and

b) if CL$(i,j) = 1$, then

α. if $\{(m,n) \in N(i,j) \mid C(m,n) \neq 0\} \neq \phi$, set $C(i,j) = \min_{(m,n) \in N(i,j)} \{C(m,n) \mid C(m,n) \neq 0$, and

β. if $\{(m,n) \in N(i,j) \mid C(m,n) \neq 0\} - \phi$, set $C(i,j) = 1 + \max_{(m,n) \in A(i,j)} C(m,n)$.

The function C assigns labels to connected regions. These connected regions, however, are not maximal. Once connected homogeneous regions are detected and assigned integer labels, we want to detect regions which are maximal with respect to connectivity and homogeneity. This is done by merging all the homogeneous and connected areas which are connected to one another.

We define a merging operator CM as follows. Let C_N and C_M be connected regions whose resolution cells were assigned by C. If there exists $(i,j) \in C_N$ and $(k,l) \in C_M$ such that (i,j) and (k,l) are nearest neighbors, CM merges those two regions by assigning min $\{(N,M)\}$ to all the resolution cells of $C_N \cup C_M$. Apply the merging.

The last operator to be defined is the clustering operator. The elements to be clustered are the labeled connected homogeneous regions.

Let

$$L_i = \{(k,l) \mid CM(k,l) = i\}.$$

L_i is the set of resolution cells belonging to the region labeled i. Let $L = \{L_1, L_2, \cdots, L_K\}$ be the set of all the regions labeled by the merging operator, and let CODE $= \{1,2,\cdots,K\}$ of a set of cluster codes. Define a clustering function CC: $L \rightarrow$ CODE. The function CC assigns a cluster code to each one of the labeled connected and homogeneous regions. We define CC in a sequential manner, as follows.

1) $N = 1$. $CC(L_i) = N$ if and only if $\#L_i = \max_{1 \leq j \leq K} [\#L_j]$.

2) $CC(L_k) = N$ if and only if $d[I(L_k), I(N)] \leq \theta_3$ where $I(L_k)$ is the mean of the grey level vectors of the resolution cells in L_k, and $I(N)$ is the mean of the grey level vector of resolution cells of regions that have been assigned to cluster N. d is a distance function defined on the set of grey level vectors, and θ_3 is a specified threshold.

3) Repeat Step 2 until no more assignments are possible. If all L_j, $1 \leq j \leq K$ were assigned, then stop. If not, then go to Step 4.

4) $N = N + 1$.

5) $CC(L_m) = N$ if and only if

$$^{\#}L_m = \max_{\substack{\text{all } j \text{ such that } L_j \\ \text{has not been assigned}}} {^{\#}L_j}.$$

6) Go to Step 2.

The threshold θ_3 is specified by the user as follows: Once the largest of the assigned regions is found, the distances between its mean and the means of the unassigned regions are computed, quantized, and displayed in a histogram. The user specifies the thresholds according to these histograms.

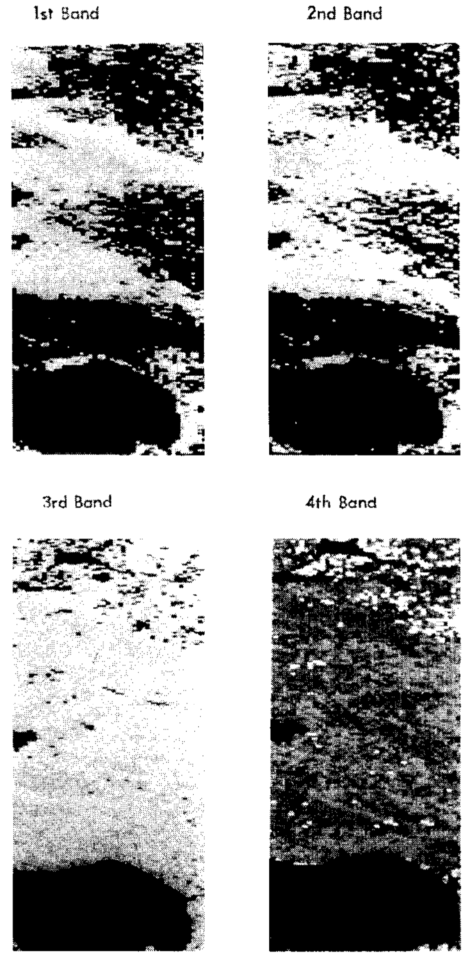


Fig. 1. Four bands of the compressed ERTS image (Monterey Bay, July 25, 1972).

IV. A Parametric Study

The basic steps in the detection of connected homogeneous regions on an image are generation and then thresholding of the gradient image, cleaning of the thresholding image, and labeling the connected regions in the cleaned image. There are many ways of generating gradient images choosing the thresholding constant and cleaning the thresholded image. Because those are user selected factors, we have carried out an experiment in order to study the effect of some of these factors on the detection of connected homogeneous regions and the clustering of those regions.

