Reprinted from

Symposium on
Machine Processing of
Remotely Sensed Data

June 3 - 5, 1975

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
75CH1009-0 -C

COMBINED SEQUENTIAL ANALYSIS OF MULTIPLE FEATURES

J.D. Nichols, W.M. Senkus

Remote Sensing Research Program
University of California
Berkeley, California 94720

I. ABSTRACT

A new technique, which greatly facilitates the computer analysis of large-area multi-feature imagery, has been developed at the R.S.R.P. Called OWNMASK, the technique permits previous classification results for an area to be used to control new classifications of that area in such a way that both the number of features and the number of classes which must be considered for each pixel are reduced significantly, thereby reducing both human labor and computer costs.

II. INTRODUCTION

Many applications of computerized processing of remote sensing data require sequential analysis of the area being studied, i.e. a series of classifications of the area using data acquired at different times. If the area in question is large, each classification can be expensive and time-consuming, requiring extraction of training areas, generation of training statistics, separability analysis, the classification itself, and perhaps repetition of this process to refine the results. Typically, however, a large part of this effort and expense is wasted, since many classes either are known to be stable or are of no interest after an initial classification. Urban areas and large bodies of water, for example, are not of primary interest in many applications; timbered and wildland areas are of little interest in agricultural applications; and agricultural areas are of little concern in forestry applications.

One method of reducing the cost of sequential analysis is stratified classification (Nichols and Senkus, 1973). Stratified analysis permits irregular shaped areas, or groups of such areas, to be classified with mutually exclusive sets of statistics and subsets of features. These areas, or strata, are chosen on the basis of gross characteristics which can be delineated by a human photo interpreter working with a photographic print or transparency of the area and a boundary digitizing device. Such characteristics might include texture, brightness, administrative boundaries, and gross land use types. The stratum to which a pixel belongs can then be used during computer classification to select from a set of independent training sets the minimal training set and feature subset necessary to classify that pixel accurately, or to specify that the pixel belongs to a stratum of no further interest, and need not be classified at all. In this latter case, the stratum number itself becomes the class for such a pixel.

Stratified analysis has one major limitation - areas on the stratification, or mask, must be relatively large, contiguous areas delineable on a photograph. Furthermore, while the use of a mask can reduce computer costs significantly, this cost reduction is offset to some degree by the added cost of generating the mask.

III. OWNMASK CLASSIFICATION

In many applications it would be convenient if previous classification results could be used as a mask. Not only would this remove the restriction that a stratum need be a large, contiguous area, it would also make possible other important applications of the concept. For these reasons, OWNMASK classification was developed.

A. DATA PREPARATION

The following data are required for OWNMASK classification:

1. A previous classification result for the area being considered.

2. Multi-feature digital data for the area, registered if necessary to correspond to the previous

classification

3. One or more sets of training statistics describing all the new classes to be found in the data.

The previous classification result can be either masked or unmasked. It need not be a detailed classification of earlier data - an inexpensive classification can be performed using a small feature set and a few broad classes to sort the data. The result of this classification can then be used to control a second, more detailed classification.

The multi-feature data submitted to OWNMASK must overlay the result of the previous classification (the mask). If data from two separate dates are involved, registration must be performed. Control points on both images are selected, regression analysis is used to determine the transformation required, and the new data are then registered to the old.

Depending on the size and complexity of the area being classified, it may be desirable to generate more than one training set. If it is known, for example, that all points which were assigned to agricultural classes on the mask are still agricultural types, all forest classes are still forest types, and so on, then it will be advantageous to develop a separate training set for each of these types. Each training set may use a different subset of the features available. This will reduce classification costs by enabling the classifier to consider for each point the minimal class and feature set required to assign it accurately to a class. It will also reduce human labor, since generation of these independent training sets will in general be simpler than generation of a single larger one.

B. CLASSIFICATION

OWNMASK classification is performed as outlined in the flowchart of Figure 1.

1. To initialize the program, the control cards prepared by the user are read. These control cards specify which strata (old classes) in the mask (old classification) are to be reclassified; to which strata each training set provided is to be applied; and which subset of the features available is to be used with each training set. The training sets are read in and verified.

2. The classification is performed line by line. A line of the mask is read in for processing.

3. The mask line is scanned, point by point. If the stratum to which a point belongs is not to be reclassified, that point is assigned at once to its result class. If a point requires reclassification, flags are set - the point itself is given a temporary value equal to the number of the training set to be applied, and note is made that the training set was referenced.

4. If no points in the current line were flagged for classification, go to step 6 to output the result line.

5. If points in the current line were flagged, the multi-feature data required is read and unpacked, and the classifier is called once for each training set referenced. The classifier scans the line for points whose value is that of the training set being applied, and reclassifies those points only.

