7.9 10077

CR151860

LOCKHEED ELECTRONICS COMPANY

AEROSPACE SYSTEMS DIVISION

16811 EL CAMINO REAL • HOUSTON, TEXAS 77058 • TELEPHONE (AREA CODE 713) 488-0080

Ref: 642-2230
Contract NAS 9-15200
Job Order 73-763

TECHNICAL MEMORANDUM

TEST AND EVALUATION OF PRINCIPAL COMPONENT CLUSTER IMAGES IN LACIE

By

R. A. Abotteen

Approved By: J. C. Minter
T. C. Minter, Supervisor
Techniques Development Section

Distribution:

JSC/L. F. Childs
J. D. Erickson (4)

LEC/M. L. Bertrand
B. L. Carroll
J. E. Davis
J. A. Quirein
D. G. Saile
P. C. Swanzy
R. E. Tokerud
REDAF (3)
Technical Library (2)
Job Order File

[redacted]

(E79-10077) TEST AND EVALUATION OF PRINCIPAL COMPONENT CLUSTER IMAGES IN LACIE (Lockheed Electronics Co.) 17 p HC A02/MF A01 CSCL 02C

N79-15356

Unclas G3/43 00077

March 1977

LEC-10052

A SUBSIDIARY OF LOCKHEED AIRCRAFT CORPORATION

CONTENTS

TABLES

1. INTRODUCTION

Image interpretation is an important method for acquiring training data to classify Landsat images in the Large Area Crop Inventory Experiment (LACIE). Interpreting a scene for classification requires that training fields containing statistically representative samples of all spectral signatures in the given scene be selected and correctly labeled. When multiple passes over a scene are to be interpreted, the variation of the spectral signatures in a multitemporal sense makes it difficult to select and identify all of the variety of signatures in the scene. To address this problem in multitemporal image interpretation of LACIE segments, the principal component (PCOMP) cluster images are introduced as an image interpretation aid.

Some analyses of PCOMP transformed Landsat data have been performed (refs. 1,2). In this report, the PCOMP transformation was applied to some multitemporal clustered LACIE data. The object was to find whether PCOMP cluster images were helpful in selecting all multitemporal signatures and how much information (in terms of probability of misclassification) was lost when the three most significant PCOMP channels were used in classifying the PCOMP cluster image.

A cluster image was generated by clustering the picture elements (pixels) in the segment and then replacing each pixel by the mean of the cluster to which it was assigned. Cluster keys were also generated as part of the cluster image (ref. 3). The keys were ordered according to the Kauth greenness number (ref. 4). A PCOMP cluster image was generated by applying the PCOMP transformation to the cluster image and the cluster keys. The PCOMP transformation was the matrix of eigenvectors obtained from the mixture covariance matrix Σ_m. A cluster covariance matrix Σ_c also was calculated using the cluster means in the cluster image.

The purpose was to have an alternative method of obtaining the PCOMP transformation. It was found that Σ_c was not significantly different from Σ_m; hence, all PCOMP transformations in this report were calculated from Σ_m.

If m acquisitions from a LACIE segment were clustered in n channels, where n = 4m, an n-dimensional cluster image (with n-dimensional cluster keys) would be generated. Note that m color infrared (CIR) film products of the cluster image could be generated using a production film converter (PFC). The first film product would use channels 1, 2, and 4. The second film product would use channels 5, 6, and 8; whereas the m*th* film product would use channels (n - 4), (n - 3), and n. A color film product of the segment PCOMP cluster image would be generated by the PFC using the first three (which are the most significant) channels of the image. CIR film products of the segment cluster image and PCOMP cluster image would also show color cluster keys.

