NASA TECHNICAL NOTE

NASA TN D-8489

NASA TN D-8489

DATA-HANDLING CONSIDERATIONS FOR ADVANCED HIGH-RESOLUTION MULTISPECTRAL SCANNERS

Sidney L. Whitley

Lyndon B. Johnson Space Center
Houston, Texas 77058

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MAY 1977

1. Report No. NASA TN D-8489	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DATA-HANDLING CONSIDERATIONS FOR ADVANCED HIGH-RESOLUTION MULTISPECTRAL SCANNERS		5. Report Date May 1977
		6. Performing Organization Code JSC-11778
7. Author(s) Sidney L. Whitley		8. Performing Organization Report No. S-472
9. Performing Organization Name and Address Lyndon B. Johnson Space Center Houston, Texas 77058		10. Work Unit No. 177-52-83-00-72
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		13. Type of Report and Period Covered Technical Note
		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The purpose of this study was to determine the scope of the data-handling problems for advanced high-resolution multispectral scanner systems. Several scanner design parameters that greatly affect data rates were varied to determine overall effect on data rates. Bit rates were calculated for scanners with 10-, 15-, 20-, and 30-meter instantaneous fields of view; with 5, 7, and 10 channels; and with swath widths of 46.3, 92.6, 138.9, and 185.2 kilometers (25, 50, 75, and 100 nautical miles). The data were analyzed to determine whether they could be (1) recorded onboard the spacecraft with currently available equipment, (2) recorded onboard with the equipment proposed for the 1980-90 period, (3) transmitted to Earth with current or proposed equipment, and (4) compressed and decompressed without significant degradation of classification accuracy. The effect of classifying 10-meter scanner data on the existing and the developing pattern-recognition systems was examined.

17. Key Words (Suggested by Author(s)) Data processing Multispectral scanner Pattern recognition Data compression Instantaneous field of view	18. Distribution Statement STAR Subject Categories: 35 (Instrumentation and Photography) 43 (Earth Resources)

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 22	22. Price* $3.50

* For sale by the National Technical Information Service, Springfield, Virginia 22161

DATA-HANDLING CONSIDERATIONS FOR ADVANCED HIGH-RESOLUTION MULTISPECTRAL SCANNERS

By Sidney L. Whitley
Lyndon B. Johnson Space Center

SUMMARY

Between 1980 and 1990, spaceborne multispectral scanners will probably be required to have resolution cell sizes as small as 10 meters, as many as 10 channels, and swaths as wide as 185.2 kilometers (100 nautical miles). The data produced by these multispectral scanners may be transmitted to Earth in real time, recorded and retained onboard the spacecraft, or recorded for subsequent transmission. Real-time transmission of the data will require extensive development of transmission and receiving equipment. Onboard recording of data during manned space flights appears to present fewer developmental problems. Data-handling problems for advanced multispectral scanners can be greatly reduced by using data-compression techniques.

Data from a 30-meter instantaneous field of view, 10-band scanner with 185.2-kilometer (100 nautical mile) swaths may be recorded onboard the manned spacecraft or transmitted from satellites with current state-of-the-art equipment. However, similar multispectral scanners with a 10-meter instantaneous field of view will present a challenge during the 1980's. Additional research is needed in the areas of quantitative data band selection, tape recorders, data compression, swath-width selectability, and data-processing procedures and systems.

INTRODUCTION

Background

The relatively new field of remote sensing has been recognized for its potential by numerous sciences and disciplines. One remote sensor, the multispectral scanner, has had a very great effect and appears to hold much potential for future application. Multispectral scanner technology is growing rapidly. Considerable progress has been made in detector sensitivity, spectral selectability, mechanical arrangement of components, data-sampling techniques, data-encoding techniques, data-recording techniques, transmission techniques, and data-processing techniques. Existing multispectral scanners, such as those used on Landsat (formerly the Earth resources technology satellite (ERTS)) and on aircraft, led to the emergence of a large community of multispectral data users. These users have developed and demonstrated numerous applications of remotely sensed data.