Three types of gradients, three thresholds for the gradient images, and two cleaning thresholds yielded 18 combinations of gradient–thresholds. A compressed strip (each resolution cell is the average of an 8×8 subimage on the original) of an ERTS image (Monterey Bay, ERTS Image identification 1002-18134, see Fig. 1) was processed using those 18 combinations. The average variance over the connected regions was computed (for each combination of gradient–thresholds) in order to compare the homogeneity

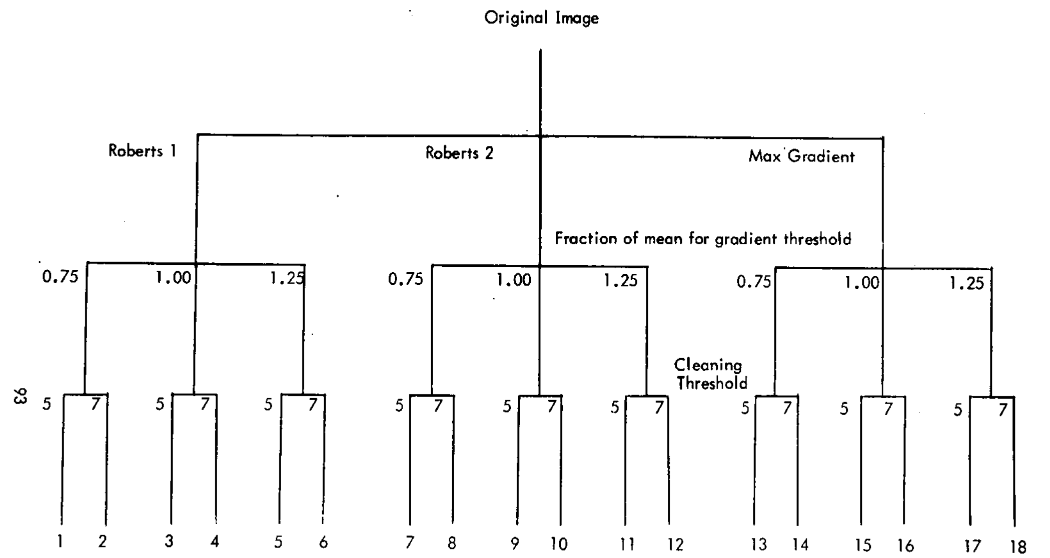
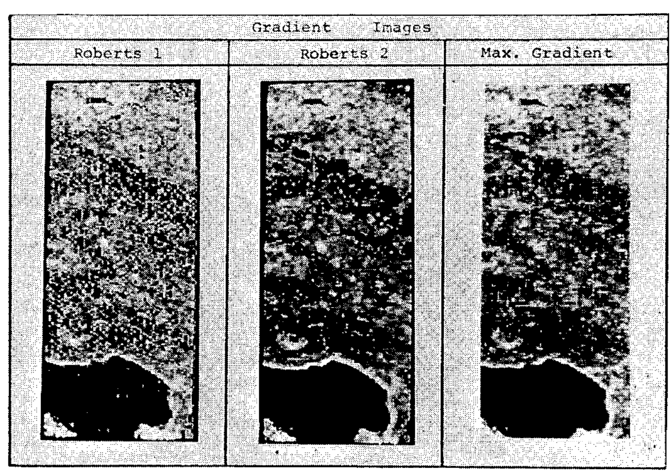


Fig 2. The 18 processing possibilities using 3 kinds of gradients, 3 threshold values, and 2 cleaning parameters.



(a) (b) (c)

Fig. 3. Gradient images. (a) Robert's distance 1. (b) Robert's distance 2. (c) Maximum gradient. (Bright grey levels represent high gradient values.)

of homogeneous regions of different combinations. The quick Robert's gradient at distance 2, a dynamic thresholding constant of the running mean, and a cleaning parameter requiring at least seven homogeneous nearest neighbors, yielded the smallest average variance of the connected regions.

The gradients selected for the study are defined as follows. Let $I(i,j) = (I(i,j,1), I(i,j,2), \cdots, I(i,j,N))$ be a vector of integer components representing the N tuple of grey levels of the resolution cell (i,j).

a) Robert's (distance 1) gradient is defined by

$$R(i,j) = \sum_{n=1}^{N} (|I(i,j,n) - I(i+1, j+1, n)| + |I(i+1, j, n) - I(i, j+1, n)|).$$

b) An extended Robert's gradient (Robert's distance 2) is defined by

$$ER(i,j) = \sum_{n=1}^{N} (|I(i-1, j, n) - I(i+1, j, n)| + |I(i, j-1, n) - I(i, j+1, n)|).$$

c) A maximum gradient is defined by

$$M(i,j) = \max_{\substack{kl \\ k=i \text{ and } l=j+1 \text{ or} \\ k=i+1 \text{ and } j-1 \leq l \leq j+1}} \left[\sum_{n=1}^{N} |I(i,j,n) - I(k,l,n)| \right].$$

For each one of those three gradients, a running mean was computed (as defined in Section III), and the fractions of that mean to be used as thresholds were 0.75, 1.00, and 1.25. The application of three thresholds to three gradient images yielded nine thresholded images. Two cleaning thresholds were applied in the cleaning process of those