2. TEST AND EVALUATION PROCEDURES

2.1 TEST PROCEDURE

The following procedural steps were carried out in performing the test:

a. For every image, the PCOMP transformation was calculated using Σ_m, where Σ_m was the mixture covariance matrix calculated from all pixels in the image.

b. For every image, a cluster image was generated using standard clustering parameters.

c. Using the PCOMP transformation, a PCOMP cluster image was generated.

d. CIR film products were generated from the cluster image.

e. One color film product was generated from the PCOMP cluster image.

f. CIR film products of the image were also generated.

g. Using products from steps e, f, and g, along with the image ancillary data, color cluster keys were labeled as wheat or nonwheat.

h. The initial wheat proportion estimate $\hat{q}_W$ and nonwheat proportion estimate $\hat{q}_N = 1 - \hat{q}_W$ were computed.

i. Using the labeled color keys in the PCOMP cluster image film product, representative wheat and nonwheat training fields were selected in proportion to $\hat{q}_W$ and $\hat{q}_N$.

j. Ground truth (GT) test fields were selected for each segment using GT maps.

k. An Earth Resources Interactive Processing System (ERIPS) batch job was submitted to classify each segment twice using the original features. One classification was performed using the training fields selected in step i with equal wheat and nonwheat *a priori* probabilities. This classification will be referred to as the *PCOMP fields/equal a prioris*. The second classification was performed using the training fields selected in step i with $\hat{q}_W$ and $\hat{q}_N$ obtained from step h, as the wheat and nonwheat *a priori* probabilities, respectively. This classification will be referred to as *PCOMP fields/user-input a prioris*.

ℓ. The ERIPS was used to classify each PCOMP cluster image interactively using the first three PCOMP features. This classification will be referred to as *PCOMP$^{(3)}$ fields/equal a prioris*.

A flow chart of the test procedure is shown in figure 1.

