[illegible]

JSC-14456
80-10251
NASA CR-
160716

CLASSIFICATION WITH
SPECTRAL-SPATIAL-TEMPORAL ARCHETYPES

Job Order 71-593

(E80-10251) CLASSIFICATION WITH SPECTRAL-SPATIAL-TEMPORAL ARCHETYPES (Lockheed Electronics Co.) 21 p HC A02/MF A01 CSCL 05B

N80-30834

Unclas
G3/43 00251

Prepared by
Lockheed Electronics Company, Inc.
System and Services Division
Houston, Texas
Contract NAS 9-15200
For
EARTH OBSERVATIONS DIVISION
SCIENCE AND APPLICATIONS DIRECTORATE

National Aeronautics and Space Administration
LYNDON B. JOHNSON SPACE CENTER
Houston, Texas

August 1978 LEC-12691

JSC-14456

CLASSIFICATION WITH SPECTRAL-SPATIAL-TEMPORAL ARCHETYPES

Job Order 71-593

Prepared By
P. J. Aucoin, Jr.
L. E. Giddings

APPROVED BY

John A Rainey

J. A. Rainey, Supervisor
Scientific Applications Section

Prepared By

Lockheed Electronics Company, Inc.

For

Earth Observations Division

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

August 1978 LEC-12691

CONTENTS

ABSTRACT

Both spatial and spectral relations are included in a new unique algorithm for the analysis of satellite data. Preexisting archetypes, essentially local spectral - spatial patterns, are used by machine to classify pixels. A human observer rejects or reinteprets certain classes, and the machine continues its analysis. The algorithm combines, in an optimal way, machine and human capabilities. It is discussed for the simplest case of a single line, and extension is made to a multiline environment. The method promises to be efficient in use of computers, and more accurate and precise for agricultural and other applications.

The human observer - computer activity tradeoff is based upon the idea that people are able to easily make decisions based upon global context information, whereas computers can rapidly classify data given local recognition patterns, or archetypes. These archetypes (flexible templates) seem to be relatively simple in structure, and few in number, for crop identification applications.

1. INTRODUCTION

1.1 BACKGROUND

A primary disadvantage of most methods of classification of remotely sensed data is their lack of accuracy. In particular, both parameteric and non-parameteric classifications normally used for classification of agricultural areas, all of which depend only on spectral and temporal data, are severely inaccurate by comparison with classification by human inspection of images. Because of such a difference, there is a tendency to judge the classification of a scene by inspection, since the human interpretations is automatically assumed to be the superior classifier.

1.2 THE LACIE EXPERIENCE

The case of LACIE is directly to this point. The entire classification scheme in LACIE depends on the human interpretation of a scene. An analyst furnishes the data which enables the classifier to function, and then he reviews the results with a view towards changing the results.

This in itself is a valid approach, except for one problem: the classifier, regardless of the correctness of the analysts' contributions, cannot provide an analysis that is as good as a human analysis. This has been variously observed. The computer + human has achieved a correctness considerably below the capabilities of a human alone.

The reason for this is clear: The human uses spatial contextual information, with spectral information, and some temporal information. The machine completely bypasses the spatial data. It is limited in the information it can use, and therefore its analysis is limited in accuracy, even when it operates in conjunction with a human.

There are signs that future crop inventory algorithms will use spatial information in some form; specifically, there are efforts directed toward incorporating the AMOEBA algorithm into Procedure 2, the proposed successor to Procedure 1. This is one of three or more principal procedures currently documented for use in remote sensing.

1.2 ADVANTAGES OF THIS ALGORITHM

The algorithm proposed here promises greater accuracy in classification at a reasonable price in machine time. It bears some resemblance to the human way of distingushing classes. It promises the following advantages

1. Optimal involvement of analysts and computers.
2. Careful modelling in the use of local and global contextual information.
3. Automatic registration among acquisitions and channels to at least a shift of one pixel.
4. Allocation at the sub-pixel level of boundary pixels to neighboring classes.
5. Detection of cloud and shadow components based on signatures and multiple acquisitions.
6. Detection of patterns close to the level of resolution, such as strip fields.
7. Reconciliation of computer-determined classes with human determined categories of interest.

