

1-1-1981

Computer Generated Maps from Digital Satellite Data: A Case Study in Florida

L. G. Arvanitis

R. M. Reich

R. Newburne

Follow this and additional works at: http://docs.lib.purdue.edu/lars_symp

Arvanitis, L. G.; Reich, R. M.; and Newburne, R., "Computer Generated Maps from Digital Satellite Data: A Case Study in Florida" (1981). *LARS Symposia*. Paper 478.
http://docs.lib.purdue.edu/lars_symp/478

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

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

Copyright © 1981

by Purdue Research Foundation, West Lafayette, Indiana 47907. All Rights Reserved.

This paper is provided for personal educational use only,
under permission from Purdue Research Foundation.

Purdue Research Foundation

COMPUTER GENERATED MAPS FROM DIGITAL SATELLITE DATA: A CASE STUDY IN FLORIDA

L.G. ARVANITIS, R.M. REICH

University of Florida
Gainesville, Florida

R. NEWBURNE

Florida Division of Forestry
Environmental Resources Management
Miami, Florida

I. ABSTRACT

The availability of satellite data and the advantages offered by automatic machine-processing of such data have opened up new and exciting possibilities for developing ground cover maps. Two Landsat analyses techniques (an unsupervised clustering algorithm called Landsat Signature Development Program, and an interactive method based on the Multispectral Image Analyzer) are used to compare computer-generated character maps to known earth-surface features. Data samples are shown and applications are discussed. Reference is made to the value of the digital computer in natural and man-made features mapping and monitoring, and suggestions are given for further research.

Key Words: Multispectral sensing, Landsat Signature Development Program (LSDP), Multispectral Image Analyzer (Image 100).

II. INTRODUCTION

Ground cover maps are important tools to a wide array of users. Rational management of land resources in particular requires an accurate assessment of the existing resource profile and respective changes over time. The need to establish compatible regional land use/land cover information systems is underlined by planning agencies of the public and private sectors as conflicting uses of land and water intensifies.

Much progress has been made over the past three decades in supplementing planimetric and topographic maps with ground cover details obtained from aerial photographs. Several approaches have already been employed in machine recognition of spatial patterns and automatic display of ground features with minimal human intervention. Photo interpretation is tedious, time consuming, and overall a costly process which at best reflects the degree of

expertise and qualitative judgement of the individual photo interpreter. Hence updating land use/land cover maps at frequent time intervals as needed is not always feasible.

This limited study, supported jointly by NASA-Kennedy Space Center and the School of Forest Resources and Conservation, University of Florida, evaluates the feasibility of using computer maps of ground cover from satellite input tapes.

III. STUDY PROCEDURE

A. SELECTION OF TEST SITES

The following criteria were used in selecting study areas (Fig. 1):

- . Budget and time constraints
- . Diversity of ground cover conditions
- . Advance knowledge of the study areas
- . Availability of Landsat input data
- . Availability of recent aerial photography.

One test site was located in Alachua, Bradford, and Union Counties of north central Florida covering about 21 x 21 miles. Ground cover conditions are characterized by deciduous and non-deciduous hardwood forests, mixed softwoods-hardwoods, natural pine stands, pine plantations of various ages, grazing lands, cultivated fields, rivers, lakes, small towns, and scattered residential areas.

The second site was located in southwestern Florida near the City of Fort Myers and covers part of Lee County; approximately 20 miles in an east-west direction and 23 miles in a north-south direction. The main ground cover features include mixed hardwood and softwood forests, cultivated and open uncultivated fields, residential areas, a portion of the City and the bay of Fort Myers, mining pits (some filled

with water), and a section of the Caloosahatchee River.

