Reprinted from

Symposium on Machine Processing of Remotely Sensed Data

June 29 - July 1, 1976

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
76CH1103-1 MPRSD

PROCAMS: A SECOND GENERATION MULTISPECTRAL-MULTITEMPORAL DATA PROCESSING SYSTEM FOR AGRICULTURAL MENSURATION*

Jon D. Erickson and Richard F. Nalepka

Environmental Research Institute of Michigan
Ann Arbor, Michigan

ABSTRACT

A prototype operational data-processing system has been defined, implemented, and tested by Multispectral Analysis Section personnel of the Environmental Research Institute of Michigan (ERIM). This system has been designed for the classification and mensuration of agricultural crops through the use of data provided by the LANDSAT satellite scanner. The specific crops for which the system was designed are the small grains including wheat, rye, oats, and barley, although the system as designed is not limited to these particular crops.

The processing system, known as PROCAMS (Prototype Classification and Mensuration System), has been built based on the experience gained and to overcome the difficulties with previous processing systems. PROCAMS takes advantage of advanced techniques and understanding to help reduce the need for classifier training information, however, in its present form it still depends on human intervention and assistance for training. The PROCAMS is designed to take advantage of multitemporal coverage while simultaneously recognizing the reality of the limited availability of complete multitemporal coverage and providing a means for accommodating multiple training sites. (This system also handles the more conventional unitemporal single training site situation.) Also addressed by PROCAMS are the data-processing problems associated with partial cloud cover, bad data lines, as well as changing sun angle and atmospheric state.

The PROCAMS, as presently defined, provides for the use of many options depending on the characteristics of the available data. Incorporated as part of PROCAMS are advanced data transformation and signature extension algorithms to allow for the use of signatures over large areas and a variety of measurement conditions.

INTRODUCTION

The need for signature extension in a large area crop inventory utilizing LANDSAT satellite multispectral scanner data is as simple and obvious as the need for other major efficiencies of time and cost. Signature extension is defined as the capability to use signatures well beyond the local time and place at which they are derived.

In the context of research and development supporting the Crop Inventory Programs,[1] testing and evaluation of algorithms and procedures for signature extension on a large number of cases covering the full range of operational conditions between area sampling segments is desirable. Both adequate data and an adequate prototype operational data processing system are necessary to the evaluation.

This paper discusses the design and implementation of an improved agricultural crop acreage mensuration system which includes signature extension as a major feature but which goes beyond simply this feature and incorporates multitemporal as well as early season unitemporal approaches and means for using multiple training sites. Also addressed are partial cloud cover and cloud shadows, bad data points and lines, as well as changing sun angle and atmospheric state, and sensor variations.

This processing system, known as Prototype Classification and Mensuration System (PROCAMS) has been designed and built to overcome the difficulties with previous processing systems and is based on experience with those systems. It has a stacked-job batch mode as well as an interactive mode. This system will be responsive to better designs as these come from research at a more reasonable level of flexibility than can a system with an operational responsibility, but at the same time provides many key features which many research systems do not. Its use features developmental feedback and interaction between testing and research.

A GENERAL DESCRIPTION

PROCAMS can be characterized as a system which incorporates unitemporal and multitemporal modes. A unitemporal mode is needed to make early season estimates. These estimates would be expected to have slightly higher variance than estimates made later in

* The effort described herein was supported by the Earth Observation Division of the NASA/Johnson Space Center under contract NAS9-14123.

the season in a multitemporal mode for sampling segments having acquisitions in later biophases. An attractive multitemporal classifier which doesn't require retraining is included if a particular data acquisition condition is met.

This PROCAMS also may operate in either observation space or one of two transform spaces reached by either a linear combinations transform or a nonlinear transform. The possibility of simplifying subsequent processing is offered through these transforms.

A major feature of the system is the partitioning and signature extension approach which employs recognition on a non-training segment (we assume that most of the segments are of this type) using "labeled" field center clusters from that segment, a form of local recognition. The signature extension to a non-training segment within a partition (spectrally "similar" area) is accomplished by the labeling of clusters (obtained in an originally unsupervised manner) by a form of cluster matching with training segment clusters. The matching does not require all the same classes to be present in training and recognition segments; missing or added classes of non-interest can be accommodated, relaxing somewhat this unnecessary constraint on partitioning. This signature extension approach also is multitemporal and can use different training segments for different times simultaneously if needed to accommodate the low multiple acquisition probability on fixed location training segments within a partition. It is recognized that the error induced in the final acreage estimate due to signature extension must be less than the sampling error with only training segments being used.

