LEC

LOCKHEED ELECTRONICS COMPANY, INC.

AEROSPACE SYSTEMS DIVISION

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

Ref.: 642-2042
Contract NAS 9-12200
Job Order 73-763

79-10076

CR-151861

TECHNICAL MEMORANDUM

ANALYSIS OF PRINCIPAL COMPONENT TRANSFORMED LANDSAT DATA

By

R. A. Abotteen

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

Distribution:

JSC/L. F. Childs
J. D. Erickson (3)

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

(E79-10076) ANALYSIS OF PRINCIPAL COMPONENT TRANSFORMED LANDSAT DATA (Lockheed Electronics Co.) 25 p HC A02/MF A01 CSCL 02C

N79-15355

Unclas
G3/43 00076

August 1976

LEC-9003

CONTENTS

TABLES

FIGURES

1. INTRODUCTION

The Landsat multispectral scanner provides data in the form of four spectral images of the particular area of the Earth under observation. Ready and Wintz (ref. 1) have shown that the principal component (PCOMP) transformation applied to airborne- and satellite-gathered multispectral data is very useful for information extraction, since the first few PCOMP images contain essentially all the information present in the original spectral bands. In this report, a PCOMP transformation was applied to single-pass and multipass Landsat data. The transformed data were analyzed visually and automatically. The PCOMP transformation is

$$\underline{Y} = M\underline{X} \tag{1}$$

where

$\underline{X}$ = vector of n spectral intensities associated with each pixel.

M = n × n unitary matrix derived from the mixture covariance matrix Σ_X of the spectral bands such that the rows of M are the normalized eigenvectors of Σ_X.

$\underline{Y}$ = vector of n principal components.

The covariance matrix of the PCOMP transformed data then becomes

$$\Sigma_Y = M\Sigma_X M^T = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \tag{2}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n$ (the variances of the principal components) are the eigenvalues of Σ_X ordered such that $\lambda_1 > \lambda_2 > \cdots > \lambda_n$ and M^T is the transpose of M.

An implementation plan for the program which outputs eigenvalues and eigenvectors is presented in section 2. Several examples and applications are also presented in section 2. Section 3 deals with the effect of the transformation on signature extension, while section 4 states the conclusions drawn from this study.

2. APPROACH

The PCOMP transformation presented in this report was calculated as follows. A field containing all the picture elements (pixels) in a site was defined. An n-dimensional covariance matrix was calculated for that field. The eigenvalues and normalized eigenvectors corresponding to the covariance matrix were calcualted. The rows of the transformation matrix M were taken to be the normalized eigenvectors such that the first row corresponds to the largest eigenvalue and the second row corresponds to the second largest eigenvalue, etc.

2.1 IMPLEMENTATION OF THE PCOMP PROGRAM

The University of Houston SYMAT program (ref. 2) was modified and implemented on the Laboratory for Applications of Remote Sensing (LARS) terminal under the file name PCOMP. The modifications made allowed the PCOMP program to read the covariance matrix from LARSYS-punched cards. It also added the capability of projecting the principal components onto the Kauth (ref. 3) space. For multipass data, the principal components were projected onto the Kauth space four channels at a time since the Kauth transformation is four-dimensional.

2.2 EXAMPLES

The eigenvalues and eigenvectors of Σ_X for some U.S. Department of Agriculture (USDA) Statistical Reporting Service (SRS) and intensive test sites were computed using single-pass and multipass data. The sites used were Marion, Dickinson, and Morton counties in Kansas. The results are given in tables 1 and 2. Each table also shows the projection of the eigenvectors onto the Kauth space. The first two principal components in table 1 projected onto the Kauth space are primarily Bright (first component) stuff and Green (second component) stuff. The remaining

TABLE 1.— EIGENVALUES, EIGENVECTORS, AND THEIR PROJECTION ONTO THE KAUTH SPACE OF SINGLE-PASS SITES

Site (Landsat date)	Eigenvalues of Σ_X	Eigenvectors of Σ_X				Projection of eigenvectors onto Kauth space			
Marion (1454)	34.4	0.42	0.56	0.65	0.30	0.99	0.11	-0.02	-0.01
	13.6	-0.34	-0.63	0.54	0.44	-0.11	0.99	0.02	-0.02
	1.4	0.78	-0.54	0.10	-0.29	-0.03	-0.01	-0.99	-0.12
	0.8	-0.32	0.06	0.52	-0.79	-0.01	-0.02	0.12	-0.99
Marion (1634)	54.2	-0.03	0.12	-0.83	-0.54	-0.57	-0.82	0.01	0.01
	49.8	0.34	0.93	0.13	-0.01	0.81	-0.55	0.20	0.01
	4.3	-0.91	0.34	-0.07	0.24	-0.16	0.14	0.98	0.04
	0.9	-0.26	0.01	0.53	-0.81	-0.01	-0.01	0.04	-1.00
Dickinson (1454)	75.4	0.35	0.50	0.70	0.37	0.98	0.22	0.02	0.00
	19.7	-0.35	-0.70	0.45	0.44	-0.21	0.98	-0.01	0.02
	1.1	-0.75	0.50	-0.20	0.40	-0.02	0.01	0.96	0.27
	0.9	0.45	-0.11	-0.52	0.72	0.01	-0.03	-0.27	0.96
Dickinson (1634)	77.5	-0.09	-0.29	0.78	0.55	0.38	0.92	0.00	0.00
	52.4	-0.45	-0.83	-0.31	-0.06	-0.92	0.38	-0.06	0.01
	3.1	-0.84	0.47	-0.09	0.24	-0.06	0.04	1.00	0.06
	1.0	-0.28	0.01	0.54	-0.80	-0.01	0.01	0.06	-1.00
Dickinson (1652)	90.4	0.35	0.68	-0.48	-0.44	0.18	-0.98	0.00	-0.01
	51.2	-0.38	-0.52	-0.69	-0.34	-0.98	-0.18	0.00	0.01
	1.7	-0.11	0.23	-0.52	0.82	0.01	-0.01	0.39	0.92
	1.4	0.85	-0.47	-0.18	0.14	0.00	-0.02	-0.92	0.39