During the next few years (in the 1980's), many advanced high-resolution multispectral scanners will be proposed for manned and unmanned spacecraft missions. The users may desire, if not require, instantaneous fields of view as small as 10 meters and scanner swath widths of at least 185.2 kilometers (100 nautical miles).

Objective

The purpose of this report is to describe the results of an investigation to determine the scope of the data-handling problems imposed by advanced high-resolution multispectral scanners. The specific questions addressed in this study are as follows.

1. What are the potential configurations of multispectral scanners for 1980-90? What instantaneous fields of view, number of channels, data-recording precisions, and scanner swath widths are likely to be required?

2. What potential data bit rates can the various multispectral scanner configurations produce?

3. What is the current state of the art in onboard pulse code modulation (PCM) data recording?

4. What will the probable onboard data-recording rate be by 1980-90?

5. What is the maximum data transmission rate at current receiver capability, and what would be the effect of high-resolution scanners on the existing data-receiving network?

6. How well could existing and advanced data-processing and analysis systems handle the high-resolution data?

7. What are the costs of classifying high-resolution multispectral data using selected automated data systems?

8. What areas of research must be pursued to handle advanced multispectral scanner system data adequately?

Approach

The approach used for this study was to define the probable scanner configurations in terms of instantaneous field of view (IFOV), number of channels, swath width, and data-recording precision. Data bit rates were calculated for the various scanner configurations. These bit rates were evaluated to determine the feasibility of recording the data with the existing equipment and the feasibility of transmitting and receiving the data at a ground receiving station.

Various alternatives were studied to determine whether the data should be recorded or transmitted in the original density or whether they should be compressed before recording or transmitting. The problems associated with ground data handling were also considered to determine whether the existing techniques were adequate for processing and analyzing multispectral scanner data. Costs of classifying a 185.2-kilometer-square (100 nautical mile square) using selected current data-processing techniques were projected.

SCANNER CHARACTERISTICS

Because several configurations may be considered for spaceborne advanced multispectral scanners, a range of instantaneous fields of view, of numbers of channels, and of swath widths was examined.

Scanner Instantaneous Field of View

Generally, very high resolution surveys are conducted over small rapidly changing areas. The most frequently mentioned instantaneous fields of view (resolution cell on the Earth's surface) for advanced multispectral scanners are those measuring 10 to 30 meters. Scanner systems with an IFOV of 10, 15, 20, and 30 meters are considered so that data will be readily available if a trade-off becomes necessary.

Scanner Swath Width

Unlike the Landsat scanner, high-resolution scanners may not be required to cover the full 185.2-kilometer (100 nautical mile) swath; therefore, swath widths of 46.3, 92.6, 138.9, and 185.2 kilometers (25, 50, 75, and 100 nautical miles) were included in this study.

Scanner Channel Assignments

Numerous studies have been conducted to provide information for the selection of multispectral scanner bands for the Earth observations satellite. No quantitative approach exists for determining the best selection of data bands for a multispectral scanner system. The techniques currently used produce atypical choices of data bands. It is likely that a quantitative method for channel selection could be developed over a 3- to 5-year period. This length of time is required because data should be considered throughout the visible, near-infrared, intermediate-infrared, and thermal-infrared portions of the electromagnetic spectrum for materials to be identified by the scanner data. Real data sets should be used in the band-selection study rather than discrete master (typical) signatures. Experience has shown that unique spectral signatures do not exist for a given material but that the signatures under various conditions produce a family of signatures. To determine whether various surface materials are distinguishable by spectral-pattern recognition, the variations in signatures must be considered, and those bands that provide the greatest separability by spectral-pattern recognition or by other classification techniques must be identified.

Number of Bands

The number of multispectral scanner data bands required to produce a reasonably accurate surface classification map is fairly small. A few well-chosen data bands generally produce classification maps with almost as much accuracy as a large number of data bands. Coggeshall and Hoffer (ref. 1) reported that, for a set of multispectral scanner data collected over a forest area, certain classification accuracies were possible with selected data bands (table I).