2. FOUNDATIONS FOR THIS ALGORITHM

2.1 THE ONE LINE RESTRICTION

It seems fundamental that an algorithm that is to handle two physical dimensions, with various spectral channels and temporal information in the form of multiple acquisitions, must base its operations on a single line. For the purpose of the discussion in this section, all concepts will be referred to a single line presumed to be the result of some kind of mapping of multiband, multitemporal data into a single line. To the extent possible, illustrations will be made in a graphic way for this environment.

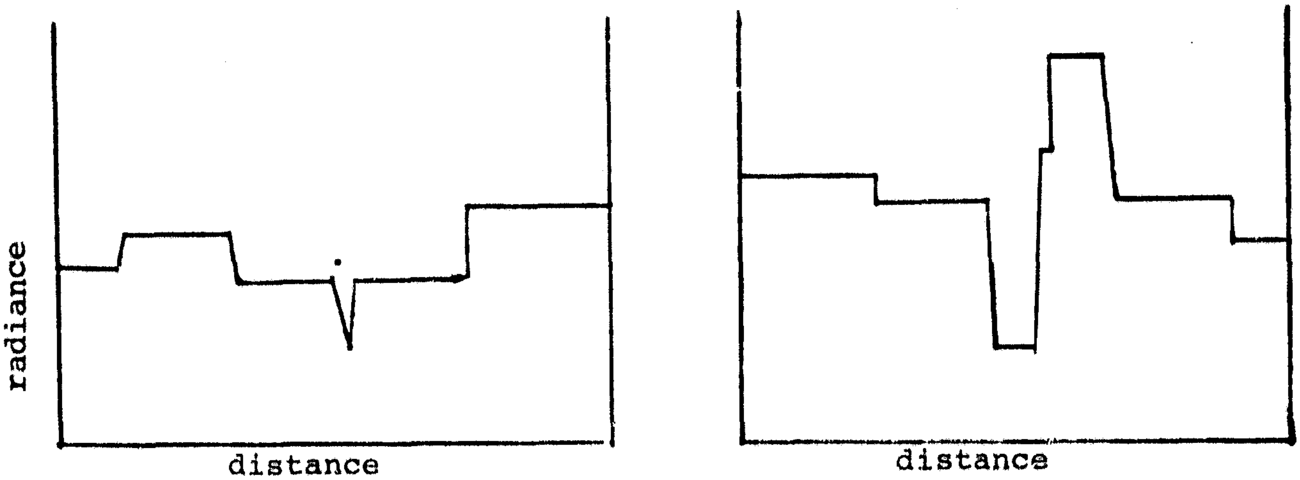
The generalization of this concept to the multiline, multichannel environment promises to be relatively straightforward. This is discussed in a later section.

For the purposes of this discussion, the data to be analyzed consist of a string of digital data. For specific references, we will consider them to consist of pixels from a line of Landsat data, with linear resolution of about 100m. However, the discussion applies to any line of digital data.

Figure 2-1 presents several lines of idealized noise-free Landsat data.

2.2 ARCHETYPES

Archetypes are patterns, flexible templates which consider spatial (contextual), temporal, and spectral information. Several simple archetypes of the type which might be used for crop and field identification are presented in Figure 2-2.