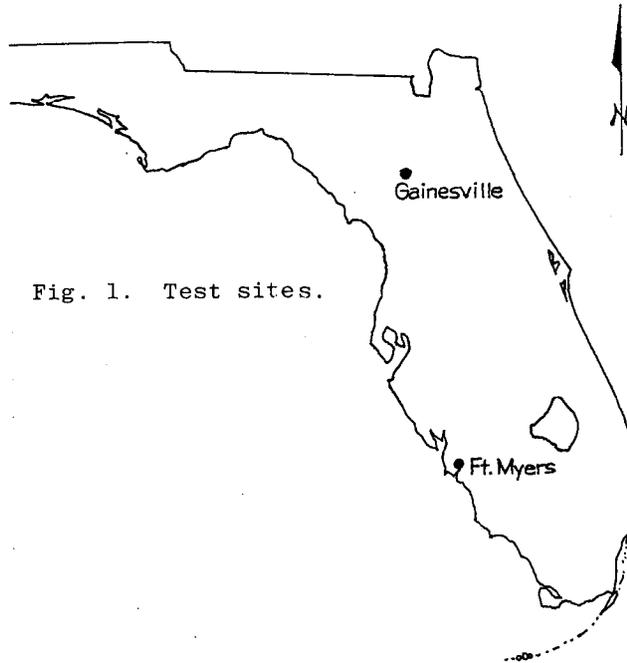


Fig. 1. Test sites.

B. LANDSAT INPUT DATA

For this study, the following satellite input data was used:

Fort Myers Test Site -

- March 4, 1975 - Landsat Scene Identification No. 2041-15174.
- February 21, 1977 - Landsat Scene Identification No. 20761-15023.

Gainesville Test Site -

- April 17, 1977 - Landsat Scene Identification No. 20816-15024.
- October 14, 1977 - Landsat Scene Identification No. 20996-14544.

The dates selected were dictated to some extent by the availability of raw data and the need to evaluate possible changes over a short time interval (1975 to 1977), as well as within-year seasonal variation (April vs. October 1977).

C. SATELLITE DATA PROCESSING SYSTEM

Presently users of the satellite data on the Kennedy Space Center (KSC) Applications Projects Branch employ several Landsat analyses techniques, two of which were evaluated in this study: The unsupervised clustering algorithm, called Landsat Sig-

nature Development Program (LSDP), and the interactive one based on the Multispectral Image Analyzer (Image 100). The LSDP and three companion programs, the Landsat Geometric Correction Program (LGCP), the Landsat Signature Comparison Program (LSCP), and the Landsat Classification and Mapping Program (LCMP) are written in FORTRAN V⁷.

D. MULTISPECTRAL IMAGE ANALYZER (IMAGE 100)

The Image 100 is designed to accommodate data in the format received from the Landsat input tapes. It enables users to interact with the data on a real-time basis. By training on small ground samples of known characteristics, all other areas of a given Landsat scene with a similar signature may be displayed on a color CRT within seconds. Up to eight themes of the same scene can be displayed simultaneously. Through a suitably scaled Gould line printer, character maps may be subsequently produced to closely approximate the 1:24000 scale of the U.S. Geological Survey (USGS) quadrangle sheets used in this study.

E. GENERAL PURPOSE COMPUTER-GENERATED MAPS

LSDP 1:24,000 computer maps of the four Landsat scenes were produced by the Honeywell 635 computer at KSC. The programs were run at several chi-square confidence levels which control the number of resultant clusters. The cluster statistics cannot change more than the selected chi-square value will allow. Thus, a confidence level of 95 percent would produce more clusters than the confidence level of 99 percent. From preliminary results, it was concluded that the 98- and 99 percent confidence levels produced the most useful number of classes for the test areas. Each of these character maps covers an area of 520 x 520 pixels (465 square miles). For an area of about 130 x 130 pixels (29 square miles), LSDP maps were also produced at the 98 percent level of confidence for both sites and dates. The purpose was to find out whether computer-generated maps of smaller areas provide a better proximity to actual ground conditions -- due to smaller variations in spectral reflectance -- than those covering relatively larger areas.

LCMP 1:24,000 maps were also produced at confidence levels corresponding to LSDP maps, but only those at the 95% confidence level were used in this study. LCMP receives its input from LSDP generated data and can improve the cluster statistic before final mapping, thus producing somewhat more accurate maps, than those of the LSDP programs.

The most promising computer maps for both test sites were subsequently compared against corresponding USGS maps and aerial vertical photographs of the same 1:24,000 scale.

F. PRELIMINARY EVALUATION OF COMPUTER MAPS

As an initial step, the LSDP, LCMP, and the Gould maps for both sites were overlaid on 1:24,000 U.S. Geological Service quadrangle maps (1966 Edition). Characteristic ground features such as lakes, rivers, highways, roads, and coast lines from the USGS maps were used to establish reference points on the computer-generated maps.

