Quarterly Status and Technical Progress Report

Landsat-D Thematic Mapper Image Dimensionality Reduction and Geometric Correction Accuracy

NASA Contract Number NAS5-27577

Principal Investigator
Gary E. Ford
Department of Electrical and Computer Engineering
University of California, Davis
Davis, California 95616

(E84-10149) LANDSAT-D THEMATIC MAPPER IMAGE DIMENSIONALITY REDUCTION AND GEOMETRIC CORRECTION ACCURACY Quarterly Status Technical Progress Report, 3 Dec. 1983 - 3 Mar. 1984 (California Univ.) 7 p
N84-26097
Unclas
G3/43 00149

For the period: December 3, 1983 to March 3, 1984

1. Problems

The progress of the investigation has been impeded during this reporting period by the heavy teaching and graduate advising load of the principal investigator. Approval has been obtained to take a leave of absence to devote full time to this project for the remainder of the contract period.

2. Accomplishments

The thematic mapper scene that is the most useful in our investigation was received during the reporting period. This is a scene from path 27, row 39 of central Texas (Scene ID E-40193-16315, acquired January 25, 1983), covering the Walnut Creek watershed east of Austin, Texas. We have studied this watershed previously for the U.S. Army Corps of Engineers (Ford, et al., 1983) using MSS data. We have a substantial collection of ground truth data for this study area, including color infrared aerial photography, USGS topographic maps, and the results of manual and computer classification studies.

We have made a preliminary study of a 1024 by 1024 subscene of this image, centered on Austin, Texas. The ranges of intensity values in all of the TM bands were found to be substantially greater than those we had observed in the other TM image available to us, a February 2, 1983 acquisition centered near Sacramento, California. Since both of these images were acquired during the winter months, the low sun angle would be expected to produce low dynamic ranges. However, we are unable to provide an explanation as to why there is such a disparity between the two images, acquired just seven days apart. We observed a 50% increase in standard deviation in the intensities in bands 1, 2, 3, 5, and 7; no change in band 4; and a 25% decrease in band 6. The decrease in the thermal band is probably due to the low percentage of water in the Texas image. Due to the heavy rainfall in California during the 1982-83 winter, the study area of interest to us on the Sacramento River was extensively flooded on the date of acquisition of the image and we are unable to perform a classification study on this image. While the late January date is not the optimum for a classification study of the Walnut Creek watershed, an interactive analysis of the image shows that the major land use classes exhibit spectral separation.

Principal components transformation has been applied to the Walnut Creek subscene to reduce the dimensionality of the multispectral sensor data. The results are summarized in Table 1. For comparison purposes, we have also applied this transformation to a Landsat 3 MSS subscene of the same area. The MSS scene was acquired in a different season and year (May 1978), but the results, as summarized in Table 2, allow for comparisons between TM and MSS data.

The correlation matrices indicate the pairwise correlation of the spectral components. The TM correlation matrix shows that visible bands 1-3 exhibit a high degree of correlation, in the range 0.92 to 0.96. While it might be expected that the reflective infrared bands 4, 5, and 7 would be highly correlated, this is only true for bands 5-7, where the correlation is 0.93. Band 4 is not highly correlated with any other band, with correlations in the range 0.13 to 0.52. Similarly, the thermal band 6 is not highly correlated with other bands, with correlations in the range 0.13 to 0.46. It should be noted that bands 1-3 exhibit a moderate degree of correlation with bands 5 and 7, with correlations in the range 0.63 to 0.83. These results are significantly different than those we reported for the Sacramento subscene in our project report for the period June 3, 1983 to September 3, 1983. This is a further indication that the Sacramento scene is not representative of data from the TM sensor.

The MSS correlation matrix shows that the visible bands (4 and 5) are highly correlated (0.96), as are the reflective infrared bands (6 and 7), with a correlation of 0.92. On the basis of this comparison between TM and MSS data, we are unable to support the expectation that there would be lower correlation among the TM spectral components than has been observed for the MSS spectral components.

The principal components transformation matrix is composed of the normalized eigenvectors of the covariance matrix ordered by eigenvalues. The weights used to generate a transformed component are given in the columns of the eigenvector matrices in tables 1 and 2. The first transformed component of the TM data is a weighted average of the seven spectral components. These weights are approximately proportional to the standard deviations of the spectral intensities. The second transformed component is roughly the sum of bands 1-4 minus the sum of bands 5-7. The third component is roughly band 4 minus bands 1-3 and 7, with negligible weights given to bands 5 and 6. Band 4 dominates this transformed component due to its relatively low correlation with all other bands. Again, these conclusions differ significantly from those based on the Sacramento scene.

For the MSS data, the first transformed component is again a weighted average of the four original components, with the weights roughly proportional to standard deviations of the spectral intensities. This component has been referred to as the brightness component since it is a sum of the original components. The second transformed component is roughly the difference between the visible and reflective infrared components. It has been referred to as the greenness component, as it exhibits a strong response to vegetation.

If the data is represented by L transformed components of the original N components, where $L<N$, then the mean-squared error of the representation is given by the sum of the N-L eigenvalues of the components not included in the representation. The percent variance in any transformed component is given by the ratio of the eigenvalue for that component to the sum of the eigenvalues. For the TM image, it is seen that nearly 80% of the variance is contained in the first transformed component. Transformed components 2 and 3 represent 9.7% and 7.9% of the variance, respectively. The variance in the remaining components is at or below 1%. Representing the image by the first three components maintains 97.4% of the variance in the original seven components. For the MSS data, 98.6% of the variance is contained in the first two transformed components. The results in both cases support the statement by Gnanadesikan (1977) that at least 95% of the variance will be contained in the first three transformed components.