Before describing PROCAMS operations and procedures, a general description of what component parts make up the system is presented. In Figures 1 and 2, the major component parts are represented. We begin with a partitioning of segments into groups that, within the same partition are "similar", but between partitions, are "dissimilar". We will not further address the partitioning component here, but believe a reasonable partition may usually be obtained in a timely manner. Since about 20 percent of the segments which are processed have some clouds, the next component is a cloud, cloud shadow, and bad-line detector which "excludes" those pixels from all steps until the proportion estimate and acreage computation. Because CITARS[2] showed the degradation on classification of differential haze effects, the next component is an atmospheric haze correction and normalization which normalizes the effects of differential haze over different segments to a uniform effect.

For the inventory of winter wheat, one of two paths can now be followed. If LANDSAT data has been acquired during each of biophases 1 and 4 as well as 2 or 3, the Delta Classifier[3],which is a regionally, prior trained, multitemporal per pixel classifier can be used. These biophases relate to the maturity of the crop with, ideally; (1) biophase 1 representing the condition where mostly bare soil is visible with, at most, only a little plant emergence; (2) biophase 2 showing significant plant emergence and growth but with only partial ground cover; (3) biophase 3 occuring at the peak in green development of the plant; and (4) biophase 4 occuring at or near harvest.

If the acquisition condition is not met, the other path must be followed. This path is also multitemporal, but can be unitemporal as well, giving estimates earlier in the growing season. If the acquisition condition is met, the choice between paths is not yet clear. The first component after the path option is a linear transform on the LANDSAT data channels. A dimensionality reduction may be achieved here, but the major reason for the transform, which is information preserving, is to put the discriminating information from the scene into 2 (or 3) linear combination channels, leaving out non-discriminating or confusing information[4,5]. The next component is a gradient-type boundary pixel identifier which eliminates from clustering (but not from classification) most of the non-field-center "mixed" pixels. Then, clustering of the field center pixels is accomplished. A capability for both unsupervised and supervised clustering modes is required. Then, the CROP-A signature extension module[6] is employed for cluster matching using an Ax+B affine transform. The next module is a reverse labeling, automatic procedure for enabling the classification of the recognition (non-local) segment to occur with labeled clusters from that recognition segment -- a type of local training. Finally, classification and proportion estimation and acreage calculations can be made.

The reverse labeling approach has the appealing feature of using locally (from the recognition segment) derived clusters for classification of the recognition segment such that the performance is expected to be as accurate as local training as long as the label extensions from the training segments do not provide label errors. Thus, the requirement for accurate cluster transformation is relaxed somewhat, but the burden is entirely on obtaining correct transformed cluster and label associations in the training segment.

PROCAMS, as presently implemented, includes procedures for locating and identifying training pixels but does not include registration overlay of multitemporal data which must be provided. Both of these steps are candidates for large improvement in efficiency and cost reduction in future operational systems.

PROCAMS OPERATIONS

PROCAMS can operate in the conventional mode of training on a portion of a scene and classifying the remainder of that single unitemporal scene. This conventional mode is referred to as "local recognition" and the data flow for local recognition is illustrated in Figure 3. Here we see the flow of the multispectral data (lines designated with a "d") and the flow of clusters (lines designated with a "c"). We also see that the cluster operation in this case is supervised since "ground truth" information is assumed to be available for local recognition scenes or sample segments. This is a relatively straightforward approach.

One of the prime goals of PROCAMS, however, is to provide the capability for handling multiple training sites or segments should they be available to improve the accuracy of signature extension by improving ones ability to label clusters from the signature extension or non-local recognition sites. By using multiple training segments, one could also overcome some of the difficulties limiting one from taking advantage of multitemporal data. Difficulties arise here since, due to cloud cover problems, a single training site in the same partition as a recognition site may not have data acquired on all of the same dates for both sites. The ability to use multiple training sites,each of which potentially matches only a subset of the dates available for the recognition site, then becomes quite important.

In any case, the system becomes much more complex in its necessary interactions to satisfy the above needs. This is obvious if one examines Figures 4, 5 and 6, which include the data and cluster flow for non-local recognition for multiple-training and multitemporal data sets with partially and fully matching biophase data acquisitions.