The Austin Cary Forest (ACF) and the Beef Research Unit (BRU) of the University of Florida in the Gainesville test site were selected for preliminary evaluation of the computer maps. The ACF includes natural and planted pine stands, bottomland hardwoods, cypress, and recently logged-planted areas. The BRU has mainly grazing lands and cultivated fields (light and dark tone). Some tree islands and cypress domes are also present. Vertical 1:10000 black and white aerial panchromatic photos taken on 10/5/77 were available for preliminary field and laboratory work.

In the Fort Myers test site, a sample area was selected within another intensive remote sensing study¹. As a reference base, we have used black and white 1:24,000 aerial photographs as well as color infrared transparencies taken in 1978. This sample area includes forest areas, open cultivated or uncultivated fields, a river, small ponds, a bay, and scattered houses.

In comparing the computer-generated maps to the aerial photos, attention was given to determining whether the maps could depict specific ground features. Such features included forests (hardwoods, softwoods, mixed), cultivated fields, grazing lands, uncultivated open fields, recently logged and/or planted parcels, large bodies of fresh and salt water, rivers, as well as residential/industrial area.

G. TEST FOR AREAL CORRESPONDENCE

Two sample areas, representing a wide range of ground cover classifications, and covering approximately 4,200 acres each, were selected from the Gainesville test site. The objective was to evaluate the one-to-one areal correspondence of major ground cover categories, as outlined on the aerial photographs, and the computer-generated map. A grid of 1,050 square plots -- each covering 4.02 acres -- was superimposed on each of the four overlays

developed from the aerial photographs and the respective computer-generated maps.

A simple random sample of 100 square plots was selected without replacement to estimate the one-to-one areal correspondence between aerial photographs and the machine processed maps. For similar comparisons, Ginevan⁴ suggests use of acceptance sampling in evaluating the accuracy of computer-processed land cover maps. Basically, this approach deals with the determination of "optimal" number of ground truth samples and the "allowable" number of misclassifications of these samples.

From the results of this comparative study it became obvious that features covering small ground areas, such as roads, narrow rivers, clusters of houses, ponds, and the like, are obscured by the edge effect of the surrounding dissimilar areas (Figs. 2, 3 and 4).

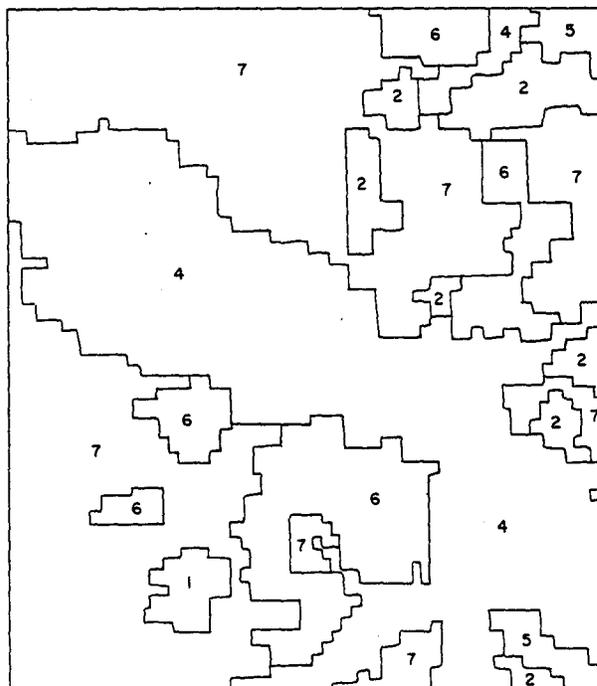


Fig. 2. A section of the April 17, 1977, LCMP computer map showing the Santa Fe River, Gainesville test site. (1-fresh water, 2-softwood forests, 3-cypress, 4-mixed hardwoods, 5-dark tone uncultivated fields, 6-dark tone cultivated fields, 7-light tone cultivated fields, 8-residential/commercial areas, 9-young pine plantations.)

