Reprinted from

Seventh International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Range, Forest and Wetlands Assessment

June 23 - 26, 1981

Proceedings

Purdue University
The Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

IDENTIFICATION OF AGRICULTURAL CROPS BY COMPUTER PROCESSING IN THE PROVINCES OF CORDOBA AND LA PAMPA - ARGENTINA

MIRTA A. RAED, CECILIA ESPOZ

Comision Nacional de Investigaciones Espaciales
Vicente Lopez, Argentina

I. ABSTRACT

This paper deals with agricultural inventory parameters from LANDSAT digital data extracted via an interactive processing system.

In this study, a man-machine interactive processing system performed the analysis of LANDSAT digital data. Specifically, multispectral agricultural crop identification and spatial area determination, within the study areas, were carried out.

A temporal coordination between the multispectral LANDSAT satellite and the ground truth of the same area was attained, from which two test fields digitally classified, in accordance with crop species, crop varieties and soil types of the same place were chosen.

The results illustrate the importance of interactive processing for analyzing LANDSAT data. It must be pointed out that they do not represent the full potential of temporal information since they are preliminary results.

II. GEOGRAPHICAL SITUATION AND DESCRIPTION OF THE AREA

The study area is located within S034.00 and S035.00 parallels and W065.00 and W063.00 meridians and corresponds to scene N° 245.084 LANDSAT data. This area involves portions of three Argentine provinces:

1) South of the Province of Córdoba
2) North and East of the Province of La Pampa
3) North and West of the Province of Buenos Aires

The mean annual temperature fluctuates from 13° to 15° reaching 17° (upper end of scene), in the South of Córdoba. Annual rainfalls are lower than 600 to 800 mm, being more frequent in Autumn and Spring. There are risks of Winter and Summer droughts.

Soils have sandy texture, with low organic material contents. Because of this and due to strong winds there is danger of eolic erosion in exposed lands (without vegetation).

The activities carried out in this area are agriculture without irrigation and cattle raising.

The most common winter crops are:

1) Wheat: In order to avoid winter droughts those varieties having a short growing season are sown. Seeding date is July-August, and harvest is carried out at the end of December.
2) Rye: It is very common since it has a deep root system which makes it drought-resistant. It is used for pasturing and it is rotated with alfalfa. Rye is sown at the end of February.

With reference to summer crops the range of possibilities is wider:

1) Sorghum: Date of sowing and variety are varied in order to avoid a possible drought during flowering time.
2) Corn: It is not very common due to the damage it suffers because of late frosts and water shortage during flowering time.
3) Sunflower: It is a drought-resistant crop. It evolves well in this area since soils are sandy and there is no possibility of diseases.
4) Soybean: It is a common crop, but during the last two years, due to market reasons, producers have changed to other types of crops. It is sown in November, and harvested in March.

Here, a brief general description has been carried out of the study area, based upon bibliographical and statistical information from governmental agencies. Along the study more details will be established regarding the area and the corresponding crops.

III. STATISTICAL PROPERTIES

In an analysis effort with these data, a maximum likelihood technique is used in an attempt to identify the different types of crops.

Ground truth information is provided initially for the training samples used to estimate the statistical characteristics for each class. Univariate histograms are compiled for each class.

During the preliminary phase of study, the conditional probability density functions of the feature measurements for each class by multivariate gaussian density functions will be approximated. The mean vectors and covariances matrices for each class are estimated by the sample means and the sample covariances calculated from training samples.

From the gaussian assumption, the following mathematical formulation can be made. For n pattern classes (n kind of agricultural crops) ω_1, ω_2, ... ω_n, the feature measurement vectors, X, for each class are distributed according to a multivariate gaussian density function, i.e.,

$$p\ (X/\omega_i) = \frac{1}{(2\pi)^{N/2}\ |K_i|^{1/2}}\ \exp\ [-\frac{1}{2}\ (X - M_i)^T\ K_i^{-1}\ (X - M_i)]$$

$$i = 1,\ 2,\ \ldots n$$

Where X is an N-dimensional vector ($N = 12$), M_i and K_i are the mean vector and covariance matrix for the i^{th} class, ω_i, respectively.