TABLE 2.– EIGENVALUES, EIGENVECTORS, AND THEIR PROJECTION ONTO THE KAUTH SPACE OF TWO-PASS SITES

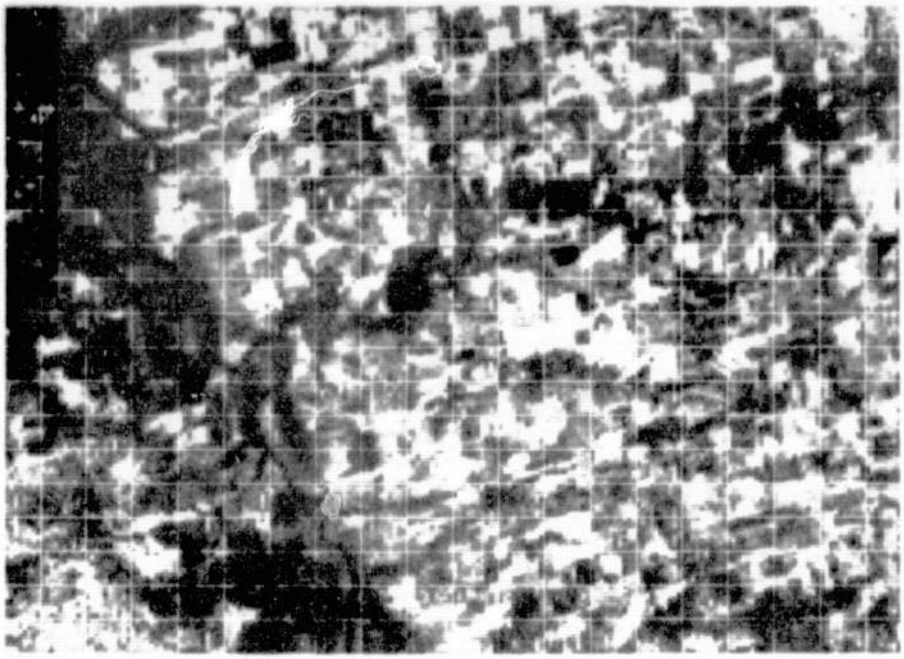
Site (Landsat dates)	Eigenvalues of Σ_X	Eigenvectors of Σ_X								Projection of eigenvectors onto Kauth space (four channels at a time)							
Marion (1454, 1634)	57.1	0.14	0.19	0.22	0.10	0.06	-0.04	0.79	0.50	0.34	0.03	0.00	0.00	0.60	0.73	0.00	-0.01
	50.0	-0.02	0.04	0.04	0.04	0.33	0.94	0.04	-0.07	0.05	0.03	0.04	0.00	0.74	-0.63	0.20	0.01
	31.5	0.40	0.52	0.62	0.28	-0.02	-0.04	-0.29	-0.18	0.93	0.10	-0.03	-0.01	-0.25	-0.23	-0.02	0.00
	13.6	0.33	0.63	-0.54	-0.44	-0.02	0.03	0.00	0.00	0.11	-0.99	-0.01	0.02	0.01	-0.02	0.03	-0.01
	4.3	0.05	-0.05	0.02	0.00	-0.90	0.34	-0.08	0.24	0.01	0.03	-0.07	0.00	-0.15	0.14	0.98	0.04
	1.2	-0.78	0.54	-0.09	0.30	-0.08	-0.02	-0.01	0.02	0.03	0.01	0.99	0.13	-0.05	0.03	0.05	0.00
	0.9	-0.11	0.02	0.18	-0.26	-0.24	0.01	0.50	-0.76	0.00	0.00	0.04	-0.33	-0.01	-0.01	0.04	-0.94
	0.8	0.30	-0.06	-0.48	0.75	-0.08	0.00	0.18	-0.27	0.01	0.02	-0.11	0.93	0.00	0.00	0.01	-0.34
Dickinson (1454, 1634)	99.8	-0.24	-0.36	-0.48	-0.25	0.03	0.13	-0.58	-0.40	-0.68	-0.14	-0.02	0.00	-0.35	-0.63	-0.01	0.00
	57.0	0.16	0.23	0.37	0.20	0.32	0.71	-0.24	-0.28	0.49	0.15	0.01	0.00	0.37	-0.77	0.05	-0.01
	48.6	0.19	0.26	0.35	0.17	-0.32	-0.51	-0.55	-0.27	0.50	0.09	0.00	-0.01	-0.86	-0.08	-0.03	0.00
	19.6	-0.35	-0.70	0.44	0.43	-0.03	-0.03	0.00	0.01	-0.22	0.97	-0.01	0.02	-0.02	0.03	0.01	0.00
	3.1	0.07	-0.03	0.02	0.00	0.84	-0.47	0.08	-0.24	0.02	0.01	-0.07	0.00	0.05	-0.05	-0.99	-0.06
	1.1	0.64	-0.44	0.19	-0.38	-0.18	0.04	0.24	-0.34	0.01	-0.01	-0.84	-0.28	0.00	0.01	0.09	-0.45
	1.0	-0.38	0.23	-0.04	0.11	-0.22	-0.01	0.48	-0.71	-0.01	0.01	0.46	0.03	-0.01	0.01	0.02	-0.89
	0.8	-0.44	0.10	0.52	-0.72	0.05	0.00	-0.05	0.08	-0.01	0.03	0.26	-0.96	0.01	0.00	-0.03	0.10
Morton (1454, 1727)	751.8	-0.11	-0.15	-0.14	-0.06	-0.49	-0.60	-0.53	-0.23	-0.23	0.00	0.00	0.01	-0.97	0.05	0.07	-0.02
	281.6	0.40	0.54	0.64	0.27	-0.16	-0.11	-0.13	-0.07	0.96	0.10	-0.02	-0.03	-0.23	0.00	0.07	-0.03
	57.3	-0.20	-0.20	0.22	0.22	-0.29	-0.45	0.58	0.45	-0.03	0.42	0.10	0.01	0.06	0.90	0.06	-0.02
	41.9	0.38	0.51	-0.51	-0.39	-0.16	-0.21	0.26	0.20	0.09	-0.90	-0.11	0.06	0.01	0.42	0.05	-0.02
	14.9	0.48	-0.49	0.38	-0.61	-0.02	-0.02	0.04	0.03	-0.04	0.06	-0.79	-0.60	0.01	0.06	0.01	0.00
	9.6	-0.02	-0.05	-0.03	0.01	-0.79	0.59	0.13	-0.10	-0.06	0.02	-0.01	0.02	0.08	-0.07	0.94	-0.32
	6.5	0.64	-0.37	-0.33	0.59	0.02	-0.02	0.02	-0.03	0.01	0.11	-0.59	0.79	0.00	0.00	-0.03	-0.03
	1.9	-0.03	0.01	0.01	-0.03	0.05	-0.18	0.52	-0.83	-0.01	0.00	0.02	-0.04	-0.01	0.00	-0.32	-0.95