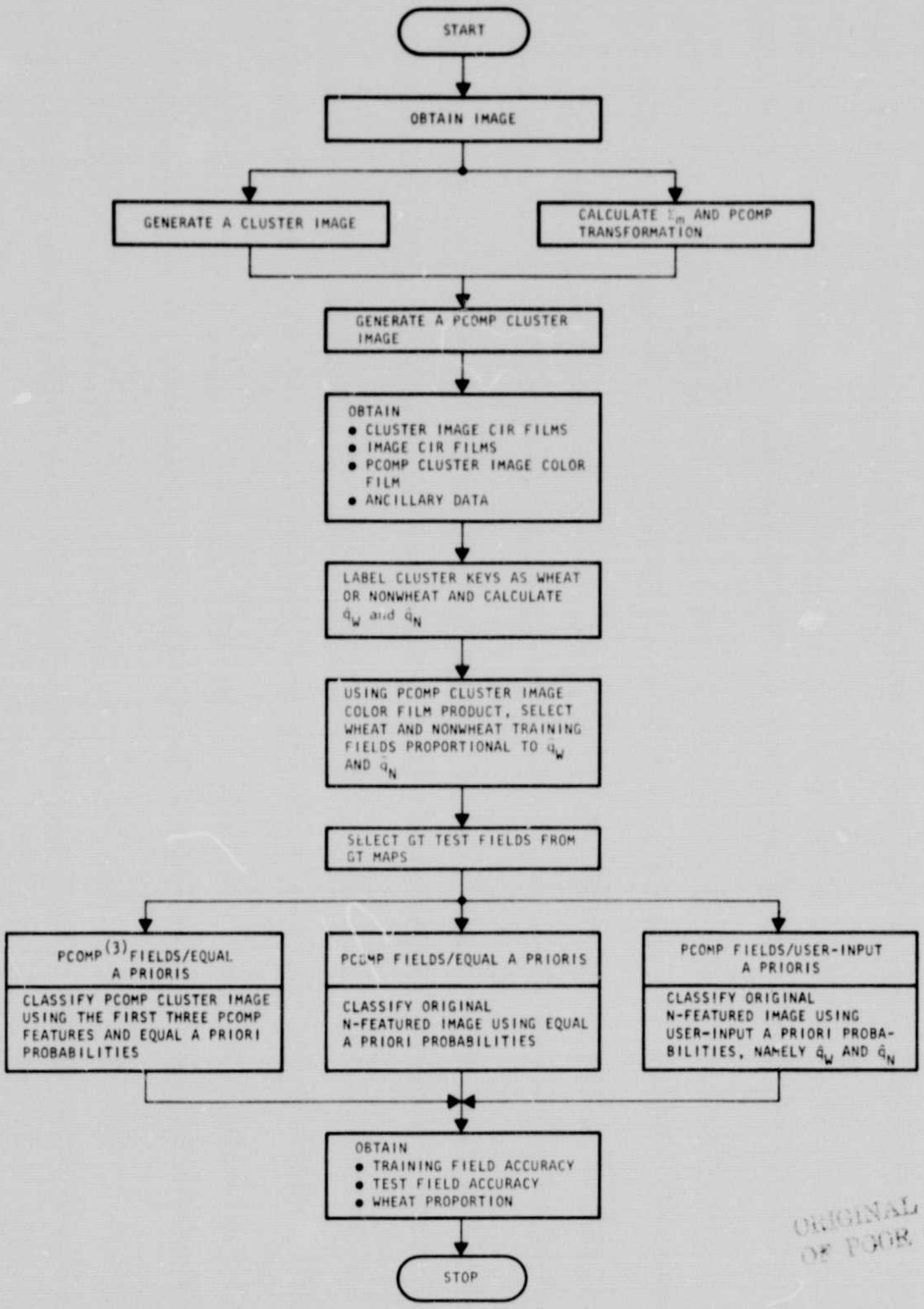

Figure 1.— Flow chart for test procedure.

2.2 EVALUATION PROCEDURE

a. Classification with original features:

(1) The average probability of correct classification (PCC_{avg}) over all test segments was calculated using *PCOMP fields/equal a prioris* procedures and compared with the PCC_{avg} values obtained using the LACIE-selected training fields standard procedures.

(2) The mean square error (MSE) in wheat proportions was calculated using *PCOMP fields/equal a prioris* procedures and compared to the MSE values obtained using the LACIE-selected training fields standard procedures. The GT wheat proportions were used as a basis for calculating the MSE.

(3) The average training field accuracy over all test segments was calculated using *PCOMP fields/equal a prioris* procedures and using the LACIE-selected training fields standard procedures.

b. Classification with the first three PCOMP features:

(1) The PCC_{avg} was calculated using *PCOMP$^{(3)}$fields/equal a prioris* procedures and compared with the LACIE PCC_{avg} and the *PCOMP fields/equal a prioris* PCC_{avg} values.

(2) The MSE in wheat proportions was calculated using *PCOMP$^{(3)}$fields/equal a prioris* procedures and compared with the LACIE and the *PCOMP fields/equal a prioris* MSE's.

3. DESCRIPTION OF THE DATA USED

Most of the LACIE sample segments used in this test contained GT information. A description of those sample segments is shown in table I. Sample segments 1270, 1087, 1576, 1046, 1224, and 1882 are blind sites. Sample segments 1976 and 1988 are intensive

TABLE I.– SAMPLE SEGMENT DESCRIPTION

Segment	Acquisitions processed (a)	Location (county/state)
1270	76072, 76162	Throckmorton/Texas
1087	76001, 76172	Crosby/Texas
1576	76162, 76198	Lancaster/Nebraska
1046	76127, 76190	Beaver/Oklahoma
1149	76099, 76189	Harrison/Indiana
1224	76080, 76162	Dewey/Oklahoma
1882	76090, 76153, 76189	Lincoln/Kansas
1976	76136, 76208, 76226	Franklin/Idaho
1988	76038, 76109, 76127, 76164	Finney/Kansas

[a]Numbers represent Julian dates; e.g., 76072 is the 72nd day of 1976.

test sites (ITS's). Sample segment 1149 is an operational LACIE segment for which no GT information is available.

Each acquisition shown in table I represents 4 channels; hence, 6 segments were processed in 8 channels, 2 segments were processed in 12 channels, and 1 segment (1988) was processed in 16 channels.