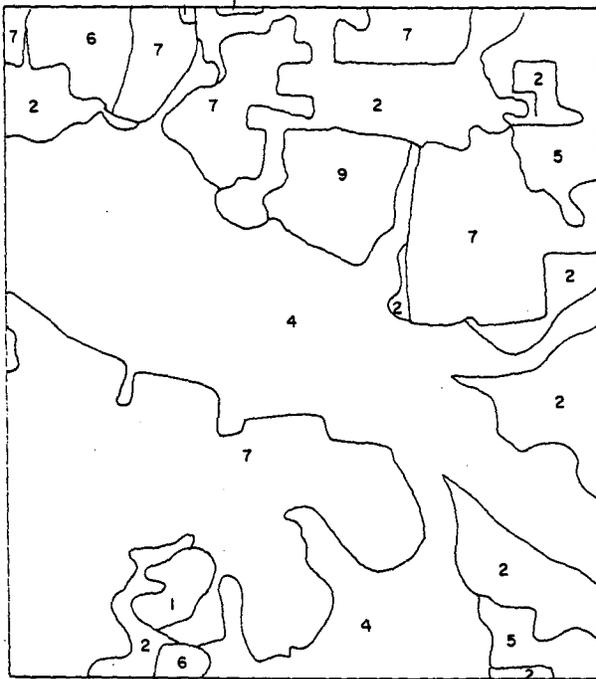


Fig. 3. A section of a Mark Hurd aerial photograph showing the Santa Fe River. The numbers refer to the classifications of Fig. 2.

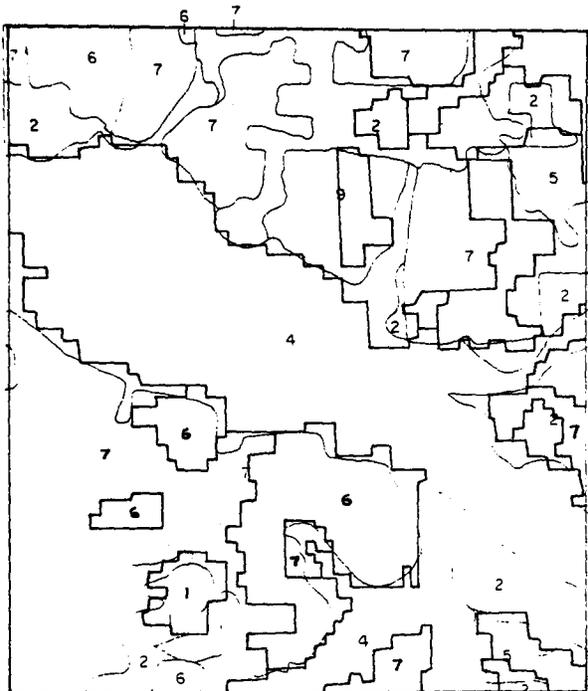


Fig. 4. Figures 2 and 3 superimposed.

The resolution of computer-generated maps diminishes beyond a certain point. However, overall we were convinced that comparing acreages by categories, as depicted by the various computer-generated maps and those delineated on aerial photographs for the same scene, may provide an insight into the capabilities and limitations of the Landsat maps.

H. PREPARATION OF OVERLAYS FROM AERIAL PHOTOGRAPHS

For the set of selected computer maps of both test sites and dates, overlays have been prepared on frosted acetate. One larger area and a smaller one were selected for detailed acreage estimation. The objective was to determine whether the size of an area affects the overall acreage estimation by categories. Because of budget constraints, the Image 100 maps were evaluated only for the smaller size areas. For the Gainesville test site, the two areas selected for a detailed evaluation were about 95,000 acres and 23,000 acres, respectively. For the Fort Myers test site, the areas were approximately 68,000 and 25,000 acres, respectively.

I. REFERENCE DATA

A sampling scheme was employed to collect reference data from the aerial photographs that would enable use to identify major ground cover types on the Landsat computer maps. A grid of 3,380 plots -- 1 square inch in size, each representing 88 acres -- was superimposed on the aerial photos and the selected Landsat computer maps. A 10 percent sample, or 338 plots, were systematically selected for evaluation. On each of the 338 plots on the aerial photos, ground cover types were recorded along with the corresponding character elements on the computer maps. This information was then used to identify the major ground cover types on the Landsat computer maps.

During this evaluation process those sections of the computer-processed maps which appeared to deviate considerably from the photointerpretation results were marked and verified in the field. Subsequently 34 plots -- 12 in the Gainesville test site and 22 in the Fort Myers test site -- were identified for field verification of the actual ground features. When applicable, data were collected on tree size, soil color (light or dark tone), and understory species. In the Gainesville test site the major differences were due to logging operations of forest areas. In Fort Myers differences were attributed primarily to the expansion of the industrial, commercial, and residential areas.