a. fields bordering a river
b. cropland, with cloud and shadow

Figure 2-1 Idealized lines of data

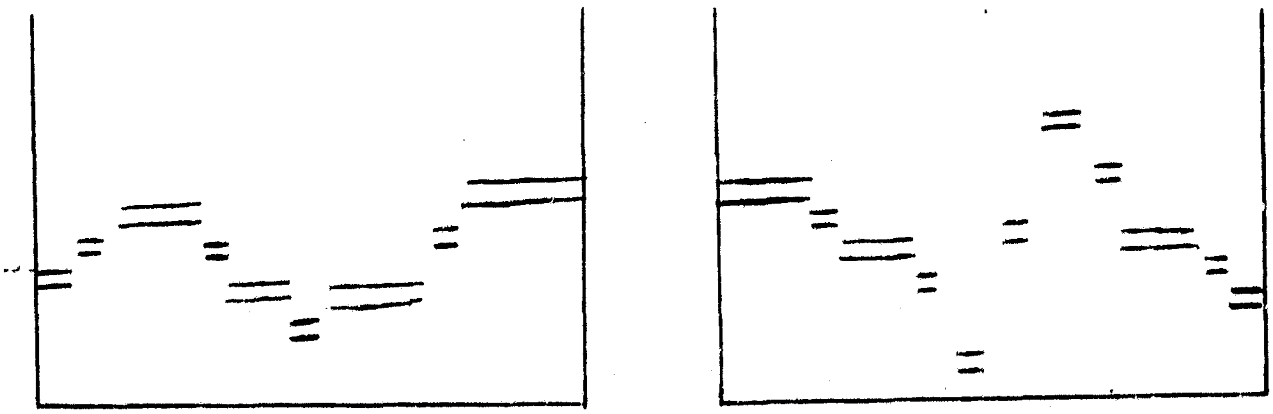


Figure 2-2 Some Idealized Archetypes

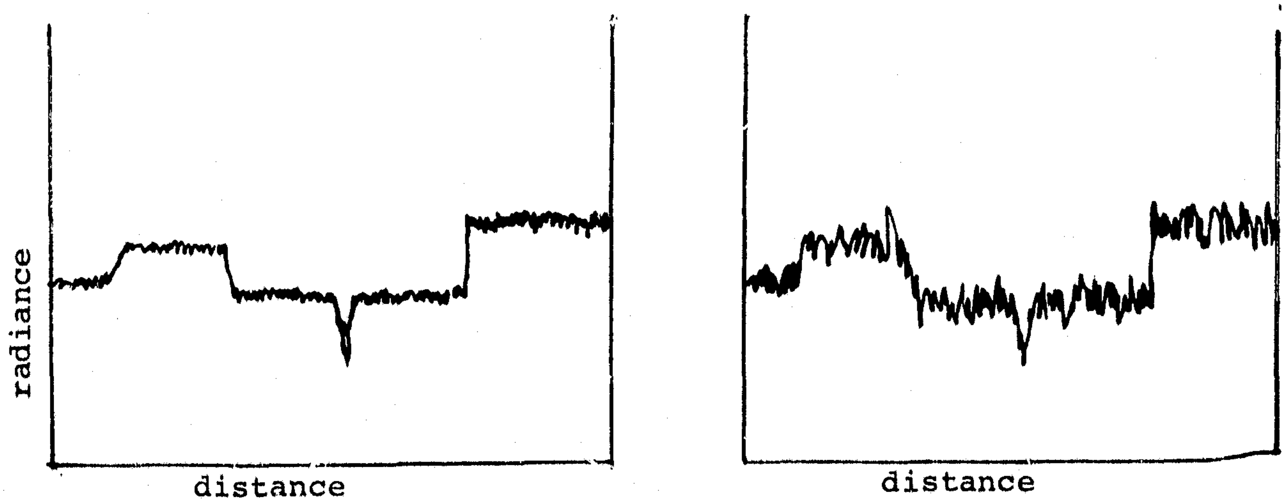


Figure 2-3 Lines of Data with Natural Variation

Note that the human analysis of the lines of data in figure 2-1 normally takes place in terms of the archetypes in figure 2-2. A person recognizes fields by their contiguous homogeneity, rivers by their V-shape, clouds and shadows by extreme difference in radiance and contiguity, etc.