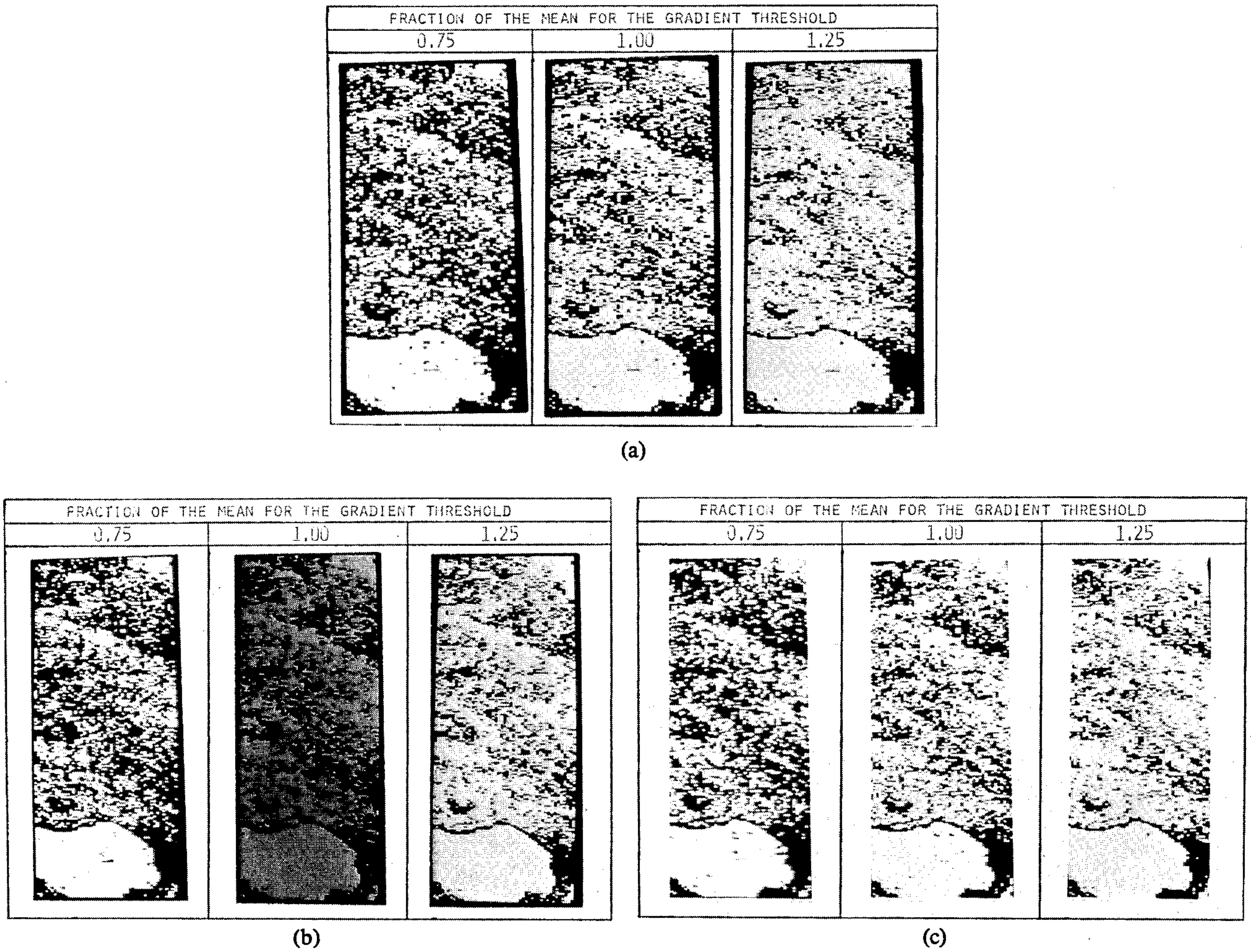


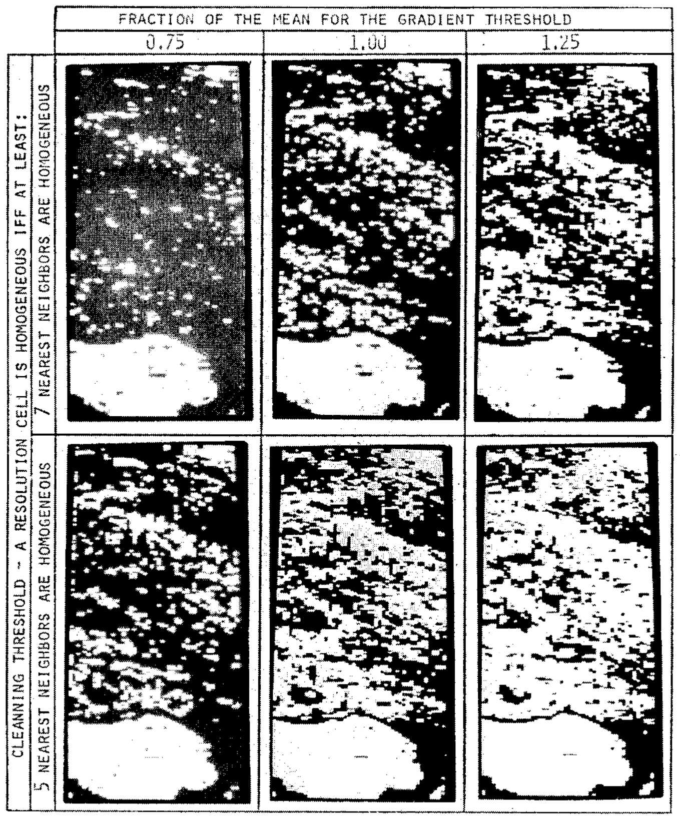
Fig. 4. Thresholded images. (a) Robert's gradient distance 1 (white regions are homogeneous). (b) Robert's gradient distance 2 (white regions are homogeneous). (c) Maximum gradient (white regions are homogeneous).

nine thresholded images to yield eighteen cleaned homogeneous images. In the cleaning process, a resolution cell was considered to be homogeneous if and only if more than K of its nearest neighbors were homogeneous. K, the cleaning threshold, was set to 5 or 7. The 18 combinations of those steps are summarized in Fig. 2. The three gradient images of the processed strip of Monterey Bay ERTS image are shown in Fig. 3. The dark areas at the bottom of the images indicate that the gradient within this region (Monterey Bay – water) is low. One can tell by looking at the three gradients, that Robert's distance 2 detected the borders of the bay better than the other two gradients. The bright strip (high gradient values) around the bay is more distinct and continuous in the Robert's 2 gradient than in the other two gradients.