J. PREPARATION OF OVERLAYS AND ACREAGE ESTIMATION OF LAND COVER TYPES ON THE LANDSAT COMPUTER MAPS

Overlays were also prepared for the individual Landsat computer maps based on the key developed previously from the reference data. Delineation of boundary lines and preparation of overlays for the Gould printer maps were made by the same person who themed the various ground cover categories. Acreages for each classification on the computer map overlays were estimated by counting the number of pixels for each ground cover category (1 pixel x 1.1 acres).

IV. RESULTS

This study^{2,3} has been limited both in scope and availability of resources. Rather than attempt to extrapolate the findings, we believe some of our observations warrant consideration for further study:

1) Several ground cover categories of the computer-generated maps (LSDP, LCMP, Image 100) evaluated in this study provided highly accurate results which could be used effectively on a large scale basis (Fig. 5).



Fig. 5. Cypress mixed with pine and hardwoods. LCMP computer map. Input data 4/17/77. Gainesville test site.

2) With one exception, results from the Gainesville test site were more satisfactory than those of the Fort Myers test site. This outcome may be attributed to the highly diversified ecological conditions, and thus, to the wider range of spectral response patterns of the Fort Myers test site as compared to those of the Gainesville site.

3) For the Fort Myers test site the 2/21/77 LCMP maps provided a separate classification for salt water (Fig. 6). This rather rare coincidence may be attributed to the wave motion at that particu-

lar time, and/or the turbidity of the merging water from Caloosahatchee River. It is known that suspended organic and inorganic materials in water bodies cause scattering and absorption of incident energy, thus affecting the spectral reflectance which is detected by Landsat⁶.

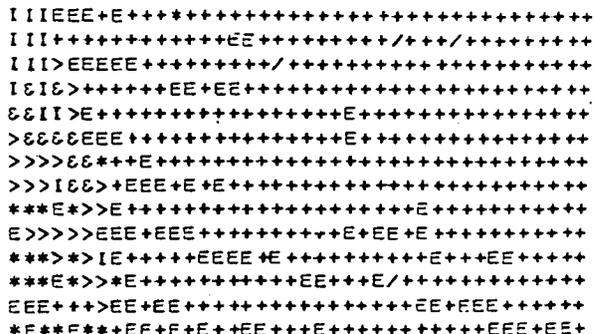


Fig. 6. Salt water. LCMP computer map. Input data 2/21/77. Fort Myers test site.

4) In the process of evaluating the various LSDP, LCMP, and Image 100 maps, difficulties were encountered in superimposing the computer-generated maps onto vertical aerial photographs and the U.S.G.S. 7-1/2 min. quadrangle sheets. Although the LSDP, LCMP, and Image 100 maps are supposed to be the same scale as the USGS ones (1:24,000), there were differences in the north-south direction due mainly to line printing and the size of individual character elements. These differences introduced problems in field orientation and area estimation by ground cover categories which must be properly corrected. A better procedure has been developed at the KSC-Applications Projects Branch after our study was concluded which allows corrections with ground reference data and produces improved LCMP classifications.

5) The exact location of specific ground features, such as small residential areas, roads, small rivers, and lakes, cannot be determined from any of the computer-generated maps in this study due to edge effect. Such features are classified in one of the surrounding cover categories.

6) Overall, computer-generated maps for relatively large areas (520 x 520 pixels) have produced better results in this study than maps covering smaller areas (such as 130 x 130 pixels).

7) Along transition zones of such ground features as shorelines, lakes, and ponds, the areas are usually left unclassified in the computer-generated maps due to noise or edge effect. As a result, locat-

ing exact boundary lines on the maps becomes a very difficult task.

8) Although specific pixel character elements of the computer-generated maps represent in some cases certain ground features such as forests, cultivated fields, open uncultivated fields, etc., the overall use of the same symbol is not consistent in a given map (Fig. 7). The spatial pattern of the specific ground cover mosaic and the reflectance from surrounding areas seem to affect the use of alternative mapping characters to denote the same ground surface features.

```

//s/++++/+++++// //+++++
////////+++++//++ +++++++
////////+++++//s+++++//s+//
////+ss//+++++//s+++++//s$$$+
////////+++++//+++++//s$$$s$+
++++//+++++//+++++//+++++
+++//++//+++++//s//+++++
++++//+++++//+++++//s//+++++
////+//+//+++++//s//+++++
////s//+++++//s//+++++//+++++
////+//+++++//s$$$+++++//+++++
////////+++++//s$$$+++++//+++++
////////+++++//+++++//s//+++++
////////+++++//+++++//s//+++++
++++//+++++//+++++//+++++

```

Fig. 7. Open uncultivated fields (natural grasses, palmetto, scattered dense patches of trees) or young pine plantations. LSCP computer map. Input data 10/14/77. Gainesville test site.