4. RESULTS

The results of the test which is the subject of this report are presented in tabular form. Tables include the classification performance for each segment using the various procedures. The classification performance is specified in terms of training field accuracy; test field accuracy, which is the probability of correct classification (PCC); and wheat proportions. When equal *a priori* probabilities were used in classification, the training- and test-field accuracies were computed as the arithmetic average of the wheatfield and the nonwheatfield accuracies. When LACIE procedures were used in processing the data set, equal *a prioris* were used in classification. When user-input *a priori* probabilities $\hat{q}_W$ and $\hat{q}_N$ were used, the field accuracy A was computed as:

$$A = \hat{q}_W A_W + \hat{q}_N A_N \tag{1}$$

where A_W was the wheatfield accuracy and A_N was the nonwheatfield accuracy. It is worth mentioning at this point that the LACIE test-field accuracies for all blind sites were computed based on analyst-selected test fields, whereas the test-field accuracies using other procedures were based on fields selected from the GT maps. Analyst-selected test fields tended to show greater accuracy.

The results of the test conducted are presented in tables II, III, and IV. Table II presents the training- and test-field accuracies for the data set using LACIE, *PCOMP fields/equal a prioris*, and $PCOMP^{(3)}$*fields/equal a prioris* procedures. It is noteworthy that (1) segment 1087 was processed manually in LACIE (i.e., the number of wheat pixels in the segment were hand counted) and (2) segments 1149 and 1988 were not processed multitemporally in LACIE. Table II shows that the LACIE and *PCOMP fields/equal a prioris* procedures yielded close training- and test-field accuracies. Table II also shows that the $PCOMP^{(3)}$*fields/equal a prioris* procedure yields remarkably close accuracies to those obtained when the *PCOMP fields/equal a prioris* procedure was used. This shows that, in terms of PCC, no significant loss of information was noted when the first three PCOMP channels were used in classifying the data set used in this test.

Table III presents the wheat proportions in the data set using GT and LACIE, *PCOMP fields/equal a prioris*, and $PCOMP^{(3)}$*fields/equal a prioris* procedures. Table IV shows the effects of using the estimated *a priori* probabilities on classification performance.

A summary of tables II, III, and IV is given in table V. For all the procedures used in this test, table V shows the average training-field accuracy, the PCC_{avg}, and the MSE in wheat proportions. The procedures displayed higher MSE in wheat proportions than the LACIE MSE, primarily because the wheat proportions in segments 1087 and 1576 were highly overestimated using the three *PCOMP* procedures. Table V indicates that the *PCOMP fields/user-input a prioris* procedure was useful in the multitemporal processing of LACIE segments. As shown also in table V, the $PCOMP^{(3)}$*fields/equal a prioris* procedure displayed a good average classification performance in comparison with the rest

TABLE II.– TRAINING- AND TEST-FIELD ACCURACIES FOR THE DATA SET USING LACIE, *PCOMP FIELDS/EQUAL A PRIORIS*, AND *PCOMP$^{(3)}$FIELDS/EQUAL A PRIORIS* PROCEDURES

Segment	Training-field accuracy, %			Test-field accuracy, %		
	LACIE	*PCOMP fields/ equal a prioris*	*PCOMP$^{(3)}$ fields/ equal a prioris*	LACIE	*PCOMP fields/ equal a prioris*	*PCOMP$^{(3)}$ fields/ equal a prioris*
1270	98.9	97.9	94.1	88.1	–	–
1087	(a)	96.7	94.4	–	82.9	80.8
1576	98.2	92.8	91.5	97.6	75.8	77.0
1046	99.3	92.8	89.5	88.8	95.6	96.9
1149	[b]97.6 (76099)	90.2	88.2	–	–	–
1224	98.0	94.0	94.6	95.1	96.3	97.6
1882	99.3	96.3	94.8	94.6	88.8	90.5
1976	99.5	92.6	89.2	36.7	36.4	46.9
1988	[b]98.0 (76109)	96.1	95.2	[b]80.2 (76109)	71.5	78.0

[a]This segment was manually processed by the LACIE.
[b]Processed by the LACIE using the Julian acquisition date shown in parentheses.