Consider processing the recognition segment data. After sequential processing through CLOUD, HAZE, and LINEAR TRANSFORM, subsets of channels are taken such that the subset represents those channels from one or more common biophase as the selected training segments. These subset channels then enter the processing chain of GRAD and unsupervised CLUSTER to provide clusters to CROP-A which also gets supervised clusters from the training segment for the subset of channels corresponding to that common biophase. CROP-A also gets another set of unsupervised clusters from the recognition segment to be transformed for use in the classification, tabulation, reverse labeling procedure. The classification is carried out on the training segment using the transformed clusters and the results are tabulated with tentative labels for recognition segment clusters. Then the results from other training segments and other times are combined into a final label determination. The recognition data is put into CLASSIFY with clusters derived from the recognition segment but labeled by the above procedure. TABULATION and PROPORTION follow.

CONCLUSION

As of this writing, the PROCAMS has been implemented and has been successfully exercized in its unitemporal local and non-local recognition modes. Efforts are now underway to gather and prepare the data necessary for the testing of the more complex aspects of PROCAMS. Specifications have been determined for a substantial data set which could serve in the long term as a classical test data set against which testing could be accomplished of PROCAMS and other agricultural processing systems including advanced signature extension algorithms .

The demonstration that such a system is more accurate and cost effectively so is a clear next step in our plans after generating some preliminary results on some Kansas LANDSAT data for wheat.

ACKNOWLEDGEMENTS

The ideas embodied in PROCAMS are in large part those of the personnel in the Multispectral Analysis Section at ERIM.

REFERENCES

1. MacDonald, R.B., F.G. Hall, and R.B. Erb, The Use of LANDSAT Data in a Large Area Crop Inventory Experiment (LACIE), Proceedings of the Tenth International Symspoium on Remote Sensing of the Environment, Ann Arbor, Michigan, October 1975.

2. Bizzell, R.M., F.G. Hall, A.H. Feiveson, M.E. Bauer, B.N, Davis, W.A. Malila and D.P. Rice, Results from the Crop Identification Technology Assessment for Remote Sensing (CITARS) Project, Proceedings of the Tenth International Symposium on Remote Sensing of the Environment, Ann Arbor, Michigan, October 1975.

3. Engval, J. and J. Tubbs, A Non-Parametric Approach to Classifying Remotely Sensed Data Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, Indiana, June 1976.

4. Kauth, R.J. and G.S. Thomas, The Tassled Cap -- A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as seen by LANDSAT, Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, Indiana, June 1976.

5. Crane, R., W. Richardson, R. Hieber and W. Malila, A Study of Techniques for Processing Multispectral Scanner Data, 31650-155-T, Willow Run Laboratories, The University of Michigan, Ann Arbor, Michigan, January 1973.

6. Lambeck. P.F. and D.P. Rice, Signature Extension Through the Application of Cluster Matching Algorithms to Determine Appropriate Signature Transformation, Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, Indiana, June 1976.