Kauth-projected PCOMP eigenvectors are primarily Yellow (third component) stuff and Nonesuch (fourth component) or a combination of them. In table 2, the first four Kauth-projected PCOMP eigenvectors are primarily Bright and Green stuff while the remaining projections are primarily Yellow stuff and Nonesuch.

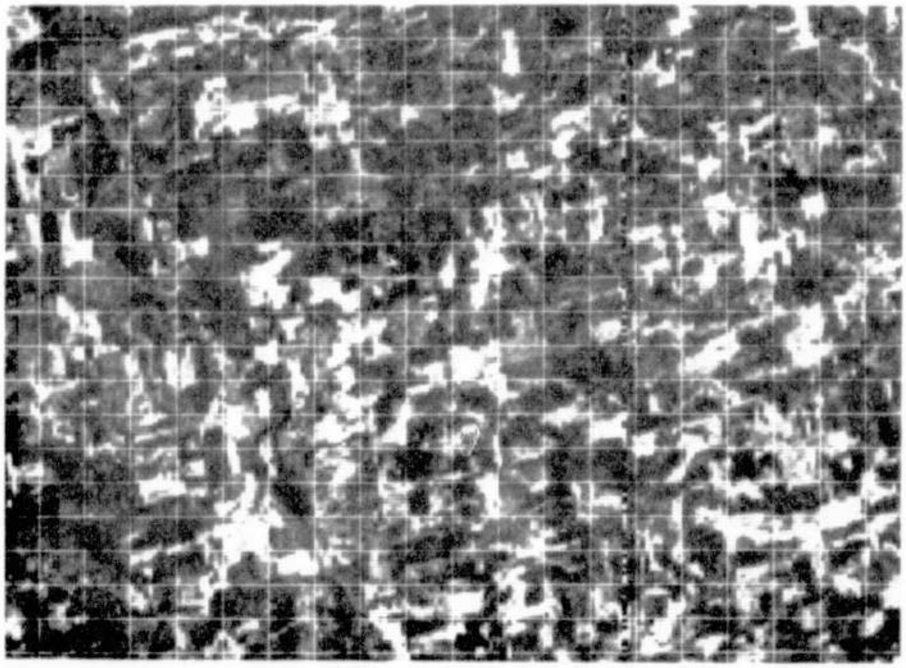
2.3 GENERATION OF PCOMP IMAGERY

Landsat data were transformed using the unitary matrix M defined in equation (1). The transformation was performed using the LARSYS data transformation processor. The processor transformed then rescaled the data to lie between 0 and 255. Black-and-white images of principal components 1, 2, 3, 4, 5, 6, 7, and 8 corresponding to the two-pass SRS Dickinson site (Landsat dates: 1454, 1634) are shown in figure 1. Figure 2 shows the first five principal component images of the Morton intensive test site (Landsat dates: 1454, 1727). There is essentially no information in PCOMP 5 or higher. This indicates that the information in the Landsat bands can be packed in fewer bands using the principal components that correspond to the largest eigenvalues of Σ_X.

A false color image of the first three principal components was generated using multitemporal data. Figure 3a shows a false color image of the original Landsat bands of Dickinson in the first pass (Landsat date: 1454). Figure 3b shows a false color image of Dickinson using another pass (Landsat date: 1634). The first three principal components of the eight-channel Dickinson site (Landsat dates: 1454, 1634) were used to generate the false color image shown in figure 4.