TABLE I.- CLASSIFICATION ACCURACY
COMPARED TO NUMBER OF CHANNELS
USED IN SPECTRAL-PATTERN-
RECOGNITION ANALYSIS

Band no.	Overall accuracy, percent
1	44.0
2	80.5
3	87.1
4	89.3
5	90.8
6	92.4
8	93.7
10	94.7
12	95.1

17Apr78
ja

ERRATA

NASA Technical Note D-8489

DATA-HANDLING CONSIDERATIONS FOR ADVANCED HIGH-RESOLUTION MULTISPECTRAL SCANNERS

By Sidney L. Whitley
May 1977

Page 5: A corrected page 5 is attached for substitution in place of the existing page 5 on which portions of the text appear in the wrong order.

Issued January 1978

Each set of bands used by Coggeshall and Hoffer was the best subset of data bands, as determined by a divergence computation. The values presented in table I are typical and have been observed in other analyses.

The scanner data should be digitized such that one 8-bit data word is acquired in each data band for each ground resolution cell (IFOV) across the swath. Samples should be collected simultaneously in all bands so that the data will be perfectly registered.

Also, an appropriate number of readings should be digitized as the scanner views its calibration sources. Overhead information, such as a synchronization pattern, is required to indicate the beginning of a scan line or data frame. For scanners producing numerous ground resolution cells per scan line, several subframe synchronization patterns should be included so that reinitialization can be achieved if synchronization is lost. Additionally, it is advisable to include some preamble bits to assist ground processing systems in locating the start of the frame.

It is difficult to justify the use of numerous multispectral scanner data bands solely by surface-classification accuracy, because the increase in accuracy is nonlinear with respect to the number of bands used. Four well-chosen bands may be used to produce a classification map almost as accurate as one produced with 12 bands. However, it is desirable to have more than the minimum number of data bands available so that the optimum bands can be chosen (by mathematical calculation) for preparation of a surface-classification map and for other products that may be derived from them. The best estimate is that approximately 7 to 10 data bands will prove adequate for multispectral scanners to be flown in the 1980-90 missions.

DATA HANDLING

Most multispectral scanners are designed such that a mirror traverses the ground scene at the nadir of the data-collection platform, normal to the direction of flight. As the mirror traverses the ground scene, it gathers energy from a wide spectral range within a well-defined surface area that is frequently called a ground resolution cell. The dimensions of the ground resolution cell can be, for example, 10 by 10 meters. As the scanner mirror rotates, it is directed toward the ground scene for some defined swath width, usually expressed in degrees of angular excursion to either side of the platform nadir or in kilometers (nautical miles) of ground-surface coverage centered along the flightpath. Because the scanner observes the ground scene for only a small portion of a complete revolution of the scanner mirror, considerable "dead time" is experienced during which the mirror is outside the active data-collection portion of the scan. During this dead time, it is advantageous to view calibration sources through the collecting optics of the scanner. At all other times in a mirror revolution, the scanner will collect information that is of no interest to the data user. Sampling should be discontinued during dead times, except for calibration. Synchronization patterns, calibrations, and preamble bits will occur in variable-size groups of bits producing variable data bit rates advantageous to view calibration sources through the collecting optics of the scanner. At all other times in a mirror revolution, the scanner will collect information that is of no interest to the data user. Sampling should be discontinued during dead times, except for calibration.

The scanner data should be digitized such that one 8-bit data word is acquired in each data band for each ground resolution cell (IFOV) across the swath. Samples should be collected simultaneously in all bands so that the data will be perfectly registered.

Also, an appropriate number of readings should be digitized as the scanner views its calibration sources. Overhead information, such as a synchronization pattern, is required to indicate the beginning of a scan line or data frame. For scanners producing numerous ground resolution cells per scan line, several subframe synchronization patterns should be included so that reinitialization can be achieved if synchronization is lost. Additionally, it is advisable to include some preamble bits to assist ground processing systems in locating the start of the frame. Synchronization patterns, calibrations, and preamble bits will occur in variable-size groups of bits producing variable data bit rates

Each set of bands used by Coggeshall and Hoffer was the best subset of data bands, as determined by a divergence computation. The values presented in table I are typical and have been observed in other analyses.