TABLE III.– WHEAT PROPORTIONS IN THE DATA SET USING GROUND TRUTH AND LACIE, *PCOMP FIELDS/EQUAL A PRIORIS*, AND *PCOMP*[3] *FIELDS/ EQUAL A PRIORIS* PROCEDURES

[Percentage]

Segment	GT	Procedure		
		LACIE	*PCOMP fields/ equal a prioris*	*PCOMP*[3] *fields/ equal a prioris*
1270	–	18.5	28.2	43.8
1087	10.0	[a]2.8	25.3	25.0
1576	11.4	21.4	43.3	45.4
1046	23.1	12.4	22.9	15.0
1149	–	[b]14.0 (76099)	34.5	41.1
1224	44.7	34.0	35.9	34.2
1882	46.7	48.2	42.3	38.9
1976	28.2	26.8	24.4	31.1
1988	33.0	[b]23.5 (76109)	34.1	29.9

[a]This segment was manually processed by the LACIE.

[b]Processed by the LACIE using the Julian acquisition date shown in parentheses.

TABLE IV.- EFFECT OF USING THE ESTIMATED *A PRIORI* PROBABILITIES ON CLASSIFICATION PERFORMANCE

Segment	Wheat proportions, %				Test-field accuracy, %	
	$\hat{q}_W$	GT	*PCOMP fields/user-input a prioris*	*PCOMP fields/equal a prioris*	*PCOMP fields/user-input a prioris*	*PCOMP fields/equal a prioris*
1270	21.0	—	25.4	28.2	—	—
1087	14.0	10.0	22.2	25.3	73.7	82.9
1576	30.0	11.4	38.2	43.3	76.6	75.8
1046	12.0	23.1	15.4	22.9	97.2	95.6
1149	11.0	—	16.9	34.5	—	—
1224	34.0	44.7	34.5	35.9	94.4	93.6
1882	23.0	46.7	39.3	42.3	90.3	88.8
1976	7.0	28.2	16.5	24.4	56.3	36.4
1988	27.0	33.0	33.1	34.1	72.2	71.5

TABLE V.— SUMMARY OF RESULTS

Procedure	Average training-field accuracy, %	PCC_{avg}, %	MSE in wheat proportions, %
LACIE	98.6	83.0	8.2
PCOMP fields/ equal a prioris	94.4	78.2	14.0
PCOMP$^{(3)}$ fields/ equal a prioris	92.4	81.1	15.3
PCOMP fields/user-input a prioris	96.0	80.0	13.2

of the procedures. This implies that most of the information contained in the multitemporal images for this test can be concentrated in the first three PCOMP's. This test indicates that the PCOMP cluster images should be useful in the multitemporal processing of LACIE segments.

5. CONCLUSIONS AND RECOMMENDATIONS

Based on the results obtained in this test, the following conclusions were reached:

a. Color film products of the first three PCOMP channels in the PCOMP cluster images are helpful interpretive aids for multitemporal processing of LACIE segments.

b. A procedure such as *PCOMP fields/equal a prioris* yielded comparable results to LACIE-selected training fields standard procedures. The *PCOMP fields/user-input a prioris* procedure displayed slightly better results than the *PCOMP fields/equal a prioris* procedure.

c. It was found that most of the information (in terms of PCC) was preserved when the first three PCOMP channels were used in classifying the multipass PCOMP cluster images.

It is recommended that the LACIE analysts be provided and become familiar with a color film of the PCOMP cluster image when multitemporal processing of the LACIE image is required.

6. REFERENCES

1. Abotteen, R. A.: Analysis of Principal Component Transformed Landsat Data. Technical Memorandum LEC-9003, Aug. 1976.

2. Ready, P. J.; and Wintz, P. A.: Information Extraction, SNR Improvement, and Data Compression in Multispectral Imagery. IEEE Trans. and Commun. 10, Oct. 1973.

3. Levy, S. A.: Evaluation of Cluster Images. Technical Memorandum LEC-9001, Oct. 1976.

4. Kauth, R.: The Tasselled Cap Revisited. NASA/JSC Internal Note TF3-75-5-190, May 1975.