2.3 NOISE, AND THE REAL CASE: GRAINESS

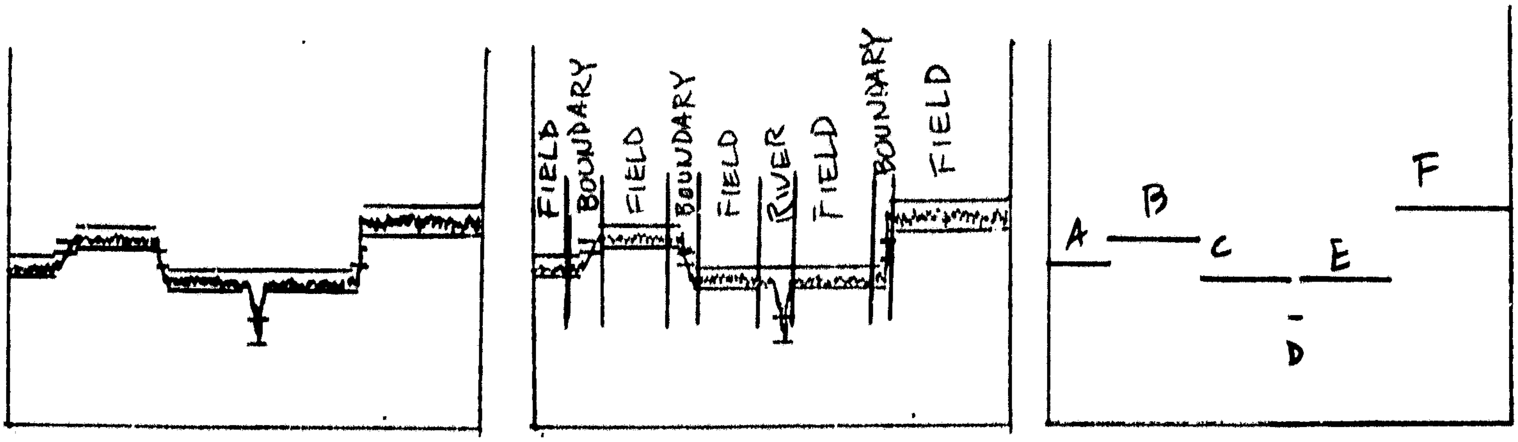
Figure 2-3 shows one of the lines in figure 2-1 under more realistic conditions, with various amounts of natural variation (noise). A human observer has as little trouble generalizing from figure 2-3 as he does from figure 2-1a. Figure 2-3b in more difficult for man or machine.

It is apparent that in addition to archetypes, the computer must be furnished some measure of graininess.

2.4 ANALYSIS OF A LINE

Given archetypes and a measure of graininess, machine analysis of a line can proceed. Figure 2-4 demonstrates that the computer will classify the line relatively into classes (figure 2-4b); interpret these classes (figure 2-4b); and derive a final classification, with boundary pixels assigned to classes.

At this time, we are not concerned with more than the straightforward classification of pixels to computer classes. This classification has resulted in the identification of fine fields (A through E) and one river. Other forms of logic can be invoked to decide where B and F represent the same crop and whether C and E, the flood plains, can be meaningfully considered to be identical.



a. Division of line

b. Assignment of Archetypes

c. Subpixel assignment and Classification

Figure 2-4

2.5 THE ROLE OF THE HUMAN

Although the method uses much computer logic, it is still necessary to have a human collaborator. In the simplest case, a human will have to approve the machine interpretation and allow processing to proceed to some kind of interpretation of results (the labelling of fields, for example). In more difficult cases, the human will have to reject a portion of the classification, or perhaps all of it, and furnish new bases for the classification. He might, for example, decide that a division into two different fields was unrealistic and therefore specific a new graininess parameter; or he might negate the use of a boundary archetype in a particular portion of the image.

The human might well fulfill other roles, although minimization of human intervention is to be strived for. He might, for example, furnish labels for certain features such as a strip field area. At this state there is little hope for complete elimination of the human from this or any other classification processor. The most that can be hoped for is the minimization of his necessary participation

2.6 MORE COMPLEX CASES

The power of this method lies in its adaptability to the diverse cases found in satellite images. A persistent problem in LACIE, for example, concerns the treatment of strip fields.

With this method, the classification of areas of strip fields is relatively straightforward.

Although the separation of strip fields as distinct classes is not especially difficult with this method, the labelling of the classes is a separate problem. Once classes have been labelled, however, it will be obvious that subpixel allocation is relatively simple.

It should be noted that with other methods, in particular with parametric classifiers based on spectral and temporal data, strip fields are classified into many distinct classes, not into a single class. As a result, subpixel allocation proves impractical with them.