9) Successful themeing of Landsat scenes on the Image 100 depends heavily on firsthand knowledge of ground cover conditions and the ability to locate specific features on Landsat input tapes as displayed on the console screen. Usually, areas with smaller ecological diversities can be more easily themed on the interactive Image 100 than those characterized by heterogeneous conditions.

10) In all computer maps, sites, and dates examined, the best results were achieved when the classification was limited to only land and water (Fig. 8). Even with three cover categories (water, forests, open fields) the 10/14/77 Image 100 map was 95.6 percent accurate.

11) Residential areas in many cases were falsely depicted by the LSDP, LCMP, and the Image 100 maps as cultivated fields.

12) Forest areas were usually underestimated by the various computer maps, while the open uncultivated fields were overestimated. The discrepancies were most likely caused by the season of the year,

```

<<#| < / //+++++//+++++
CD<# T //+++++//+++++ %
D/T T|s //+++++//+++++ .#
D / //sT<<<< //+++++//+++++ %w
D //s// //+++++//+++++ $=
D // <<#// //+++++//+++++ =#
// D< //+++++//+++++ <
D< T // //+++++//+++++ ww
/< //TD D //+++++//+++++ $$#
# // //+++++//+++++ |###
<#s //+++++//+++++ //T
//+++++//+++++ //s$#
*//+++++//+++++ w|=
//+++++//+++++ |.%

```

Fig. 8. Caloosahatchee River. A bridge in northwestern direction is denoted by blank spaces. LCMP computer map. Input data 2/21/77. Fort Myers test site.

but other factors such as the size of ground areas covered by these two categories, the interchanging landscape schemes on the ground (spatial patterns), and tree species may also be important.

13) As one may anticipate, the results obtained from computer-generated maps are better when they refer to major ground cover types such as forest areas, lakes, large agricultural and/or uncultivated fields. Small residential areas, and fields cannot be delineated with adequate accuracy. Small towns like Waldo and Starke in the Gainesville test sites are confused with cultivated fields.

14) The Image 100 allows only for eight different themes at one time for the same scene. In highly diversified sites, where more than eight ground cover categories may be present, one ends up with a relatively large number of "unclassified" and overlapping areas.

V. DISCUSSION

In using unsupervised, computer-aided pattern recognition methods (such as the one employed by the LSDP and LCMP maps) good results may be expected only when the features of interest have distinct spectral signatures. Unfortunately, in the real world of renewable natural resources such desirable features are not abundant. Data analysts and resource specialists are confronted with highly variable and often overlapping spectral patterns even when dealing with seemingly simple resources such as bare soil or forest cover. It is not sufficient to know the specific spectral characteristics of a single resource, such as a given tree species, but also spatial and temporal variations, along with

the dynamic factors influencing such variations. Therefore, to make effective use of Landsat data, and the available processing methods, there is a need to develop reference data from the same areas at different times of the year and over a period of years.

Powerful interactive devices, such as the Image 100, depends heavily on man/machine interface. If knowledge of dynamic spectral characteristics for the study area is available, one would expect to produce reliable results.

In sensing ground cover conditions, Landsat depicts the broader scene. As a result, the presence of an earth feature may be obscured by another one. Such cases were found, for example, in the test sites where relatively open forest stands were classified as uncultivated fields. Apparently, strong reflectance from the understory overshadows that of an open overstory. Thus land cover computer maps derived from Landsat data may not always be closely related to the actual use of a given piece of land.

The ease of converting a version of the LSDP family of Landsat analyses programs to the University of Florida computing system (Amdahl 470 V/6-II) during the course of this study and their subsequent accessibility by the other eight Universities of the Florida State University System suggests that the KSC programs have the potential of becoming readily available to a wide range of users. This Landsat analysis tool can run on any available general purpose computer system that accepts FORTRAN IV and has an associated tape reader and a display device. The novel feature of this technique is that it is very simple to utilize. Once the programs are operational, all a user need specify is the center of the scene to be analyzed and the level of confidence desired. Although these programs could be most effectively employed by a sophisticated remote sensing analyst who could store and refine signatures via the LSCP ancillary program, the technique's widest appeal would be for an individual user who is neither a computer nor a remote sensing expert. This feature makes these programs especially suitable for training students in the rudiments of remote sensing by satellite.