Based upon the above formulation, the classification task can be performed by applying the maximum likelihood classification rule.

IV. ANALYSIS OF AGRICULTURE DATA

The LANDSAT MSS data, processed through computer, were quantitatively evaluated, with the support of ground truth data, using a software system called ERMAN II (Earth Management System).

The ERMAN II system is based on NASA's ERIPS (Earth Resources Interactive Processing System) and consists of a large set of software programs which perform a wide variety of functions related to digital image processing, and a display system for interacting with the user. It runs on an IBM/360 or IBM/370 computer, with at least 512 K bytes of real storage.

The system was designed to analyse data from various imaging sensors including LANDSAT data and an optical-mechanical 12-channel scanner.

Data provided by LANDSAT satellites can be directly used in its 4 channels, or else, used to make band ratios for determining the best sets of 4-bands each.

Within the Laboulage area two test fields were located: MELO and REYNAL, with 10,246 and 11,373 samples, respectively.

The LANDSAT data analysed were scene N°245-084 from January 8, 1981.

The study of Melo's area was carried out with ground truth support which consisted of identification of the crop species: sunflower, sorghum, millet, sown pastures, soil, stubbles, alfalfa, in the four LANDSAT bands. Figure 1.

Table 1 show 10,246 samples, indicating how many of them correspond to each class, with 1.0 and 15 threshold values. It can be observed that the unclassified samples directly increase with the threshold number.

The different kind of pastures were mixed, and alfalfa was not distinguishable from soil and pastures.

When a classification map was obtained, there was an unclassified class; that is to say, if the value of the discriminant computed $g(X)$ function assigned to this sample is less than some threshold value, then a rejection class is formed when a sample is not classified into any of the considered classes.

Then, mathematically, a sample X is classified as from class ω_i if,

(1) $g_i\ (x)\ >\ g_j\ (x)$ for all $j \neq i$

(2) $g_i\ (x)\ \geqslant\ T_i$

Where, T_i is the threshold for the class ω_i.

It was difficult to find the best threshold for each class, so that most of the known samples fell into the correct class.

In order to improve the determination of statistical separability of multispectral measurements from agricultural cover types, band ratios for determining the best 4 bands were made.

It was necessary to study the subset selection of feature measurements from the complete set once the sixteen feature measurements were found to differentiate, in a better way, the kinds of crops. Table 2.

Divergence is defined for any of two density functions. In the case of normal variables which have unequal covariance matrices, the divergence in n spectral channels c_1, c_2, ...c_n is given by the following formula:

$$D\ (ij/c_1,\ c_2,\ \ldots c_n) = \frac{1}{2}\, tr\, [\, (\Sigma_i - \Sigma_j)(\Sigma_j^{-1} - \Sigma_i^{-1})\,] + \frac{1}{2}\, tr\, [\, (\Sigma_i^{-1} + \Sigma_j^{-1})(U_i - U_j)(U_i - U_j)^T]$$

Where U and Σ represent the mean vector and covariance matrix, respectively; A (trace A) is the sum of the diagonal elements of . Although divergence only provides a measure of the distance between two class densities, the average overall pair classes can be taken. So, the subset of features could be selected for which the average divergence was maximum or, to maximize the minimum divergence to select the feature combination which provides the greatest separation between pair of classes.

As in the first classification carried out, alfalfa appeared mixed with pastures. Therefore, it was not considered as a different class in this new image ratio classification.

Through this type of analysis, soil and stubble appear intermixed. On the other hand, sown pastures and sorghum are distinguishable in many areas. Figure 2.

Through the 11,373 samples, it was observed how classified crop samples vary with threshold values. Table 3.

More unclassified samples appeared among pastures, stubbles, and soil, perhaps, because, there were other kinds of vegetations which were not taken into account.

But the most important fact of this classification is that the three kinds of crops: sorghum, sunflower and millet are well classified.