3. THE TWO-DIMENSIONAL CASE

The generalization to more than one line is not particularly difficult. There are two principal considerations. The first in merely the agglomeration of lines into a complete image. The other is the compensation for alignment of features with scan lines.

3.1 COMBINATION OF LINES INTO A IMAGE

An image is the side-by-side collection of scan lines. Therefore, the job of amalgamating scan lines involves the relatively straightforward common identification of classes on adjacent lines, and their extension into other parts of an image. There are several simple strategies for doing this. The most elementary involves the processing of of single lines sequentially, and the extension of class labels from one line to the next. For crop applications, this may . ultimately feasible.

This appears to be a relatively straightforward operation, provided only that a line is not itself a boundary. If a boundary is nearly perfectly aligned with a scan line, other problems occur.

3.2 ALIGNMENT OF BOUNDARIES WITH SCAN LINES

The most straightforward way to compensate for such an alignment is the analysis of an image that has been rotated. A 90° rotation conversion of lines into pixels, and vice versa, for example, is a relatively inexpensive procedure. The data might well be analyzed both ways, with results compared, before a classification can be accepted.

It might be preferable to perform a rotation which is not a multiple of 90°. This is more expensive in computer time, but it would eliminate most cases of alignment with scan lines, given the human tendency to construct right angles. In cases of highly aligned images, several rotations may be useful.

For many reasons it would appear to be preferable to perform analyses in small overlapping subimages. Because of the elementary nature of operations on individual lines, pixel by pixel, there is little disadvantage to doing so.

Advantages of this process include the case of the sequential operations on small blocks:

analysis
rotation
analysis
comparison

Afterwards, a quice similar logic could be used for assembling the subimage classifications into a whole-image classification.

4. THE PROCEDURE IN PRACTICE

There is not yet a consensus on the nature of man-machine interaction for the analysis of imagery data. Certainly this subject becomes more important as more spatial information is incorporated into algorithms and the computers more approximate the human way of analyzing images.

4.1 STEPS IN A PROCEDURE

The following are steps that appear appropriate for this procedure. Note that the analyst himself acts only as an executive to guide the machine analysis. He makes very high level decisions; in a well-functioning system he will rarely need to descend to operations within classification; rather, he decides whether the bases of the classification are correct, and how to make use of the classification results.

1. To begin, the computer is presented with a set of spatial/spectral archetypes, a mapping from channel spectral values in an acquisition to a scalar measure, ie greenness, and an interval on this measure (say +10 on a range 0-255) for determination of computer classes. (The wider the interval, the fewer computer classes will occur). The interval width is a measure of graininess within the scene.
2. Next, the computer analyses a multi-temporal segment and produces a segment class map for inspection by the analyst. Trajectories for these computer classes are also produced. A computer interpretation for each class will be presented; for example, a boundary pixel will occupy a separate class and will be interpreted as such. A pre-scan of the data would optimally position the pure pixel archetypes up and down.

3. At this time the analyst would decide whether to accept or reject the analysis. If he does not accept it, he must direct a reanalysis of the scene. He might direct that classes be reassigned or combined, or change the parameters of analysis (such as graininess, or available archetypes). At this stage, he might well direct a complete reanalysis of the image.

   When he accepts the analysis he might then direct the application of some kind of logic to identify or agglomerate classes. He might for example use labelled dots and existing algorithms for labelling the classes and agglomerating them by categories. There are many options that could be applied here. Choice of any one is not critical to this discussion.

4. Based on these inputs, a category map of the segment is generated by the computer, and proportion estimates of each category are listed. Sub-pixel allocation among categories is used in preparing the proportion estimates.

   Note that, in this procedure, the analyst is involved both at the beginning, and toward the end, of each analysis. Regarding context information, the computer does as well as it can using the provided archetypes. The analyst makes final decisions owing to his superiority in handling spatial/contextual information.

4.2 A SPECIFIC EXAMPLE

This form of interaction between man and computer is especially powerful because the computer's logic is compatible with man's logic. For example, there might be a "strip field", diagrammed (for one line, one acquisition) as shown in Figure 4-1.