It is difficult to justify the use of numerous multispectral scanner data bands solely by surface-classification accuracy, because the increase in accuracy is nonlinear with respect to the number of bands used. Four well-chosen bands may be used to produce a classification map almost as accurate as one produced with 12 bands. However, it is desirable to have more than the minimum number of data bands available so that the optimum bands can be chosen (by mathematical calculation) for preparation of a surface-classification map and for other products that may be derived from them. The best estimate is that approximately 7 to 10 data bands will prove adequate for multispectral scanners to be flown in the 1980-90 missions.

DATA HANDLING

Most multispectral scanners are designed such that a mirror traverses the ground scene at the nadir of the data-collection platform, normal to the direction of flight. As the mirror traverses the ground scene, it gathers energy from a wide spectral range within a well-defined surface area that is frequently called a ground resolution cell. The dimensions of the ground resolution cell can be, for example, 10 by 10 meters. As the scanner mirror rotates, it is directed toward the ground scene for some defined swath width, usually expressed in degrees of angular excursion to either side of the platform nadir or in kilometers (nautical miles) of ground-surface coverage centered along the flightpath. Because the scanner observes the ground scene for only a small portion of a complete revolution of the scanner mirror, considerable "dead time" is experienced during which the mirror is outside the active data-collection portion of the scan. During this dead time, it is (e.g., no data are collected during dead times). The data should be time-buffered so that a continuous stream of data bits is produced at a much reduced data rate. The time-buffered data are recorded on tape or telemetered to ground receiving stations.

Bit Rates Produced by Selected Sanner Configurations

Data bit rates were calculated for the various configurations of multispectral scanners described previously. It was assumed for these calculations that all data had been time-buffered, as described previously. It was also assumed that the groundspeed would be 7.4 km/sec (4 n. mi./sec) and that all data would be digitized to 8-bit precision. The data bit rates include the active data-collection portion of the scan, calibration samples, synchronization patterns, subframe synchronization patterns, and preamble bits to aid ground-station equipment in decoding the data. The resulting data bit rates are presented in table II.

TABLE II.- DATA BIT RATES PRODUCED BY SELECTED SCANNERS[a]

IFOV, m	Bit rate, megabits/sec											
	10-channel swath width, km (n. mi.)				7-channel swath width, km (n. mi.)				5-channel swath width, km (n. mi.)			
	185.2 (100)	138.9 (75)	92.6 (50)	46.3 (25)	185.2 (100)	138.9 (75)	92.6 (50)	46.3 (25)	185.2 (100)	138.9 (75)	92.6 (50)	46.3 (25)
10	1103	828	552	276	772	579	386	193	552	414	276	138
15	490	368	245	[123]	343	258	172	[86]	245	185	[123]	62
20	277	208	[136]	[69]	194	[146]	[97]	[49]	[139]	[105]	[70]	35
30	[123]	[93]	[62]	[31]	[86]	[65]	[43]	[22]	[62]	[47]	[31]	16

[a] Bracketed entries are within current tape-recording state-of-the-art capabilities.

Onboard Data Recording

The data produced by multispectral scanners on manned spacecraft, such as the Space Shuttle, should probably be recorded onboard using direct recording techniques in cases in which the data are not required immediately. The data tapes recorded onboard the spacecraft could be returned to Earth as on conventional aircraft missions.

Current tape-recording state of the art can maintain reasonable recording and playback accuracy (1 bit error in 10^8) at a maximum of 13 000 bits/cm/tape track (33 000 bits/in/tape track) in biphase level or extended nonreturn-to-zero (NRZ) PCM format, with 42 tracks on a 2.54-centimeter-wide (1 inch wide) tape (the present capability of one commercial magnetic tape system). This recording capability translates to a bit rate of 166.32 megabits/sec for 3.05 m/sec (120 in/sec) tape-movement speed. Clearly, state-of-the-art tape recording is inadequate for multispectral scanners with 10-meter IFOV's, 10 channels, and a 185.2-kilometer (100 nautical mile) swath width. Improvement by a factor of 7 over present capability would be required to record data from this worst-case scanner configuration.