6. The classified line is written out.

7. If more lines remain in the area, go to step 2.

8. If all data have been processed, stop.

IV. APPLICATIONS

OWNMASK classification has been utilized in a juniper inventory conducted by the RSRP for the Bureau of Land Management of the U.S.D.I. An area of 1,800,000 acres was analyzed with the stratified classification technique using ERTS-1 MSS digital data. In this initial classification the data were assigned to 58 classes, 6 of which were juniper types.

One goal of this analysis was an inventory of the volume of juniper wood in this study area on a per acre basis, independent of the classification of the area into its vegetation types. For this purpose, a separate training set applicable only to the juniper pixels was developed. This training set contained 8 classes representing the different juniper densities present. Using OWNMASK classification, this training set was applied only to the pixels assigned to juniper classes in the original classification, or approximately 20% of points classified.

Computer costs for the initial, stratified classification were approximately $200.00. Computer costs for the OWNMASK classification of the juniper pixels were approximately $60.00. It is estimated that the cost of a single, stratified classification of the entire area using a combination of the original training sets with the juniper training set would have been in excess of $300.00.

A simple comparison of these figures indicates that use of the OWNMASK technique provided a significant reduction in cost. The advantage of the OWNMASK technique is even greater than this comparison implies, however, since a single classification of the area would not have produced the same result as the two-step method described. The 8 juniper density classes in the training set applied via OWNMASK, being specific to the juniper areas, were not sufficiently separable from the classes in the original training sets. Development of training sets adequate to separate the juniper density classes from the other vegetation types present would have required significant additional effort, and the cost of classification with the resulting larger training set would have been significantly larger as well.

Use of OWNMASK classification in this case, then, not only reduced the cost of obtaining the desired results, it made it possible to achieve results that might have been unattainable without the technique.

Several other applications of OWNMASK are being tested, some of which have been mentioned above. For change detection analysis, OWNMASK is being used to follow the transition from bare soil to mature agriculture. For analysis of large areas such as that used in the BLM study described, investigation is being conducted of the procedures necessary to develop the crude training sets necessary to permit a two-step classification: the first crude classification being used to sort the data into broad classes which can then be used to control a second, detailed classification. As an extension of this application, as well as to provide an inexpensive method for refining results obtained by conventional classification, OWNMASK is being extended to permit the reclassification of pixels to be controlled by the probability of mis-classification associated with the assignment of each pixel to a class, either independently of or in conjunction with the class choice itself.

At present stratified and OWNMASK classifications are available only as options of CALSCAN, the R.S.R.P.'s Gaussian maximum likelihood classifier. Work is in progress to implement these techniques as options of our implementation of ISOCLAS, the clustering program developed at JSC in Houston.

V. CONCLUSIONS

OWNMASK classification is shown to be a unique and powerful information extraction tool which can reduce human labor and computer costs in multi-feature data classification. Human labor is reduced by simplification of training set generation. Computer costs are reduced by selecting for each point the minimal class and feature sets required to distinguish that point from classes of the same gross type.

REFERENCES

Nichols, J.D. and Senkus, W.M., "Combining Human and Computer Interpretation Capabilities to Analyze ERTS Imagery", presented at Conference on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana, October 16 - 18, 1973.

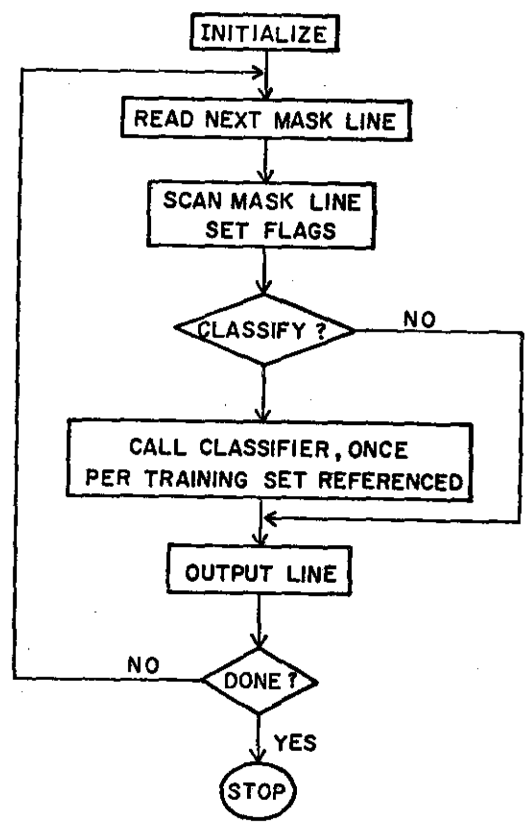

Figure 1 - Ownmask Classification Flowchart