The results for the TM data are in close agreement with those reported by Bernstein and Lotspiech (1983). Working with a subscene of the San Francisco Bay acquired on December 31, 1982, they also performed principal component processing to reduce the dimensionality of the data. Their data exhibited lower spectral standard deviations, possibly due to the even lower sun angles at that time of the year. The correlation results are similar, with two exceptions. The correlations between bands 4 and 5 and bands 4 and 7 are higher than that observed for the Texas scene, at 0.88 and 0.81, respectively. The thermal data in band 6 has a small negative correlation with bands 2-5 and 7, which was not observed in the Texas scene. Both of these differences are probably due to the large amount of water in the San Francisco Bay scene. Due to the observed differences in the correlation matrix, the weights used in forming the second and third transformed components differ considerably from those observed for the Texas scene.

The principal components transformation of the Texas image is also in close agreement with the preliminary formulation of the thematic mapper tasseled cap transformation developed by Crist (1983). The first principal component is similar to the brightness feature, which is also a weighted average. The second principal component is similar to Crist's third component, tentatively termed wetness, in that the weights are positive for original bands 1 to 4 and negative for bands 5 and 7. The second principal component is similar to Crist's greenness component, with a large positive weight for band 4, negative weights for bands 1-3 and 7, and a small weight for band 5. Thus, the physical interpretation given by Crist may apply to the Texas scene.

References:

(1) Berstein, R. and Lotspiech, J. B. (1983). "Landsat-4 thematic mapper sensor evaluation and advanced information extraction experiments," Proceedings of the Eighth Pecora Memorial Remote Sensing Symposium (preprint copy).

(2) Crist, E. P. (1983). "The thematic mapper tasseled cap -- a preliminary formulation," Proceedings of the 1983 Machine Processing of Remotely Sensed Data Symposium, pp 357-364.

(3) Ford, G. E., Algazi, V. R., and Meyer, D. I. (1983). "A noninteractive procedure for land use determination," Remote Sensing of Environment, vol. 13, pp 1-16.

(4) Gnanadesikan, R. (1977). Methods for Statistical Data Analysis of Multivariate Observations, New York: Wiley.

3. Significant Results

The significant results obtained this quarter are those involving principal components analysis, as reported in section 2.

4. Publications

No publications were released during this period.

5. Recommendations

No recommendations are made concerning changes relating to maximum utilization of the Landsat-D system.

6. Funds Expended

Salary - Research Assistant

Employment Benefits

Clerical

Computer Time

Table 1. TM Principal Components Transformation

MEANS

77.46	28.63	31.97	41.88	61.89	102.48	28.54

STANDARD DEVIATION

8.897	5.042	7.630	8.663	17.878	2.826	9.906

CORRELATION MATRIX

1.0000	0.9574	0.9273	0.4420	0.6370	0.1701	0.7515
0.9574	1.0000	0.9478	0.5174	0.6803	0.2070	0.7791
0.9273	0.9468	1.0000	0.4885	0.7463	0.2646	0.8333
0.4420	0.5174	0.4885	1.0000	0.5236	0.1345	0.4288
0.6370	0.6803	0.7463	0.5236	1.0000	0.4666	0.9273
0.1701	0.2070	0.2646	0.1345	0.4666	1.0000	0.4445
0.7515	0.7791	0.8333	0.4288	0.9273	0.4445	1.0000

EIGENVECTORS

0.3052	0.6309	-0.2189	-0.2135	0.1062	-0.5668	-0.2877
0.1807	0.3275	-0.0664	-0.0200	0.0635	0.0392	0.9218
0.2887	0.3996	-0.1475	0.0267	0.1526	0.8044	-0.2533
0.2208	0.2289	0.9320	0.1518	-0.0622	-0.0305	-0.0486
0.7543	-0.5177	0.0196	-0.3745	0.1472	-0.0191	0.0200
0.0520	-0.1092	-0.0310	0.5739	0.7965	-0.1434	-0.0100
0.4128	-0.0521	-0.2364	0.6787	-0.5491	-0.0913	-0.0230

EIGENVALUES

529.48	64.35	52.60	6.76	5.40	3.83	1.13

PERCENT VARIANCE IN TRANSFORMED COMPONENTS

79.8	9.7	7.9	1.02	0.81	0.58	0.17

CUMULATIVE PERCENT VARIANCE IN TRANSFORMED COMPONENTS

79.8	89.5	97.4	98.4	99.2	99.8	100.0
Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7

Table 2. MSS Principal Components Transformation

MEANS			
25.48	28.10	52.87	26.17
STANDARD DEVIATION			
5.621	8.852	9.455	4.992
CORRELATION MATRIX			
1.0000	0.9616	0.5484	0.2529
0.9616	1.0000	0.5170	0.2186
0.5484	0.5170	1.0000	0.9164
0.2529	0.2186	0.9164	1.0000
EIGENVECTORS			
0.3780	0.3490	0.7900	0.3335
0.5886	0.6052	-0.5349	-0.0336
0.6618	-0.5525	0.1337	-0.2683
0.2696	-0.4546	-0.2683	0.8054
EIGENVALUES			
162.15	59.18	1.72	1.21
PERCENT VARIANCE IN TRANSFORMED COMPONENTS			
72.3	26.3	0.77	0.54
CUMULATIVE PERCENT VARIANCE IN TRANSFORMED COMPONENTS			
72.3	98.6	99.5	100.0
Band 1(4)	Band 2(5)	Band 3(6)	Band 4(7)