The rates in table II vary from 1103 megabits/sec for a 10-meter IFOV, 10-channel, 185.2-kilometer (100 nautical mile) swath-width scanner to 16 megabits/sec for a 30-meter IFOV, 5-channel, 46.3-kilometer (25 nautical mile) swath-width scanner.

The data in table II indicate that several scanner configurations are within the capabilities of current recording technology. For instance, the 30-meter IFOV, 10-channel, 185.2-kilometer (100 nautical mile) swath-width scanner produces a bit rate of 123 megabits/sec, which is within present tape-recording capability. Also, the 20-meter IFOV, 5-channel, 185.2-kilometer (100 nautical mile) swath-width scanner produces a bit rate of 139 megabits/sec, which is also within the present tape-recording capability of 166.32 megabits/sec. Bracketed entries in each column of table II are within current tape-recording capability.

The data bit rate produced by the Landsat-A and Landsat-B multispectral scanner is 15.06 megabits/sec, and the bit rate for the Skylab S192 multispectral scanner was 23.30 megabits/sec. Discussions with representatives of tape-recording manufacturers indicate that it is reasonable to expect an improvement (by a factor of 2.4) in tape-recording capabilities by 1980 or 1982 to a data bit rate of approximately 400 megabits/sec. This enhanced capability will depend on improvement of tape-recording and tape-reading technology and, possibly, on improvement of tape-movement speed. An examination of the data rates in table II reveals that several scanner configurations exceed even the tape-recording capability projected for 1980-82. For instance, the 15-meter IFOV, 10-channel, 185.2-kilometer (100 nautical mile) swath-width scanner exceeds the projected 400-megabit/sec rate by 90 megabits/sec.

Data Transmission

Data transmission will be a requirement for unmanned satellites and may be considered for certain applications in manned spacecraft. In the worst-case scanner configuration (10-meter, 10 channels, 185.2-kilometer (100 nautical mile) swath width), the data bit rate is 1103 megabits/sec. This rate exceeds the capability of the existing space tracking and data acquisition network (STDN) and the tracking and data relay satellite system (TDRSS) planned for 1980.

Examination of the STDN downlink capabilities indicates a data-transmission rate of 128 kilobits/sec for one link and 256 kilobits/sec for another link. Data dump capability is 1.02 megabits/sec (time shared with other downlinks).

The data given in table II indicate that all scanner configurations exceed the current STDN capabilities. These STDN data rates are candidates for use in the Space Shuttle Program.

The STDN receiving equipment could be modified to accept the 1103-megabit/sec rate, but the present bandwidth allocation would require redesignation or combination or both. Estimates for such a modification to the entire STDN are approximately $2.6 to $3.1 million.

The TDRSS downlink is planned to have a capability for handling a data rate of 50 megabits/sec. Table II indicates that data from a few of the scanner configurations (i.e., a 20-meter, 7-channel, 46.3-kilometer (25 nautical mile) swath-width scanner) can be transmitted over the TDRSS downlink. Data compression will enable transmission of data from many scanner configurations on the TDRSS downlink. Data-compression techniques are discussed in subsequent paragraphs.

It is technically possible to transmit data at 1103 megabits/sec and to maintain an acceptable signal-to-noise ratio, but the existing and planned ground-based tracking and receiving stations would require extensive and expensive modification to receive this volume of data. An almost identical problem exists in recording high bit rate data onboard the spacecraft or at ground receiving stations. Onboard recorders must be lightweight and reliable, use little power, and produce little heat, whereas ground-based recording systems must be capable of receiving a weaker and noisier signal. In either case, many of the data-recording rates presented in table II exceed the bit rates that can be recorded on a single magnetic tape. It is strongly recommended that splitting the data stream and recording it on two or more tape recorders be avoided. At the bit rates shown in table II, it is operationally difficult to remove the skew in a single tape when two recording heads are used; the problem would be even more difficult, or perhaps impossible, if the data were recorded on two or more tapes.

Assessment of Onboard Recording and Transmission

The decision to record or transmit data from multispectral scanners should be related to, if not determined by, the intended application of the data. If the data are urgently needed to survey some dynamic condition, they should be transmitted. However, if the data will be used exclusively for studies in which the data return date is not overly critical, it is advisable to record the data onboard the spacecraft.