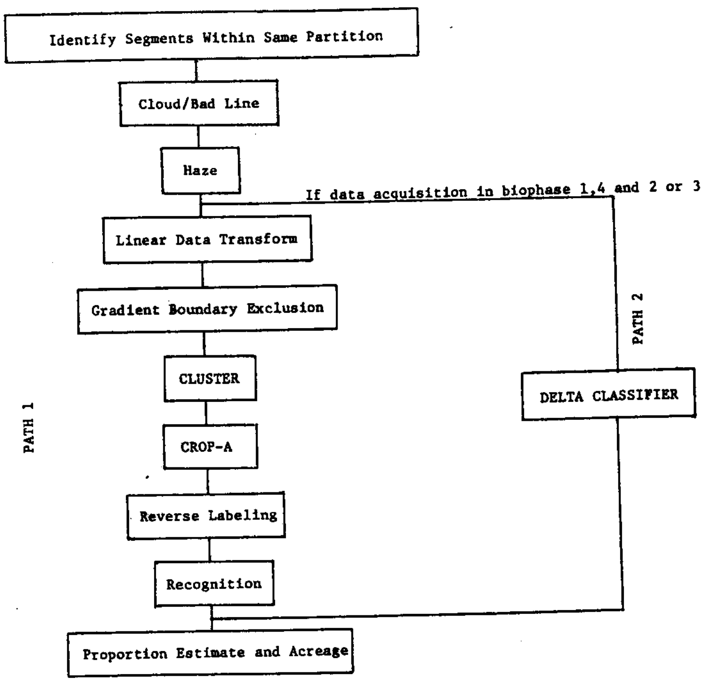

Figure 1. Components In Prototype CAMS System

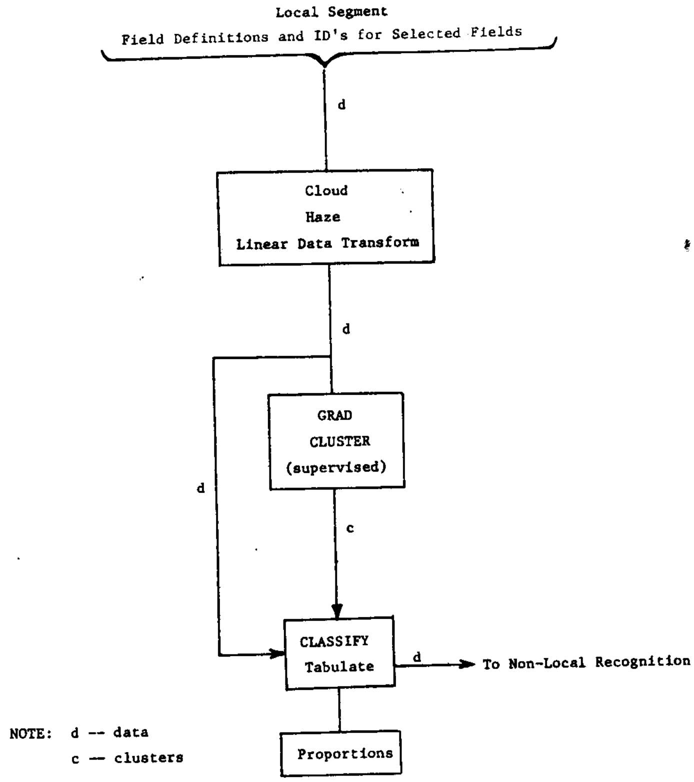

Figure 3. Prototype CAMS Procedure for Local Recognition

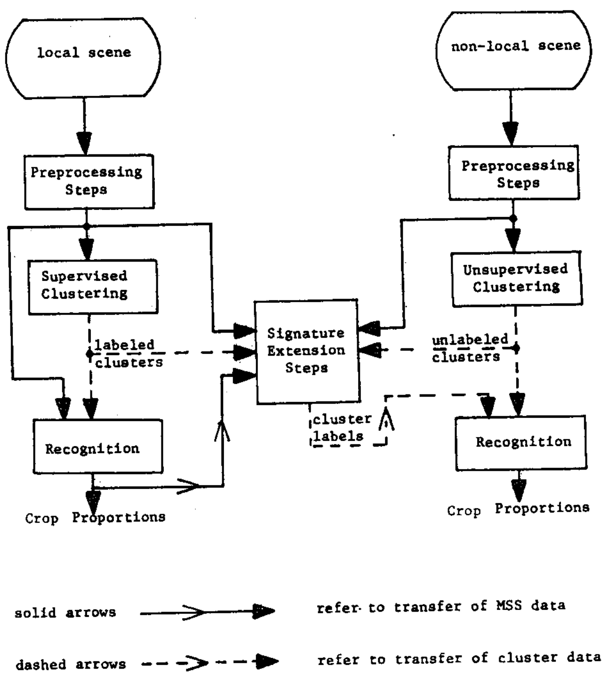

Figure 2. Overview Of The PROCAMS Data Processing System

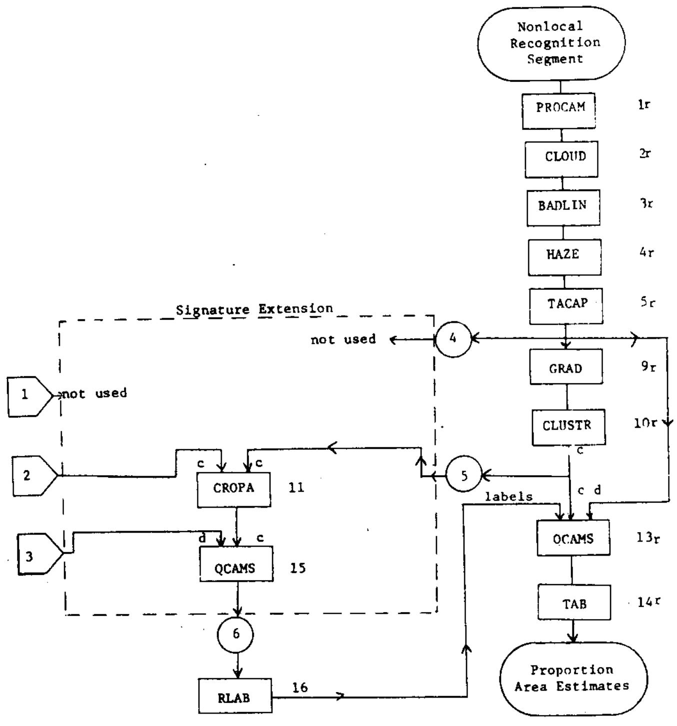

Figure 4. PROCAMS System - Nonlocal Recognition (Case 1: Fully Matching Biophases)