Based on archetypes, the computer may flag these pixels as belonging to a strip field. The analysis could continue as follows.

Image that the pixels representing this strip are smeared, and that if smearing were removed, the figure would show more definition, as shown in figure 4.

Then a search could be made for "pure" pixel classes with tolerances (30-40) and (10-20) within which the strip field pixels could be included.

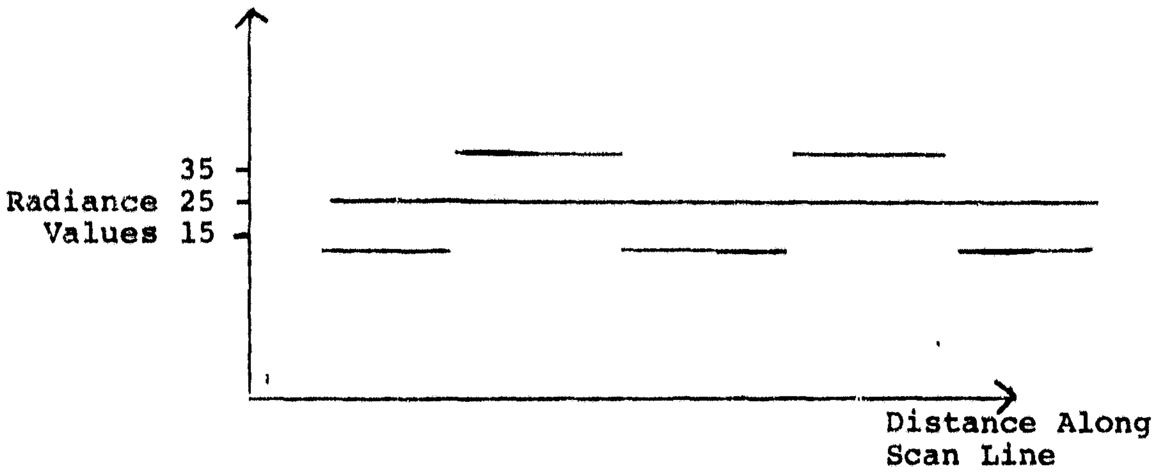
The computer could not reasonably be expected to go through this process unassisted. The analyst, for example, might reach one of the following conclusions regarding the computed-interpreted "strip field".

a. He might decide that no strip field was present, and that the signature was based on other effects (i.e. excessive noise).

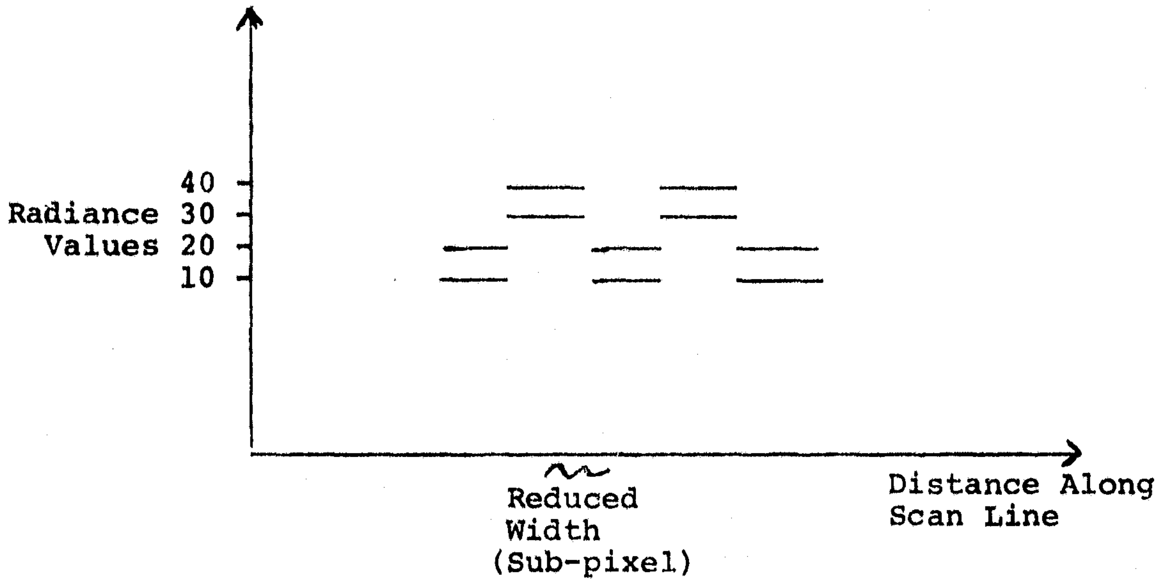
b. He might decide that the strip field was present, but that the crops should be allocated to other pure pixel classes in the scene. (A resolution decision is implicit here).