When data are recorded onboard the spacecraft, as opposed to being transmitted, the recorded signal has a higher signal-to-noise ratio and is less likely to contain noise than the transmitted signal. Reliability should be somewhat higher because fewer components are involved in data handling; i.e., no transmitter and no ground receiver are needed. Also, the position of the spacecraft with respect to ground receiving stations is of no concern. The onboard data-handling system will be very expensive because it must have high reliability and must operate in a more hostile environment than the ground-based receiving station (with the exception of locally inclement weather). The facilities for maintenance of the onboard data system will be much less elaborate than at ground receiving stations, and thus onboard maintenance capability will be more limited. In some cases, it may be desirable to tape record the multispectral scanner data onboard the spacecraft in real time. After the data-collection period is over, and perhaps when the lighting conditions or groundtrack is unsuitable for further data collection, the data may be played back from the onboard tape for transmission to ground receiving stations at reduced data rates. A reduced playback rate could be chosen such that the present state-of-the-art facilities could be used for transmission, receiving, and data recording.

If the data from high-resolution advanced scanners are to be transmitted to Earth in real time, the ground receiving equipment may require extensive modification. Whether these modifications are required will depend on the scanner resolution used, on the swath width required, and on whether the data can be compressed before transmission.

Several positive and negative aspects of both onboard recording and data transmission must be considered. A decision as to the data-handling approach must be made for each application.

DATA COMPRESSION

Before seriously considering the actual compression of data, one should consider whether the full number of data bands is warranted. As shown in table I, the improved accuracy achieved by the addition of data bands comes slowly and expensively in terms of sensor complexity and data processing. Therefore, the accuracy required for the survey should be considered.

Studies have shown that multispectral scanner data can be compressed. The data can be compressed spectrally because of a significant degree of correlation between bands or channels and spatially because of the frequent similarity of adjacent ground-resolution cells.

Crawford et al. (ref. 2) have shown that principal component analysis and canonical analysis could effectively reduce the dimensionality of multispectral scanner data. It was shown that, when a 9-band set of data was translated by canonical analysis, the first three or four eigenvalues (dimensions in the new basis) contained the greatest share of the information content, as shown in table III.

Some promising research work on the compression of multispectral scanner data by principal component analysis is described in reference 3. This work seems particularly applicable for the compression of advanced multispectral scanner data and is discussed in the following paragraphs.

TABLE III.- USE OF CANONICAL ANALYSIS FOR TRANSFORMATION TO REDUCE DIMENSIONALITY OF DATA

Factor	Eigenvalue
1st	53.7
2d	2.2
3d	1.6
4th	1.4
5th	.8
6th	.7
7th	.6

Approach to Data Compression

Although several approaches can be used for multispectral scanner data compression, only one approach to spectral compression will be addressed; i.e., reduction of the dimensionality of the original set of multispectral scanner data. A data point describing a ground resolution element may be expressed as vector

$$\underset{\sim}{X} = \left[X_1, X_2, \ldots, X_N\right]^t \qquad (1)$$

where N is the dimensionality of the original data, or the number of bands in the multispectral scanner. The data are then projected onto a new lower dimensional basis (n < N) by the transformation $\underset{\sim}{T}$,

$$\underset{\sim}{Y} = \left(\underset{\sim}{T}\,\underset{\sim}{X} - \underset{\sim}{U}\right) \qquad (2)$$

where $\underset{\sim}{U}$ is defined as

$$E\left\{\left[X_1, X_2, \ldots, X_N\right]^t\right\} \qquad (3)$$

Here, E is the expectation operator, $\underset{\sim}{U}$ is the mean vector of $\underset{\sim}{X}$, and $\underset{\sim}{T}$ is the n × N matrix in which the n rows are the first n eigenvectors corresponding to the n largest eigenvalues of the covariance matrix $\underset{\sim}{C}$ of $\underset{\sim}{X}$.