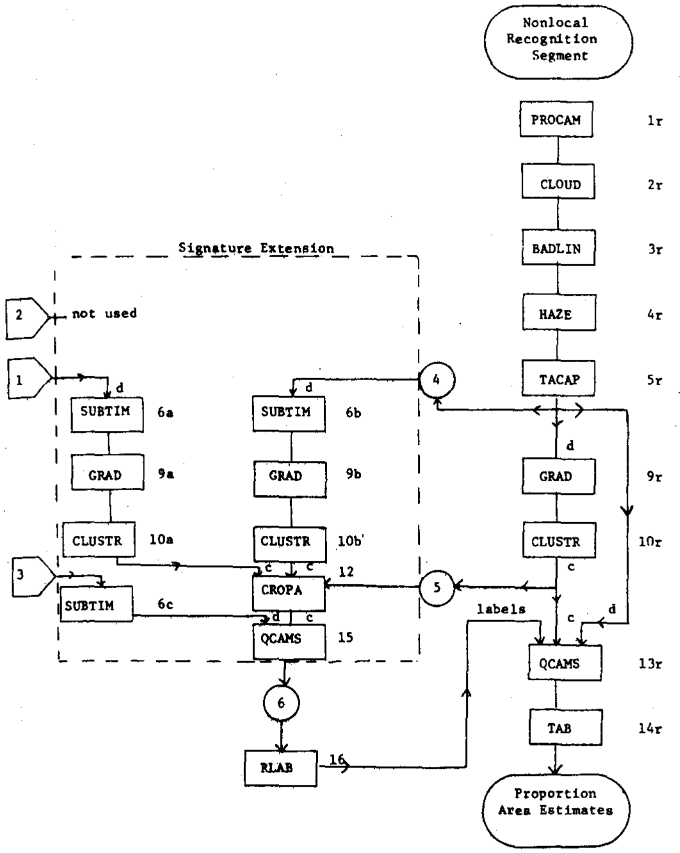

Figure 5. PROCAMS System - Nonlocal Recognition (Case 2: Paritally Matching Biophases)

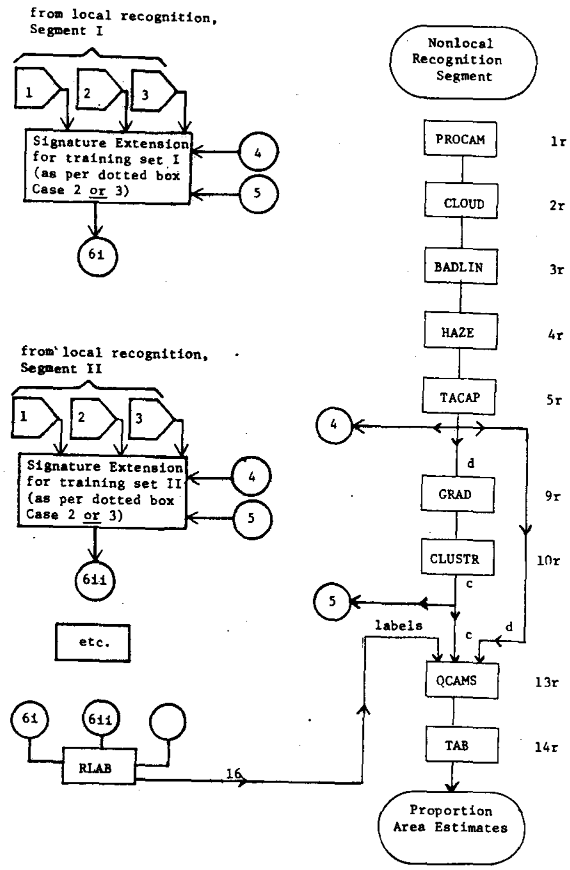

Figure 6. PROCAMS System - Nonlocal Recognition (Case 3: Multiple Training Segments)