Afterwards, another Laboulage subset named REYNAL was studied.

In this test field a good separation among training fields was observed. As the boundary among fields was very clear and besides, these were more homogenous fields than in MELO subset, a decision was taken to study the kind of crops in the four original LANDSAT bands.

Through ground truth different pastures and soil areas were known, so several training fields of the same class were averaged and then, the test field was classified with these statistical data.

When the classification map, having a treshold value of 1.0, is observed sunflower appears well delineated, but there were problems with sorghum fields.

In order to improve this classification, several sown pasture samples were taken and a better division among the reflectance values of recently sown pastures and bare soil were obtained. Figure 3.

When crop histograms were compared, i.e., sorghum, millet and sunflower, it was possible to distinguish them in the four LANDSAT bands and they had unimodal normal distributions. So these three kinds of crops were quite different from each other.

It must be pointed out that when several training fields of the same class were computed, there appeared less unclassified class samples.

It was impossible to differentiate sown pastures from natural pastures.

Besides, a problem arose between statistically classified sorghum and ground truth.

In order to improve the classification, two sorghum training fields, two sown pasture training fields and a big area like natural pastures were chosen.

The same sorghum areas. as in the former classification, were obtained. It was observed that the forest was well delineated and it was differentiated from natural and sown pastures in each classification.

One of the conclusions was that sorghum reflectances, in some areas, were similar to natural and sown pasture reflectances.

A comparison of the different crop classes of one test field with the other one was carried out in order to finish with the preliminary work.

Training fields of the three kind of crops of the test field MELO were considered and with these data, REYNAL test field was classified. Sunflower and millet areas appeared with the same distribution in both classifications.

It should be observed that millet, through ground truth, is in different growing states but the reflection in each area is the same.

With regard to sorghum areas, the identification of them is not so clear as it happens with the other crops.

Then, MELO test field was classified with REYNAL training fields.

In this classification, the number of sunflower samples was less than the number of the initial crop classification.

Sorghum areas were not well delineated, though, through ground truth they were in the same growing state. It must be noticed that, at that time, fields underwent water excess (inundations) and that perhaps this was the reason which led to the mentioned missclassification.

V. CONCLUSIONS

The difference which appeared in the classification accuracy is probably due to the fact that training samples used were not completely representative of all variations of multispectral response patterns of crop species and the number of training samples is inadequate.

It is necessary to find better and more efficient means of collecting ground truth data. Variations in spectral patterns caused by the different stages of crop maturity, variety differences, moisture conditions, soil temperature, water content, and other parameters must be thoroughly investigated.

Therefore, computer analysis of LANDSAT MSS data is an effective method in Argentina for identifying agricultural crops. This capability should lead to improvements in precision and timeliness of crop production estimates.

VI. REFERENCES

1. Kettig, R.L.; Landgrebe, D.A.; Classification of Multispectral Image Data by Extraction and Classification of Homogenous Objects. LARS Information Note 062375 - Purdue University, West Lafayette, Indiana, 1975.

2. Todd, W.J.; Baumgardner, M.F.; Land Use Classification of Marion County, Indiana, by Spectral Analysis of Digitized Satellite Data LARS Information Note 101673. Purdue University, West Lafayette, Indiana, 1973.

3. Kumar, R.; Silva, L.; Statistical Separability of Spectral Classes of Blighted Corn. LARS Information Note 041973. Purdue University, West Lafayette, Indiana.

4. Swain, P.H.; Pattern Recognition: A Basis for Remote Sensing Data Analysis. LARS Information Note 111572, Purdue University, West Lafayette, Indiana, 1973.

5. Landgrebe, D.A.; Min P.J.; Swain, P.H.; Fu, K.S.; The Application of Pattern Recognition Techniques to a Remote Sensing Problem. LARS Information Note 080568. Purdue University, West Lafayette, Indiana.

6. Kumar, R., Silva, L.F.; Separability of Agricultural Cover Types by Remote Sensing in the Visible and Infrared Wavelength Regions. LARS Information Note 032977. Purdue University, West Lafayette, Indiana, 1977.