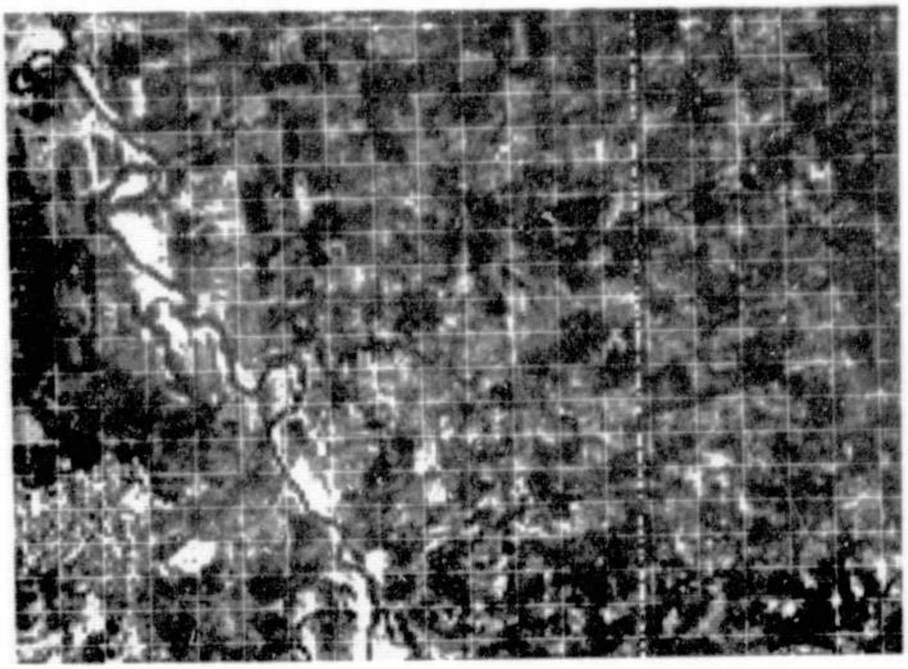


(a) PCOMP 1.

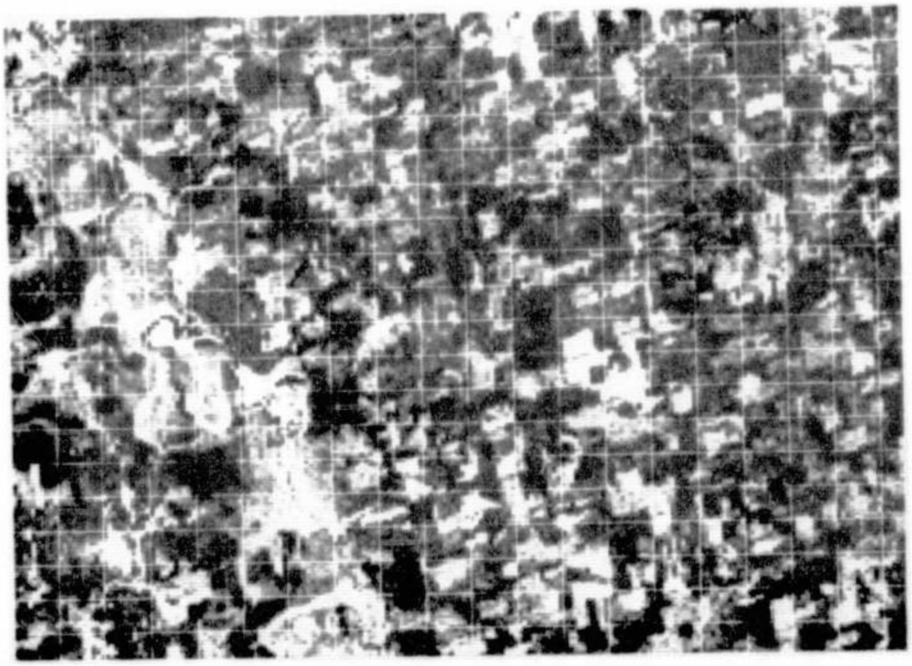


(b) PCOMP 2.

Figure 1. — Eight principal component images of the two-pass Dickinson site.

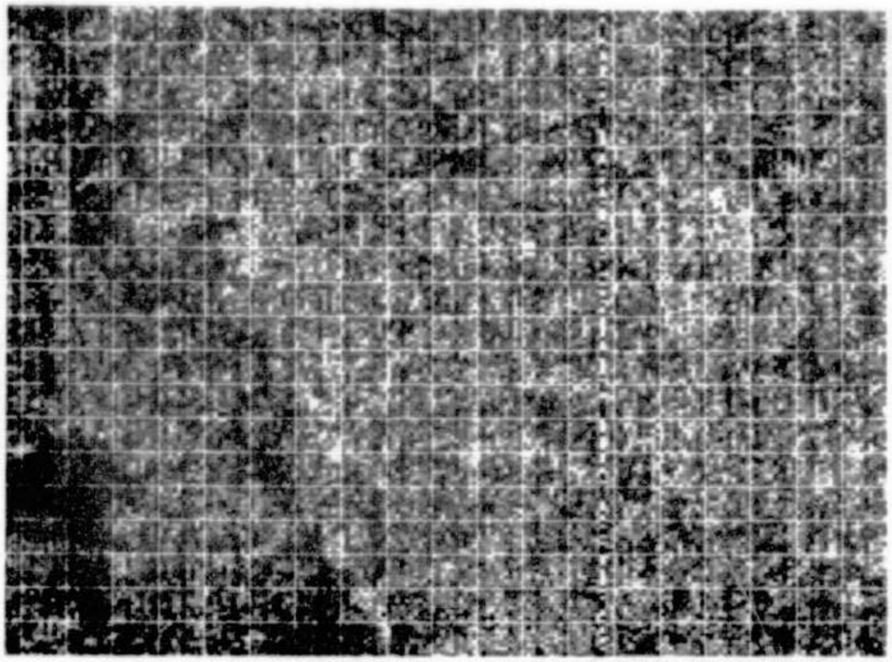


(c) PCOMP 3.