If $\underset{\sim}{C} = E[(\underset{\sim}{X} - \underset{\sim}{U})\,(\underset{\sim}{X} - \underset{\sim}{U})^t]$, then the n rows of $\underset{\sim}{T}$ are the n normalized solutions of $\underset{\sim}{C}\,\underset{\sim}{V}_i - \lambda_i \underset{\sim}{V}_i$, where $i = 1, 2, \ldots n < N$ with $\lambda_1 > \lambda_2 > \lambda_3 > \ldots > \lambda_n$ and V is a vector. The covarience matrix for $\underset{\sim}{Y}$ is given by

$$E\left(\underset{\sim}{Y}\,\underset{\sim}{Y}^t\right) \quad \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & \lambda_n \end{bmatrix} \tag{4}$$

The covariance matrix for $\underset{\sim}{Y}$ is diagonal and the elements of the transformed data vector are uncorrelated. Each λ_i represents the variance of the i^{th} element of $\underset{\sim}{Y}$.

The transformed data vector $\underset{\sim}{Y}$ is projected back on the original basis by transformation $\underset{\sim}{T}^t$.

$$\underset{\sim}{\hat{X}} = \underset{\sim}{T}^t \underset{\sim}{Y} + \underset{\sim}{U} \tag{5}$$

where $\underset{\sim}{\hat{X}}$ represents the recovered data points in their original formats and are a good approximation of $\underset{\sim}{X}$.

Ready et al. (ref. 3) give a technique for approximating the mean square error of the transformation and inverse transformation described. They conclude that the mean square error is equal to the sum of N − n discarded λ_i's. For highly correlated data, the eigenvalues λ_i decrease rapidly and the error produced by discarding λ_i's is small.

When Ready et al. applied the transformations to a set of 12 bands (all in the visible portion of the spectrum) of multispectral scanner data, they found the eigenvalues shown in table IV. This set of eigenvalues suggests that the first three or four components of the coordinate system (basis) to which the data were transformed contain almost all the information of the original basis. The work of Ready et al. showed that 12 spectral bands from the visible part of the spectrum could be compressed into 4 spectral bands with a root mean square (rms) error of only 2 percent and with no discernible degradation in classification accuracy when classified by maximum-likelihood-ratio spectral-pattern-recognition techniques.

TABLE IV.- EIGENVALUES FOR DATA SETS TRANSFORMED BY PRINCIPAL COMPONENTS ANALYSIS

Factor	Eigenvalue
1st	1760
2d	600
3d	125
4th	46
5th	25
6th	20
7th	18
8th	15
9th	14
10th	13
11th	12
12th	10

Their work also showed that data compression could be achieved because of spatial redundancy or spatial correlation. They described two approaches to spatial compression and recommended one approach in which a 10:2 spatial compression with a 2-precent rms error with no discernible degradation in pattern-recognition classification accuracy could be achieved.

It should be observed that, because all the data bands used by Ready et al. were in the visible part of the spectrum, greater correlation in the original data basis and hence greater compression without classification-accuracy degradation could be expected. The compression would probably have been less impressive if the data had included visible, near-infrared, intermediate-infrared, and thermal-infrared bands.

The limit to which data can be compressed without significant degradation of usefulness is data-dependent and is directly proportional to the statistical correlation of the data. It is difficult to extrapolate the results obtained from visible data bands only to the infrared, intermediate-infrared, and thermal-infrared bands. Additional research on the compression of multispectral scanner data is required to determine the practicality of compression as a means of improving (reducing) data bit rates.

Compressed data bit rates have been calculated for the selected scanner configurations, as shown in table V. Because the limits of data compression are not known, table V contains the original data rates and the corresponding 10:5, 10:4, and 10:3 compression rates for each of the multispectral scanner configurations presented in table II.