7. Kumar, R.; Silva, L.F.; Statistical Separability of Agricultural Cover Types in Subsets of One to Twelve Spectral Channels. LARS Information Note 041974. Purdue University, West Lafayette, Indiana, 1974.

8. Bauer, M.E.; Cipra, J.E.; Identification of Agricultural Crops by Computer Processing of ERTS MSS Data. LARS Information Note 030173. Purdue University, West Lafayette, Indiana, 1973.

VII. ACKNOWLEDGEMENT

The authors are most grateful to Dr. Adolfo D'Onofrio from the IBM Scientific Center for his significant assistance and guidance, and to Eng. Agr. J.E. Lucesole, from the Ministerio de Agricultura y Ganadería (Ministry of Agriculture and Livestock) for his contributions to this paper.

Table 1. Number of Samples for Each Class with Different Thresholds.

Training Fields

Field Name-	Sunflower	Soil	Rest	Sorghum	Pasture	Millet	Millet 2	Sorghum 2	Alfalfa	Stubble	Thresh
Class Name-	GI1	SU1	R1	SO1	LL1	MI1	MI2	SF1	AL1	RA1	
	226	333	1058	780	584	1900	1849	1189	450	373	1504
	150*	223*	720*	547*	402*	1344*	1086*	985*	305*	306*	4178*

MELO Test Field, N° of samples: 10,246

Note: (*) Percent 15
Percent 1.0

Table 2. Computed Best Channel Sets.

1.	1	10	13	14
2.	5	9	13	15
3.	5	11	13	15
4.	5	13	15	16
5.	5	7	13	15
6.	5	7	9	11
7.	1	9	13	14
8.	1	6	13	15
9.	1	13	14	16
10.	5	9	11	12
11.	5	8	9	11
12.	7	8	9	11
13.	1	13	14	15
14.	1	2	13	14
15.	5	7	11	12
16.	1	3	13	14
17.	6	7	9	11
18.	7	9	11	12
19.	5	6	7	11
20.	5	6	9	11
21.	5	7	8	11
22.	5	6	13	15
23.	1	5	13	14
24.	5	9	15	16
25.	4	5	13	15
26.	5	9	11	15
27.	1	6	14	15
28.	7	11	13	16
29.	1	7	9	11
30.	5	9	11	13

Table 3. Number of Samples for Each Class with Different Thresholds

Field Name-	Sunflower	Soil	Pasture	Millet	Sorghum	Stubble	Thresh
Class Name-	Gira	SU	PA	MI	SO	RA	
	453	864	1750	1676	2907	639	3448
	1014*	1292*	2612*	2358*	2775*	686*	0*