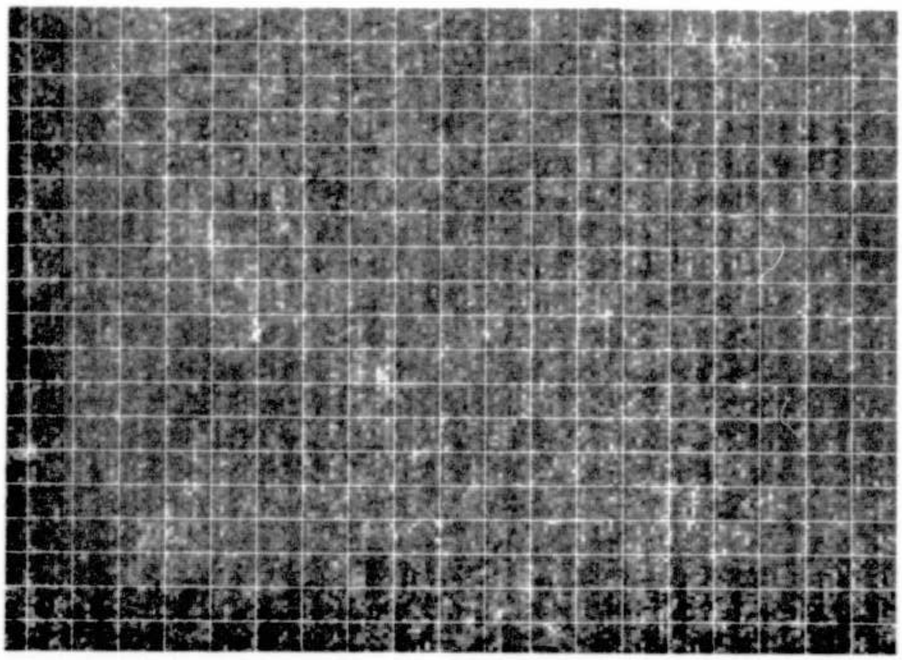


(d) PCOMP 4.

Figure 1. — Continued.

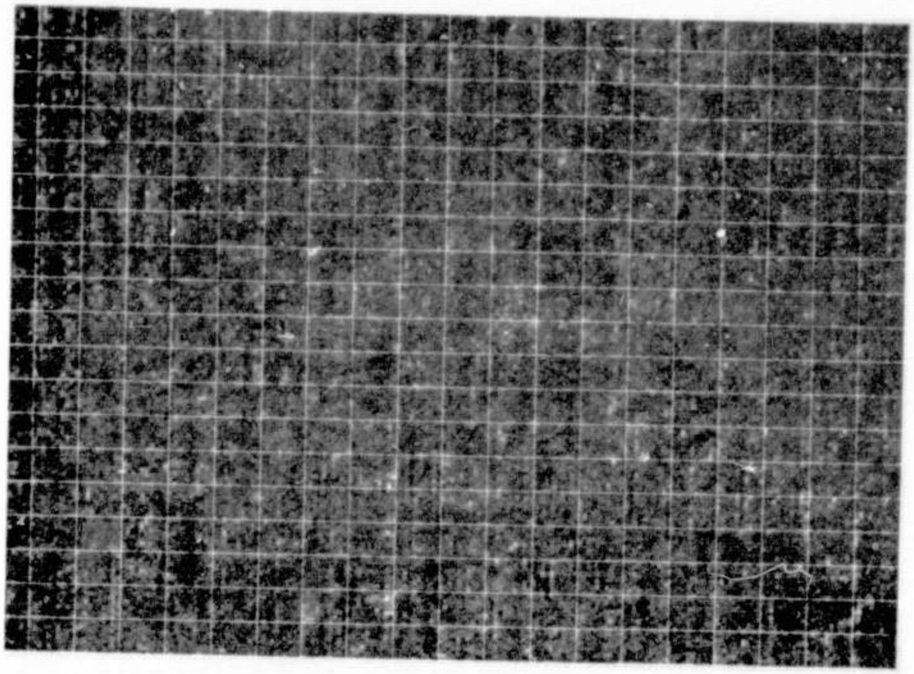


(e) PCOMP 5.