VI. CONCLUSIONS

There is every indication to suggest that digital multispectral image processing systems based on Landsat input data will play an increasingly important role in pattern recognition and mapping land cover in the years to come. Repeatability

and versatility are but two of the attractive features of this approach. Qualified answers to ever present questions of renewable natural resources and respective changes through time may be provided by rapid processing of Landsat data.

To make such an approach a cost-effective one on an operational basis there is a need for close cooperation between resource analysts and those familiar with multispectral processing systems similar to the one investigated in this study. Recent studies⁵ suggest the minimum area for which this approach can become cost-effective is between one and two million acres.

Computer-produced maps from Landsat provide a synoptic appraisal of terrain features. The ease of their frequent update may greatly assist rational planning, especially in areas characterized by rapid changes of land and water use due to human activities.

Overall, the degree of fidelity of the evaluated maps to the actual ground conditions is considered to be satisfactory. The results are in line with reported work which has been conducted under comparable conditions.

More research is needed to refine the whole approach from the machine-processing of Landsat input data to the ground feature extraction. The study was convincing enough that computer classification of digital Landsat multispectral data, supplemented with auxiliary information, such as vegetation species, soil types, and microclimate, may soon become an indispensable tool in the hands of the skillful analyst of renewable natural resources. Simulated parallax to produce stereoscopic Landsat scenes would further enhance the use of this technique, especially with the future availability of the advanced multispectral scanner (thermatic mapper) of the forthcoming Landsat.

VII. ACKNOWLEDGEMENTS

Funds and needed facilities for this project have been provided by NASA/Kennedy Space Center, Science Technology and Application Branch. Reed Barnett and Roy A. Bland have provided valuable technical support. We wish to thank H. Royce of the Computer Sciences Corporation at Kennedy Space Center, for computer programming assistance. We are grateful to Sandra Buby for her editorial comments, and to Diane Muscato who carefully typed the manuscript.

VIII. LITERATURE CITED

1. Arvanitis, L. G. 1978. Distribution Mapping of Melaleuca in South Florida via Remote Sensing. Univ. Fla., School of For. Res. and Conservation. Research Report, prepared for the U.S. Forest Service, SE For. Exp. Sta. and the Florida Division of Forestry, 113p.
2. Arvanitis, L. G., R. Newburne and R. Reich. 1980a. Evaluation of Remote Sensing Techniques on Selected Forest Sites in Florida. Res. Report NASA, CR-154635 (N80-32805/NSP).
3. Arvanitis, L. G., R. Newburne and R. Reich. 1980b. Evaluating Computer-Drawn Ground Cover Maps. NASA Tech. Briefs. Math. and Info. Sciences:5(4): 512.
4. Ginevan, M. E. 1979. Testing Land-Use Map Accuracy: Another Look. Photogram. Eng. 10:1371-1377.
5. Harding, R. A. and R. B. Scott. 1978. Forest Inventory with Landsat. Report, State of Wash., Depart. of Nat. Res., 221p.
6. Hoffer, R. M. 1978. Biological and Physical Considerations in Applying Computer-Aided Analysis Techniques to Remote Sensor Data. In P. H. Swain and S. M. Davis (Ed.) "Remote Sensing. The Quantitative Approach." McGraw-Hill, Inc.
7. Royce, H., K. McGuire, and R. Bland. 1976. Landsat Signature Development Program. In Purdue Univ. Proc., LARS Symp. Machine Proces. of Remotely Sensed Data, 299p.

AUTHOR BIOGRAPHICAL DATA

Loukas G. Arvanitis is Professor, School of Forest Resources and Conservation, University of Florida, Gainesville, and Professor, School of Agriculture and Forestry, University of Thessaloniki, Greece.

Robin M. Reich is Graduate Student, School of Forest Resources and Conservation, University of Florida, Gainesville.

Robert Newburne is County forester with the Florida Division of Forestry, Miami. When this study was conducted, he was Assistant in the School of Forest Resources and Conservation, University of Florida, Gainesville.