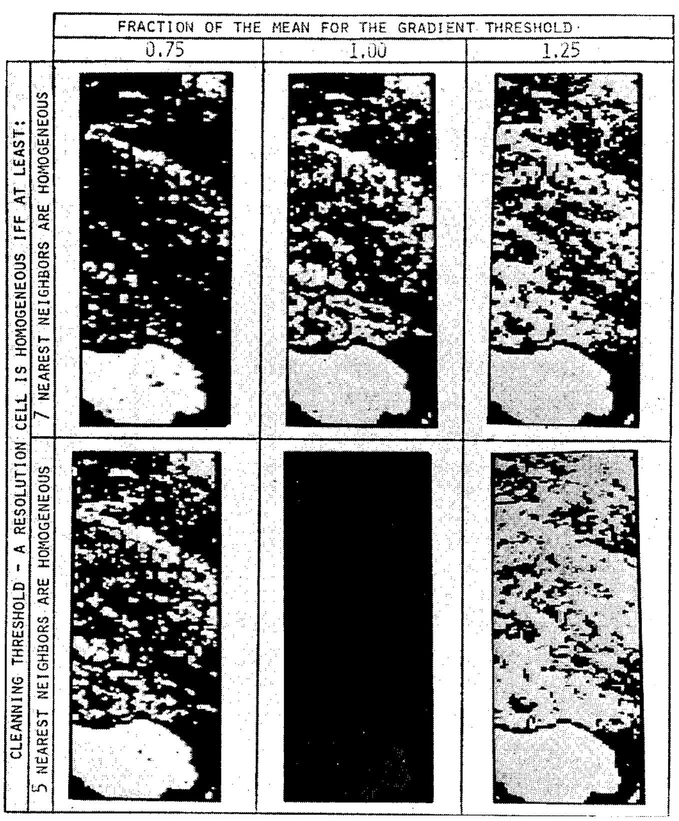
The thresholded images of Robert's distance 1, Robert's distance 2, and maximum gradient are given in Fig. 4(a), (b), and (c), respectively. The white regions represent resolution cells with gradient value less than the corresponding threshold. Obviously, the higher the threshold, the larger are the white regions. Note that nice closed boundaries like those in the artificial block world or those in face recognition are not present in our satellite imagery.

The cleaned thresholded images are shown in Fig. 5(a), (b), and (c). The cleaned images obtained with gradient threshold 0.75 of the running means and cleaning threshold of 7 (the upper left image in Fig. 5(a), (b), (c) show that beside the region of Monterey Bay, only scattered cells were considered as homogeneous. Images obtained with gradient threshold of 1.25 of the running means and cleaning threshold of 5 show that most of the resolution cells were considered as homogeneous and, besides some scattered resolution cells, there are only two connected homogeneous regions in those images. The conclusion is, of course, that both these extreme threshold combinations are not "good" for the clustering.

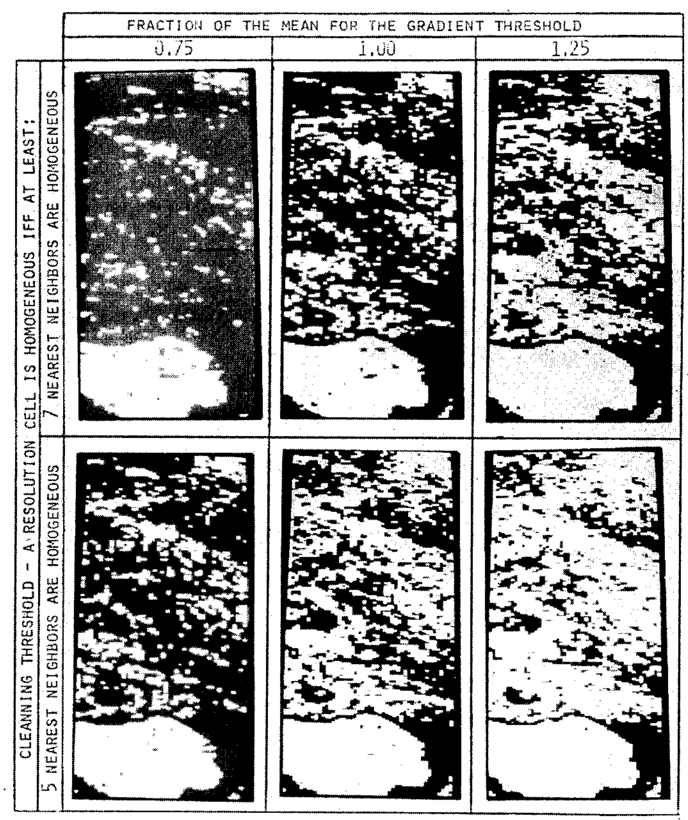
As to the other combinations of gradient–thresholds, the decision as to "what combination is the best one" is not that obvious. We want the detected homogeneous regions to represent homogeneous regions in the original data; but how does one measure that representability? This is a complicated problem that might involve considerations of shapes of regions, contrast within regions, human perception and other factors. Such a study is beyond the scope of our study. Since our spatial clustering procedure is based on the detection of homogeneous regions, we want a quantitative



(a)



(b)



(c)

Fig. 5. Cleaned thresholded images. (a) Robert's gradient distance 1 (white regions are homogeneous). (b) Robert's gradient distance 2 (white regions are homogeneous). (c) Cleaned threshold images.