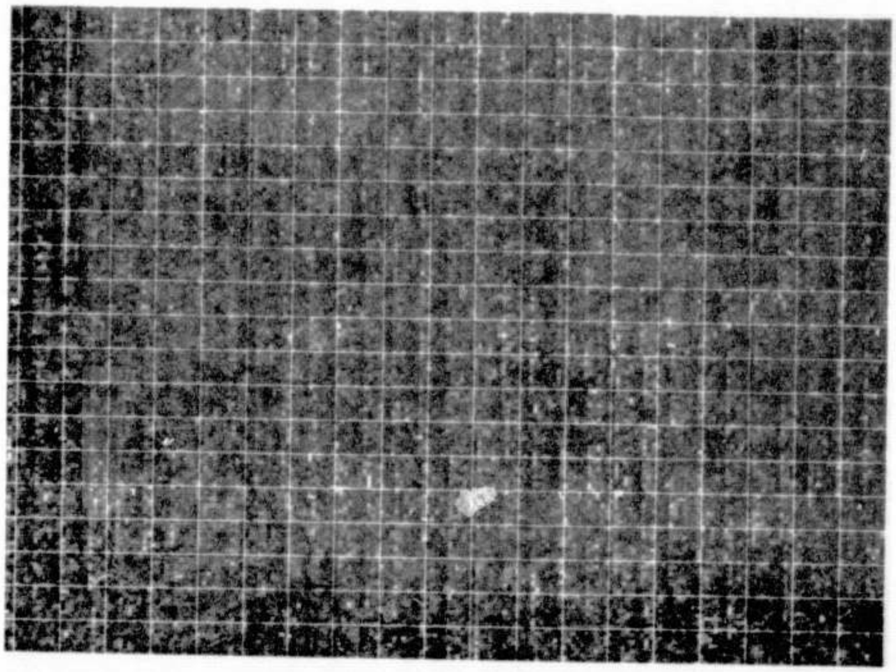


(f) PCOMP 6.

Figure 1. – Continued.

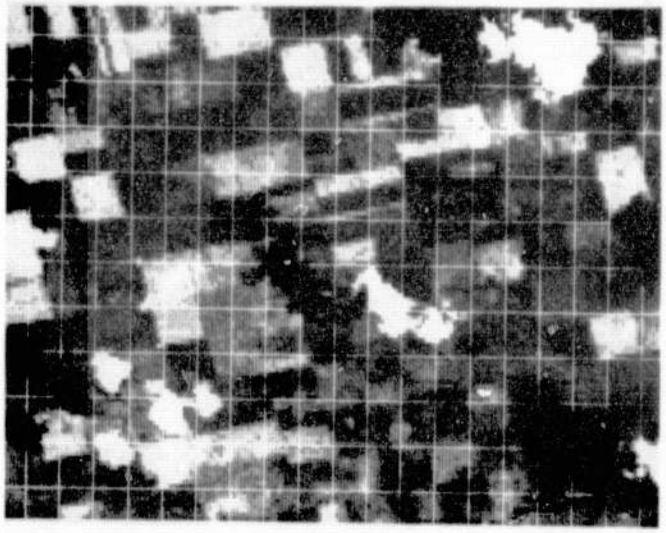


(g) PCOMP 7.

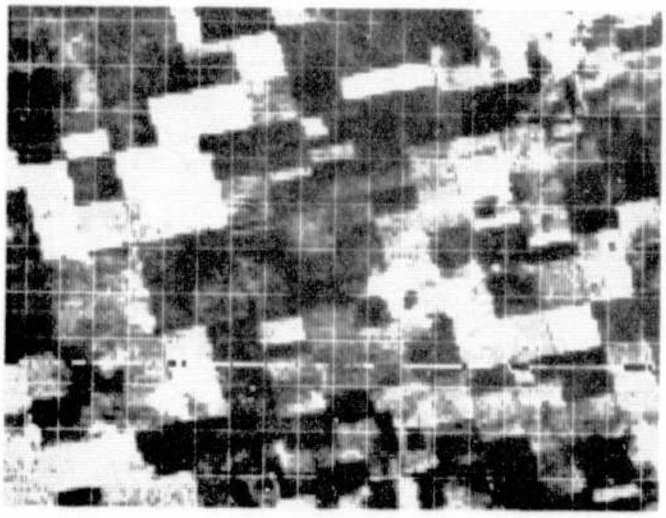


(h) PCOMP 8.

Figure 1. – Concluded.

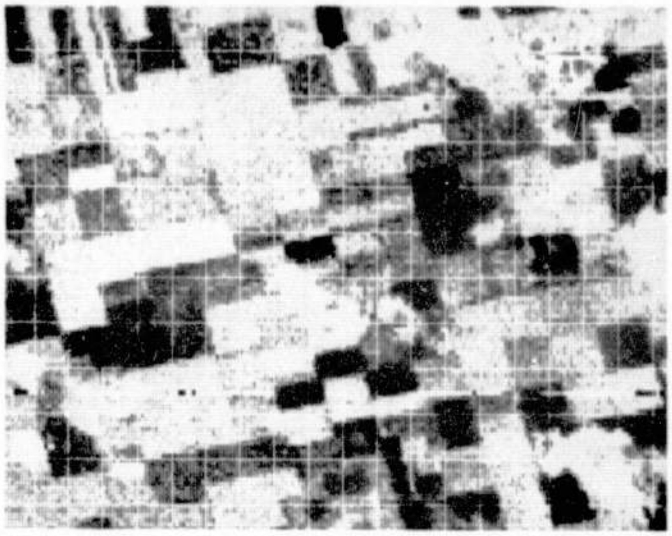


(a) PCOMP 1.

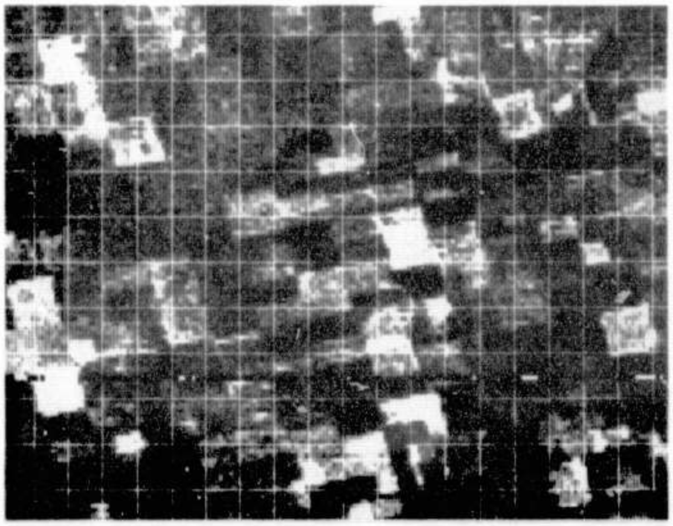


(b) PCOMP 2.