4.3 COMMENTS

Note that conventional clustering and classification are not used. The analysis is basically non-parametric. The parameters that actually exist are embedded in the archetypes, in the interpretive actions of the analyst, and in any mapping of channel (spectral) data.



Raw Values, Basic Archetypes, and Detection



Sub-pixel Allocation and Radiance Value Shifting

Figure 4-1
Strip Fields

5. OTHER COMMENTS

5.1 POSSIBLE ARCHETYPES

Some possible archetypes are presented in figure 5-1. Lines indicate upper and lower limits of the scalar measure. These are shown for a single acquisition. Side-by-side replicates shifted up or down would indicate the multi-temporal case.

These are only a few of the possibilities. These structures are solely intended for the computer interpretation. Final acceptance by an analyst will depend on his personal archetypes, or, more objectively, on the type of scene and the type of recognition desired.

Note that the shape of the archetypes is important, but that in general, absolute values are not important. However, some can be visualized in which absolute values might be used, as with the identification of clouds in thermal imagery.

5.2 MISREGISTRATION ERRORS

Registration errors can be detected when more than one acquisition per segment is used, especially for cases 2, 4, and 5 of figure 5-1.

5.3 Preprocessing

Given the constraints of agricultural analyses, the use of temporal information (from multiple acquisitions) is necessary for optimal results.

The exact method of preprocessing is not critical. It will probably be preferable to use a vegetation index transformation (map) to reduce essential information for one acquisition to a single channel.

At this point, an algorithm will be required to agglomerate the single channels.

In general, sun angle and haze corrections will not be needed unless segment-to-segment mosaicking is required.

5.4 THE MEASURE OF GRAININESS

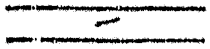
There are several ways of obtaining a graininess measure. With use, experience will certainly suggest reasonable starting values. They may also be derived from automatic analysis of histograms. Perhaps the best method is from training fields, which can be scanned automatically or defined by the analyst.

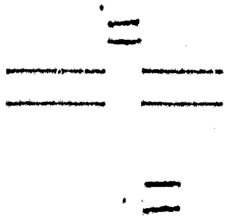
5.5 DISTANCE PARAMETERS

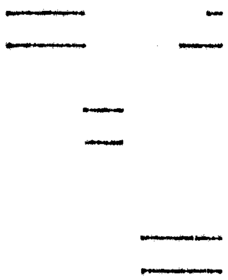
The method is not sensitive to choice of distance parameters, either in space or in spectrum. Spatial distance must recognize contiguity, and spectral distance must be significant to relative distances.

5.6 CHOICE OF CHANNELS

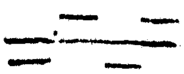
The method is also insensitive to the means used to reduce various channels into one. In some cases, it may be preferable to select a single channel with given characters. For example, for the automatic classification of water, this method will probably do as well or better with a single near-infrared channel.

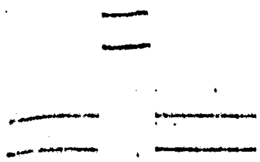
1.  large pure pixel class (one computer class) (could be 2 or 3 if pre-scanning not performed).

2.  "river"
(2 to 3 computer classes)

3.  cloud/shadow
(up to 5 computer classes)

4.  boundary
(3 computer classes)

5.   strip field
(2 or 3 computer classes)

6.  cloud, no shadow
(up to 3 computer classes)

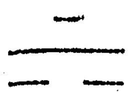
7.  noise pixel or tree

Figure 5-1

Possible Archetypes