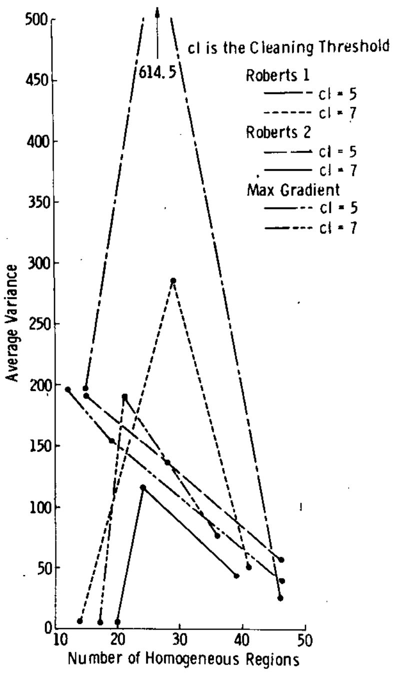


Fig. 6. Average variance versus number of homogeneous regions.

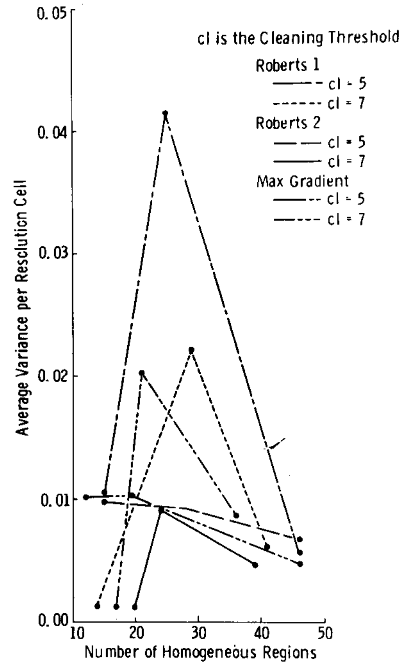


Fig. 7. Average variance per point versus number of homogeneous regions.

criterion that will enable us to compare the different gradients and thresholds and to decide which one yielded regions which are more homogeneous.

We assume that the more homogeneous a region is, the less is the variance of the measurements within the region. Let $R = \{R_1, R_2, \cdots, R_K\}$ be K connected homogeneous regions of an image, and let $I(i,j) = (I(i,j,1), \cdots, I(i,j,N))$ be the measurement vector of resolution cell (i,j).

The kth component of the mean vector of the measurements within R_m is defined by

$$\overline{X}_{mk} = \frac{1}{\#R_m} \sum_{(i,j) \in R_m} X(i,j,k)$$

where $\# R_m$ is the number of resolution cells in the region R_m. The trace of the covariance matrix of the measurement vectors within R_m is given by

$$V_m = \frac{1}{\#R_m} \sum_{(i,j) \in R_m} \sum_{k=1}^{K} [X(i,j,k) - \overline{X}_{mk}]^2.$$

The value that we have chosen to indicate the homogeneity of an image with labeled connected homogeneous regions is the average of the traces of the covariance matrices of the connected homogeneous regions, it is defined by

$$VH = \frac{1}{\sum_{m=1}^{N} (\#R_m)} \cdot \sum_{m=1}^{N} \#R_m \cdot V_m.$$

We refer to VH as the average variance of the connected homogeneous regions. VH is a function of the number of homogeneous regions and of the number of homogeneous resolution cells. Consider the extreme case in which each connected homogeneous region consists of a single resolution cell. VH for such an image is zero. Indeed, such regions are as homogeneous as can be but this is not the kind of