Figure 2. — First five principal component images of the two-pass Morton site.



(c) PCOMP 3.

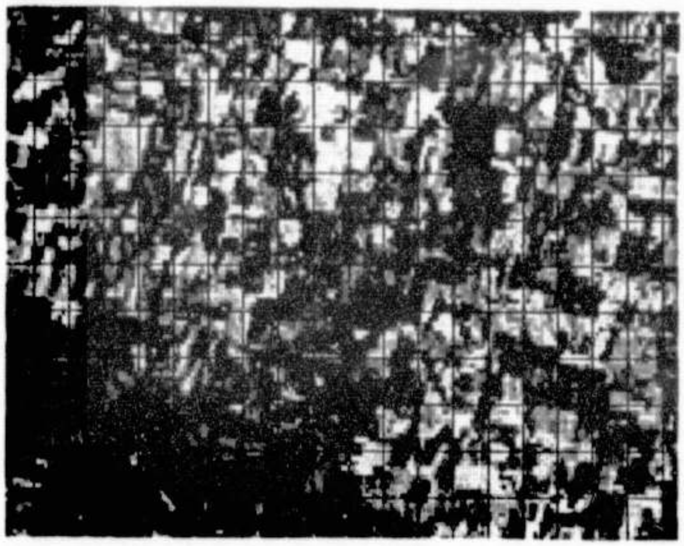


(d) PCOMP 4.

Figure 2. — Continued.

(e) PCOMP 5.

Figure 2. – Concluded.



(a) Landsat date: 1454.

(b) Landsat date: 1634.

Figure 3. - False color images of Dickinson site.

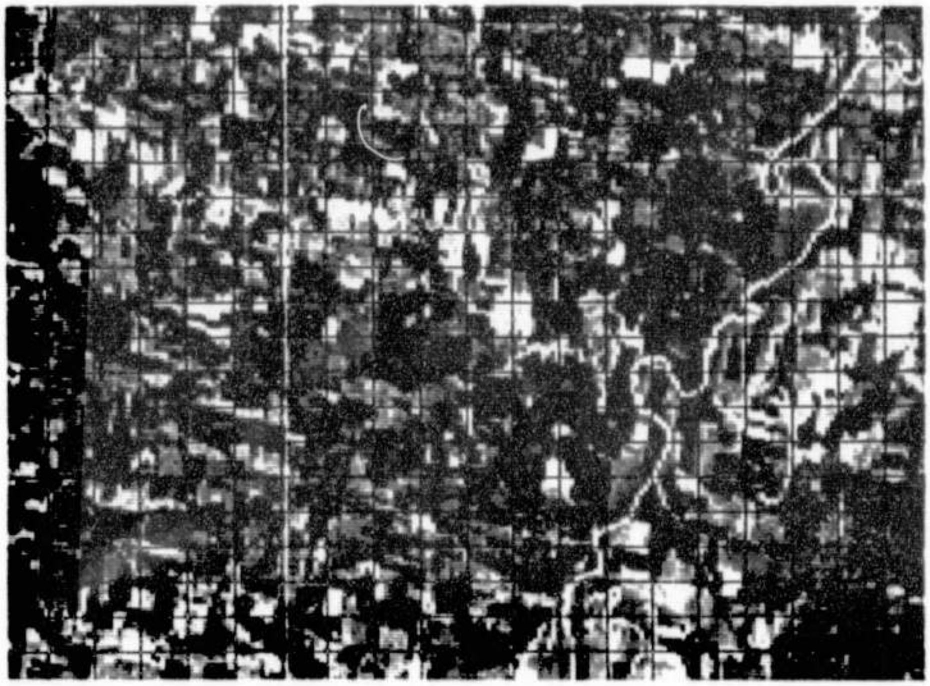


Figure 4. — First three PCOMP false color image of Dickinson site (Landsat dates: 1454, 1634).

3. THE EFFECT OF TRANSFORMATION ON CLASSIFICATION AND SIGNATURE EXTENSION PERFORMANCE

A consecutive-day acquisition (CDA) for Large Area Crop Inventory Experiment (LACIE) segment 1183 was chosen to study the effect of PCOMP transformation on classification and signature extension performance. This segment was acquired in the crop year 1975. The image with acquisition date 75110 was clear, while the 75111 acquired image was hazy. This segment along with several other CDA segments was used previously for testing signature extension algorithms (refs. 4 and 5). This segment was chosen because there is a large difference in haze level between the two acquisitions.

As mentioned previously in this report, the modified LARSYS data transformation processor transformed the data then rescaled it to have a range from 0 to 255. Various combinations of data manipulations were applied to both acquisitions of segment 1183.