REYNAL Test Field, N° of samples: 11,373

(*) Threshold .1
Threshold 1.0

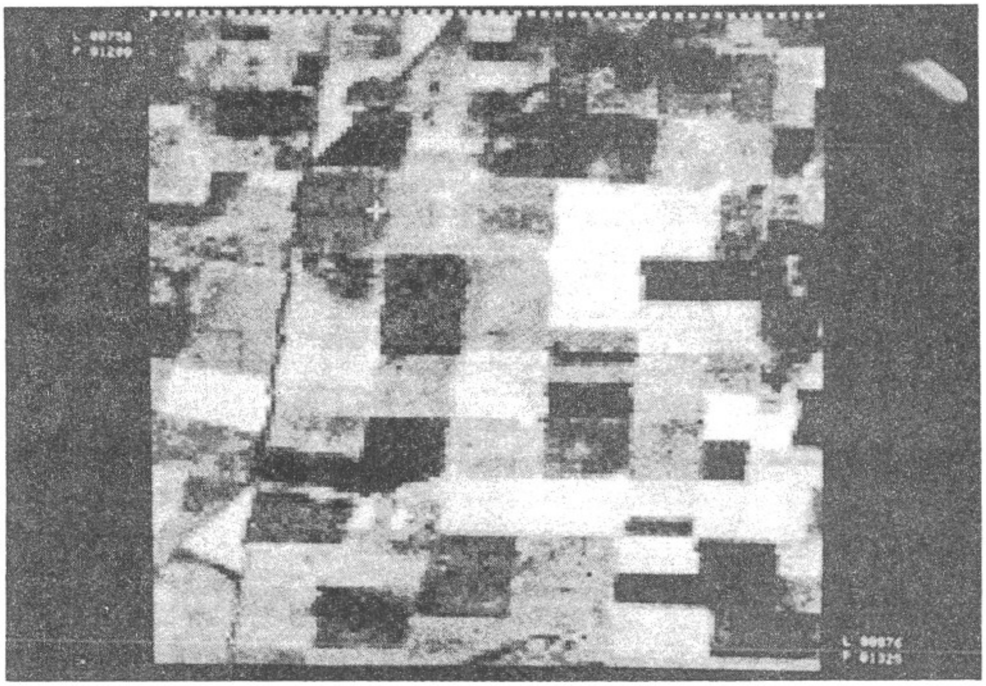

Figure 1. Image from MELO Test Field

Figure 2. Classification Character Map N°2

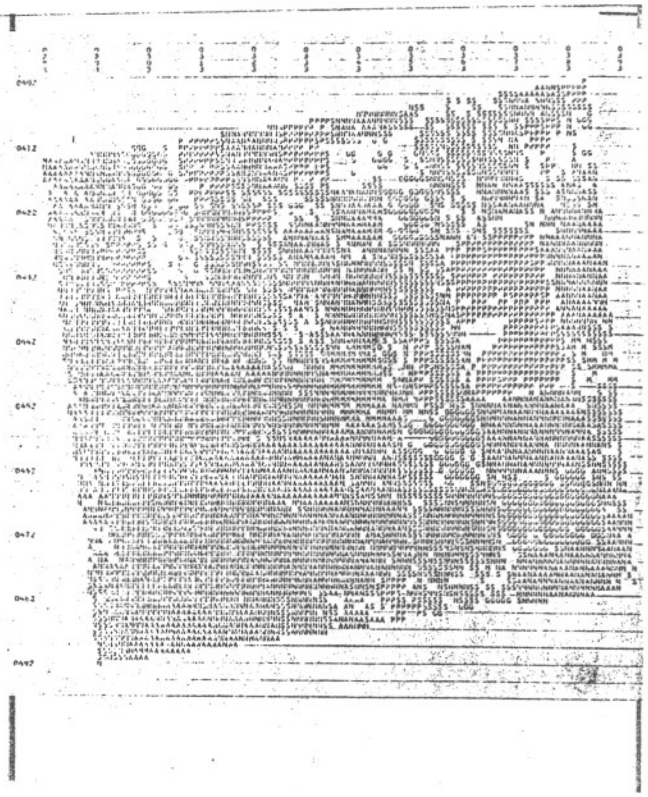

Figure 3. Classification Character Map N°3

AUTHOR BIOGRAPHICAL DATA

Mirta A. Raed was born in Buenos Aires, Argentina. She received the degree of Licentiate in Physics from the National University of Buenos Aires, in 1975. For two years, she worked in the Remote Sensing Group of the Ministerio de Agricultura y Ganadería (Ministry of Agriculture and Livestock) of Argentina. She attended the International Symposium on Remote Sensing and subsequent Seminar, held in La Paz, Bolivia, between November 28 and December 8, 1977. This Symposium was sponsored by the United Nations and the Organization of American States. At present, she is a staff member of the Comisión Nacional de Investigaciones Espaciales of Argentina.

Cecilia Espoz. Agronomist, graduated from the University of Buenos Aires. She assists teaching at the Faculty of Agronomy. At present, she is a staff member of the Comisión Nacional de Investigaciones Espaciales of Argentina. She has carried out Courses on Agrometeorology in Canada, as member of the joint Project between CNIE and PNUD (Crop Evaluation).