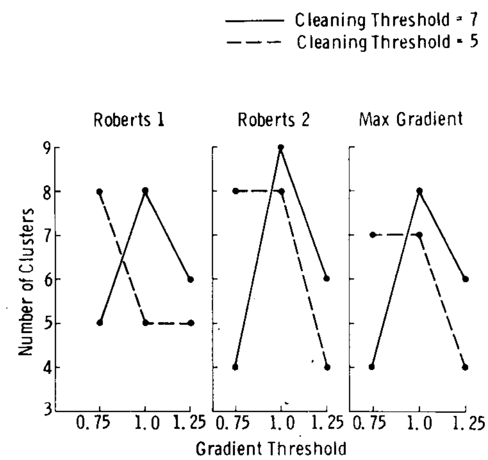
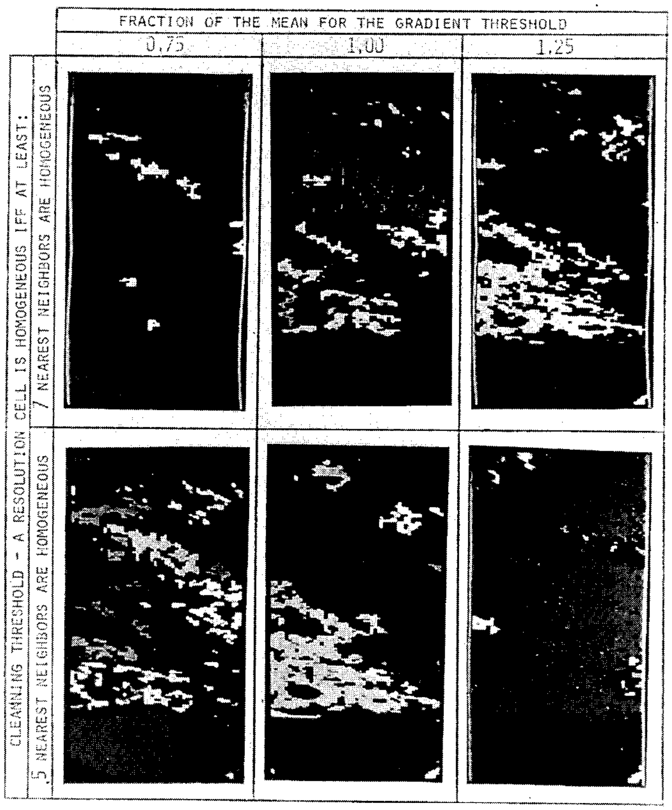
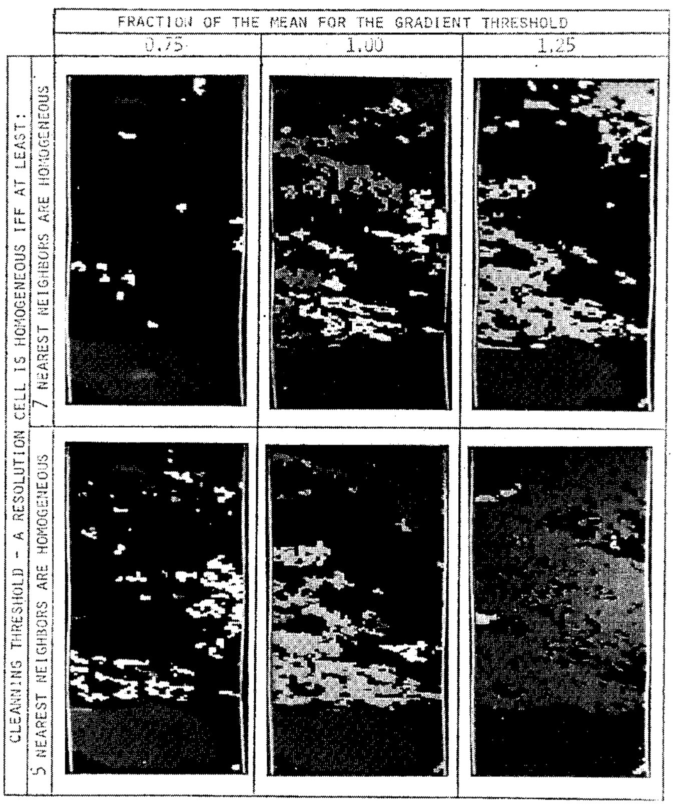


Fig. 8. Number of clusters for combination of gradient–thresholds.

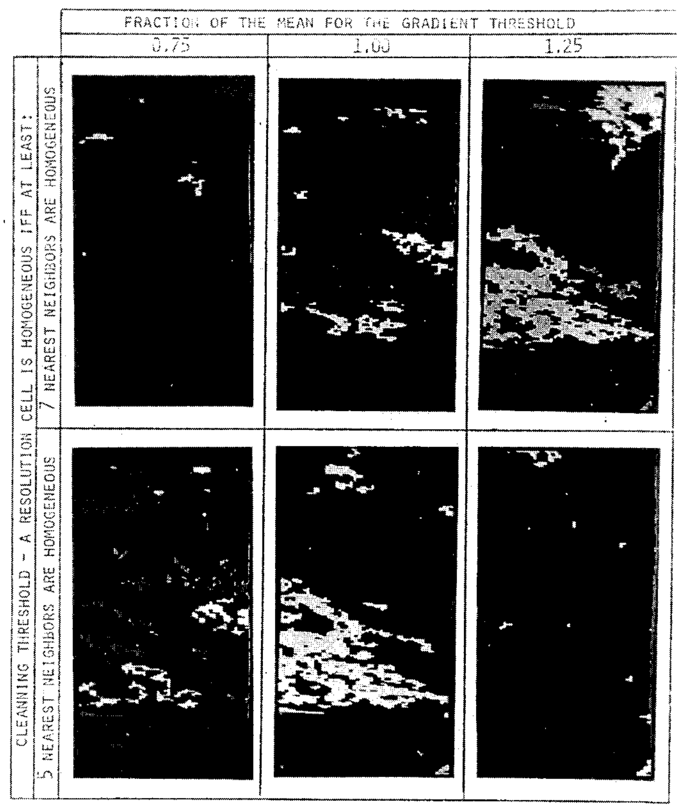
homogeneity we are interested in. For this reason, we ignored connected homogeneous regions containing less than 20 resolution cells, and we have plotted (for each image) the values of VH and the values of VH per homogeneous resolution cell versus the number of connected homogeneous regions. These graphs are shown in Figs. 6 and 7, respectively. As can be seen from those graphs, the Robert's distance 2 gradient, with cleaning threshold 7, yielded smaller average variances than the other gradients for given numbers of connected homogeneous regions. The graphs show that there is a tendency of convergence when the number of connected homogeneous regions is less than 20 or more than 40. But the clustering results show that the clustering is finer when the number of connected homogeneous regions is between 20 and 40. The number of clusters versus the number of connected homogeneous regions is plotted in Fig. 8. Based on Figs. 6, 7, and 8, we



(a)



(b)



(c)

Fig. 9. Clustered images. (a) Robert's gradient distance 1. (b) Robert's gradient distance 2. (c) Maximum gradient.