The PCOMP transformation used for both acquisitions of segment 1183 was derived as follows. Both acquisitions were merged to form an image of 234 lines with 196 pixels in each line. Each acquisition has four channels, and thus the merged image also has four channels. A mixture covariance matrix Σ_c was calculated using all the pixels in the merged image. Eigenvalues and eigenvectors of Σ_c were calculated using the PCOMP program. The PCOMP transformation was the matrix of the first two eigenvectors of Σ_c. Once the data in each acquisition of segment 1183 were transformed, calculations of statistics and all classifications were performed in the two-dimensional transformed space. Each acquisition of segment 1183 was classified using local statistics and statistics from the other acquisition. The classification results are shown in tables 3 and 4. These tables also show some results obtained from reference 5 in which the original Landsat channels used for

TABLE 3.– CLASSIFICATION ACCURACIES FOR SEGMENT 1183 WITH VARIOUS STATISTICS AND DATA TRANSFORMATIONS

Acquisition	Crop	Field accuracy, %								
		Original Landsat data			Transformed Landsat data					
		Local	ROOSTER R(S)	UT	PCOMP/SC local	PCOMP/SC UT	SC local	SC UT	PCOMP/no SC local	PCOMP/no SC UT
75110	Wheat	89.1	77.3	6.6	88.3	86.9	89.1	68.4	89.2	4.5
	Non-wheat	95.6	75.8	44.8	96.0	94.8	94.8	94.4	95.2	64.3
75111	Wheat	–	–	–	88.2	88.3	89.5	53.6	88.6	10.7
	Non-wheat	–	–	–	93.3	89.7	91.3	58.7	94.4	80.1

NOTES: ROOSTER R(S) = Rank Order Optimal Signature Transformation Estimation Routine that uses subclass means.
UT = untransformed statistics from training segment used in classification.
PCOMP/SC = principal component transformed and rescaled data.
SC = rescaled data.
Local = local statistics used in classification.

TABLE 4.– WHEAT AND NONWHEAT PERCENTAGES IN SEGMENT 1183 WITH VARIOUS STATISTICS AND DATA TRANSFORMATIONS

Acquisition	Crop	Proportion, %								
		Original Landsat data			Transformed Landsat data					
		Local	ROOSTER R(S)	UT	PCOMP/SC local	PCOMP/SC UT	SC local	SC UT	PCOMP/no SC local	PCOMP/no SC UT
75110	Wheat	55.1	43.1	9.1	53.7	49.2	54.7	30.8	54.0	8.6
	Non-wheat	35.2	39.6	22.6	37.8	42.5	35.8	55.9	39.1	26.9
75111	Wheat	—	—	—	51.2	55.7	52.0	40.1	50.4	4.3
	Non-wheat	—	—	—	40.7	36.0	39.4	11.9	41.1	67.7

NOTES: ROOSTER R(S) = Rank Order Optimal Signature Transformation Estimation Routine that uses subclass means.
UT = untransformed statistics from training segment used in classification
PCOMP/SC = principal component transformed and rescaled data.
SC = rescaled data.
Local = local statistics used in classification.

classification purposes and the training statistics were transformed with the Rank Order Optimal Signature Transformation Estimation Routine (ROOSTER).

Table 3 shows the training field accuracy for segment 1183 using original and transformed Landsat data. The local training field classification accuracy obtained by using the two-dimensional PCOMP transformed Landsat data is almost the same as the accuracy obtained by using the original Landsat data. The same observation can be noted in table 4 when proportions obtained by using the two-dimensional PCOMP transformed and original Landsat data are compared. Based on this very good classification performance of PCOMP transformed data, the transformation can be used as a feature selector.

The classification performances presented in tables 3 and 4 indicate that PCOMP transformation followed by rescaling of the data resulted in excellent signature extension performance. When the PCOMP transformation was used alone, the signature extension performance was bad; however, the signature extension performance was fair when scaling of the data alone was used. Apparently scaling of the data corrected for some of the atmospheric differences between acquisitions 75110 and 75111, but the PCOMP/SC transformation corrected for all differences between the acquisitions.

4. CONCLUSIONS

1. The first few principal components corresponding to the largest eigenvalues contain most of the information of a multispectral Landsat image. The local classification results indicate that the PCOMP transformation is an excellent feature selector which is in agreement with the results obtained by Ready and Wintz (ref. 1).

2. Based on the examination of two analyst interpreters (AI's), the multitemporal principal component false color imagery may be helpful for selecting training fields in LACIE.

3. A PCOMP transformation followed by a rescaling of the data gave excellent signature extension results; therefore, it is assumed that most of the atmospheric differences between acquisitions 75110 and 75111 of segment 1183 were eliminated. Apparently a significant part of this was due to scaling alone, since scaling of the original data caused a significant improvement in the PCOMP/SC signature extension performance. Further studies in this area are recommended.

5. REFERENCES

1. Ready, P. J.; and Wintz, P. A.: Information Extraction, SNR Improvement, and Data Compression in Multispectral Imagery. IEEE Transactions on Communications, no. 10, Oct. 1973.

2. Morris, W. L.; Wiginton, C. L.; and Lowell, D. K.: SYMAT, COVAR-Test Procedures for Matrix Calculations. University of Houston, Department of Mathematics, report no. 17, contract NAS 9-12777, Oct. 1972.

3. Kauth, R.: The Tasselled Cap Revisited. TF3-75-5-190, May 1975.

4. Finley, D. R.; and Coon, J. E.: Consecutive Day Acquisition Performance Test of Signature Correction Algorithms. Lockheed Electronics Company, Inc., LEC-6880, Sept. 1975.

5. Abotteen, R. A.: Test and Evaluation of Rank Order Optimal Signature Transformation Estimation Routine (ROOSTER). Lockheed Electronics Company, Inc., LEC-8067, Apr. 1976.