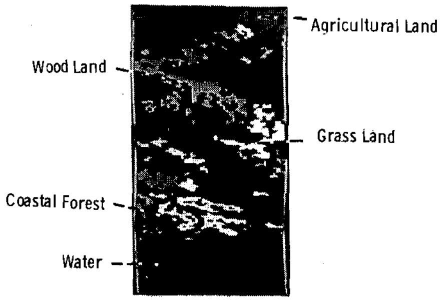


Fig. 10. The categories to which the clusters correspond for Robert's gradient distance 2, cleaning parameter 7, and gradient-threshold of 1.

can conclude that the combination of Robert's distance 2 gradient, gradient threshold equal to a running mean of the gradient image, and a cleaning threshold of 7 yielded a combination of a small average variance in the connected homogeneous regions and a large number of clusters.

Fig. 9(a), (b), and (c) shows the clustered image for various parameter combinations.

The final value of a clustering procedure is the extent to which the clusters are interpretable. In this regard by examining Fig. 10, we find that cluster 1 corresponds to water, cluster 2 to grassland, cluster 3 to coasted forest, cluster 4 to woodland, and cluster 5 to agricultural land.

There is yet much more refinement and experimentation which needs to be done with spatial clustering procedures. What happens to the clustering when an equal probability quantizing is done on the individual images in the multi-image set? How much spatial averaging should be done on the multi-image before clustering is begun? What can be done about the potential problem of distinct spatial clusters merging because of a low gradient bridge between them? What can be done about clustering the less homogeneous areas? We hope to be able to answer some of these questions in a future paper.

V. Concluding Remarks

A cluster analysis procedure which considers the spatial distribution of the measurements in image data, as well as their distribution in measurement space, has been introduced. This spatial clustering procedure is based on the detection and labeling of connected and homogeneous regions in an image, and the assignments of those regions to clusters. This is based on the assumption that resolution cells within a connected and homogeneous region do belong to the same object or field. The detection of connected homogeneous regions in an image is accomplished by computing a gradient image, thresholding the gradient image, and cleaning the thresholded image. A parametric study was carried out to evaluate the effect of the choice of the gradient, the gradient threshold, and the cleaning parameter on the homogeneity of the detected homogeneous regions. The combination of Robert's gradient of distance 2, a running mean as a threshold for the gradient image, and a cleaning procedure that considered a resolution cell to be homogeneous if at least 7 of its nearest neighbors were homogeneous yielded regions with the least average variance among the combinations that were studied.

References

[1] G. H. Ball, "Data analysis in the social sciences; what about the details?" *Proc. Fall Joint Computer Conf.*, Las Vegas, Nevada, Dec. 1965, pp. 535–559.
[2] R. E. Bonner, "On some clustering techniques," *IBM J.*, vol. 8, no. 1, pp. 22–32, Jan. 1964.
[3] C. R. Brice and C. L. Fennema, "Scene analysis using regions," *Artificial Intelligence*, vol. 1, pp. 205–226, 1970.
[4] I. Dinstein, "Cluster analysis and its application to imagery data," unpublished Ph.D. dissertation, University of Kansas, Jan. 1974.
[5] J. S. L. Gilmour, "Taxonomy and philosophy," *The New Systematics*, J. Huxley, Ed. New York: Oxford, 1940.
[6] E. E. Gose, L. V. Ackerman, A. N. Mucciardi, and F. S. Alcorn, "Computer diagnosis of breast cancer," *Proc. Computer Image Processing Recognition*, Columbia, Missouri, Aug. 1972.
[7] R. M. Haralick and G. L. Kelly, "Pattern recognition with measurement space and spatial space for multiple images," *Proc. IEEE*, vol. 57, no. 4, pp. 654–665, Apr. 1969.
[8] R. M. Haralick and I. Dinstein, "An iterative clustering procedure," *IEEE Trans. SMC*, vol. SMC-1, no. 3, pp. 275–289, July 1971.
[9] R. M. Haralick and J. G. Haralick, "Behavioral problems of deaf children: clustering of variables using measures of association and similarity," *Pattern Recognition J.*, vol. 3, pp. 269–280, 1971.
[10] M. H. Hueckel, "An operator which locates edges in digitized pictures," *J. Association Comput. Machinery*, vol. 18, no. 1, pp. 113–125, Jan. 1971.
[11] Robert R. Jayroe, Jr., "Unsupervised spatial clustering with spectral discrimination," NASA Tech. Note TN D-7312, Marshall Space Flight Center, Alabama, May 1973.
[12] G. Kaskey, P. R. Krisnaiah, and A. Azzari, "Cluster formation and diagnostic significance in psychiatric symptom evaluation," *AFIPS Conf. Proc.*, vol. 22, Fall Computer Conf., 1962.
[13] R. B. McCammon, "The dendrograph: a new tool for correlation," *Geological Soc. America Bulletin*, vol. 79, Nov. 1968, pp. 1663–1670.
[14] G. Nagy, "State of the art in pattern recognition," *Proc. IEEE*, vol. 56, no. 5, pp. 836–862, May 1968.
[15] G. Nagy and J. Tolaba, "Nonsupervised crop classification through airborn multispectral observations," *IBM J. Res. Develop.*, vol. 16, no. 2, Mar. 1972.
[16] Judith M. S. Prewitt, "Object enhancement and extraction," in *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld, Eds.: New York: Academic Press, 1970, pp. 75–149.
[17] L. G. Roberts, "Machine perception of three-dimensional solids," *Optical and Electro-Optical Processing of Information.* Cambridge, Mass.: MIT Press, 1965, pp. 159–197.
[18] T. V. Robertson, K. S. Fu, and P. H. Swain, *Multispectral Image Partitioning*, LARS Inf. Note 171373, Purdue University, Indiana, 1973.
[19] D. J. Rogers and T. T. Tanimoto, "A computer program for classifying plants," *Science*, vol. 132, pp. 1115–1118, Oct. 21, 1960.
[20] A. Rosenfeld and J. Pfaltz, "Sequential operations in digital picture processing," *J. ACM*, vol. 13, no. 4, pp. 471–494, Oct. 1966.
[21] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis," *IEEE Trans. Comput.*, vol. C-20, no. 5, pp. 562–569, May 1971.
[22] H. W. Smedes, H. J. Linnerud, L. B. Woolover, M. Y. Su, and R. R. Jayroe, "Mapping of terrain by computer clustering techniques using multispectral scanner data and using color aerial film," *NASA 4th Ann. Earth Resources Program Rev.*, MSC-05937, vol. III, Houston, Texas, Jan. 1972, pp. 61-1–61-30
[23] P. H. A. Sneath, 'Some thoughts on bacterial classification," *J. Gen. Microbio.*, vol. 17, pp. 184–200, 1957.
[24] R. R. Sokal and P. H. A. Sneath, *Principles of Numerical Taxonomy.* San Francisco, Calif.: W. H. Freeman and Company, 1963.
[25] K. Sparck and D. Jackson, "Current approaches to classification and clump-finding at the Cambridge Language Research Unit," *Comput. J.*, vol. 10, no. 1, 1967.
[26] A. G. Wacker and D. A. Landgrebe, "Boundaries in MSS imaging by clustering," Proc. 9th IEEE Symp. Adaptive Processer, Dec. 1970.
[27] J. H. Ward, Jr., "Hierarchical grouping to optimize an objective function," *J. Amer. Stat. Ass.*, vol. 58, pp. 236–245, Mar. 1963.

[28] Satosi Watanabe, "Pattern recognition as an inductive process," *Methodologies of Pattern Recognition*, S. Watanabe, Ed. New York: Academic Press, 1969, pp. 521–533.
[29] D. E. Wolf, D. J. Hall, and E. M. Endlich, "A status report on an objective system for determining cloud motions," *Proc. Comput. Image Processing Recognition*, Columbia, Missouri, Aug. 1972.

Robert M. Haralick (S'62–M'69) was born in Brooklyn, N.Y., on September 30, 1943. He received the B.A. degree in mathematics, the B.S. degree in electrical engineering, the M.S. degree in electrical engineering, and the Ph.D. degree in electrical engineering from the University of Kansas, Lawrence, in 1964, 1966, 1967, and 1969, respectively.

He has worked with Autonetics and IBM. He joined the Center for Research, Inc., University of Kansas, in 1965, and is presently an Associate Professor of Electrical Engineering at the University of Kansas. He has done research in pattern recognition, adaptive self-organizing systems, multi-image processing techniques, texture analysis, clustering, and general systems theory.

Dr. Haralick is a member of Sigma Xi, Eta Kappa Nu, the Association for Computing Machinery, the Pattern Recognition Society, and the Society for General Systems Research.

Its'hak Dinstein (S'70–M'73) was born in Haifa, Israel, on January 15, 1939. He received the B.S.E.E. degree in 1965 from the Technion—Israel Institute of Technology, Haifa, Israel, and the M.S.E.E. degree in 1969 from the University of Kansas, Lawrence, where he is now a Ph.D. candidate.

From 1965 to 1967, he was employed by Elron-Elbit, Israel, where he was engaged in development of digital instrumentation. He was a Teaching Assistant at the University of Kansas during 1967–1968, and is currently with the Center for Research, Inc., the University of Kansas, working in the area of pattern recognition.

Separation of Man-Made and Natural Patterns in High-Altitude Imagery of Agricultural Areas

ALFRED S. SAMULON

***Abstract*—A nonstationary linear digital filter is designed and implemented which extracts the natural features from high-altitude imagery of agricultural areas. Essentially, from an original image a new image is created which displays information related to soil properties, drainage patterns, crop disease, and other natural phenomena, and contains no information about crop type or row spacing.**

A model is developed to express the recorded brightness in a narrowband image in terms of man-made and natural contributions and which describes statistically the spatial properties of each. The form of the minimum mean-square error linear filter for estimation of the natural component of the scene is derived and a suboptimal filter is implemented. Nonstationarity of the two-dimensional random processes contained in the model requires a unique technique for deriving the optimum filter.

Finally, the filter depends on knowledge of field boundaries. An algorithm for boundary location is proposed, discussed, and implemented.

Manuscript received April 1, 1974; revised August 21, 1974. This work was supported by the National Aeronautic and Space Administration under Grant NASA-NGR-05-003-404.
The author was with the Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, Calif. He is now with TRW Systems, Redondo Beach, Calif., 90278.

I. Introduction

TWO PROJECTS designed to explore the possibilities of application of extremely high-altitude imagery to the study of earth resources have born fruit in the last two years. The Earth Resource Technology Satellite (ERTS) and Skylab programs have provided thousands of images taken at altitudes of several hundred miles.

Image processing of various types can be a valuable tool in preparing images for viewing, substantially increasing their utility. Several categories of image processing exist.