TABLE V.- COMPRESSED DATA BIT RATES

FOR DIFFERENT IFOV SCANNER CONFIGURATIONS

Swath width, km (n. mi.)	No. of channels	Data rates, megabits/sec			
		Original measurement	10:5 compression	10:4 compression	10:3 compression
10-meter IFOV					
185.2 (100)	10	1103	552	442	331
138.9 (75)	10	828	414	330	248
92.6 (50)	10	552	276	221	166
46.3 (25)	10	276	138	111	83
185.2 (100)	7	772	386	309	232
138.9 (75)	7	579	290	232	174
92.6 (50)	7	386	193	155	116
46.3 (25)	7	193	97	78	58
185.2 (100)	5	552	276	221	166
138.9 (75)	5	414	207	166	125
92.6 (50)	5	276	138	111	83
46.3 (25)	5	138	69	56	42
15-meter IFOV					
185.2 (100)	10	490	245	196	147
138.9 (75)	10	368	184	148	111
92.6 (50)	10	245	123	98	74
46.3 (25)	10	123	62	50	37
185.2 (100)	7	343	172	138	103
138.9 (75)	7	258	129	104	78
92.6 (50)	7	172	86	69	52
46.3 (25)	7	86	43	35	26
185.2 (100)	5	245	123	98	74
138.9 (75)	5	185	93	74	56
92.6 (50)	5	123	62	50	37
46.3 (25)	5	62	31	25	19

TABLE V.- Concluded

Swath width, km (n. mi.)	No. of channels	Data rates, megabits/sec			
		Original measurement	10:5 compression	10:4 compression	10:3 compression
		20-meter IFOV			
185.2 (100)	10	277	139	111	84
138.9 (75)	10	208	104	84	63
92.6 (50)	10	139	70	56	42
46.3 (25)	10	69	35	28	21
185.2 (100)	7	194	97	78	59
138.9 (75)	7	146	73	59	44
92.6 (50)	7	97	49	39	30
46.3 (25)	7	49	25	20	15
185.2 (100)	5	139	70	56	42
138.9 (75)	5	105	53	42	32
92.6 (50)	5	70	35	28	21
46.3 (25)	5	35	18	14	11
		30-meter IFOV			
185.2 (100)	10	123	62	50	37
138.9 (75)	10	93	47	38	28
92.6 (50)	10	62	31	25	19
46.3 (25)	10	31	16	13	10
185.2 (100)	7	86	43	35	26
138.9 (75)	7	65	33	26	20
92.6 (50)	7	43	22	18	13
46.3 (25)	7	22	11	9	7
185.2 (100)	5	62	31	25	19
138.9 (75)	5	47	24	19	15
92.6 (50)	5	31	16	13	10
46.3 (25)	5	16	8	7	5

Effect of Data Compression on Recording Capability

The data in table V indicate that the bit rate reduction achieved through spectral or spatial compression or both greatly simplifies the data-handling problem. With 10:3 data compression, several configurations of 10-meter multispectral scanners can be recorded onboard using existing tape-recording systems; e.g., with a data rate less than 166.32 megabits/sec. With a 10:3 data compression, all configurations of 15-meter scanners are within current recording capability. Table V indicates that data from all configurations of the 30-meter scanners may be recorded on state-of-the-art recorders without data compression.

If the tape-recording capability can be advanced to 400 megabits/sec by 1980-90, it will be possible to record 10-meter IFOV, 10-channel, 185.2-kilometer (100 nautical mile) swath-width scanner data onboard the spacecraft with 10:3 compression.

GROUND DATA HANDLING

System for Compressed Data

The ground data-handling system for a multispectral scanner must be capable of reading the PCM-encoded flight- or ground-recorded tapes and of projecting the compressed data back to their original basis. The transformed data must be recorded either on digital computer-compatible tape or on some high-density digital tape for further ground data processing.

Because of the excessive volume of data from all advanced multispectral scanner systems, new data-processing techniques and systems must be considered for nationwide and worldwide surveys. The existing multispectral scanner data-analysis techniques and systems, such as those using the monitor program developed at the Laboratory for Applications of Remote-Sensing (LARSYSAA) (ref. 4) and the elliptical table/assign classification program (ELLTAB) (ref. 5), will still be useful for many statewide and regional surveys.

The computer times (central processor unit (CPU)) required to process scenes from 10-, 15-, 20-, and 30-meter scanners with 185.2-kilometer (100 nautical mile) flight segments and with swath widths of 46.3, 92.6, 138.9, and 185.2 kilometers (25, 50, 75, and 100 nautical miles) are given in table VI. For purposes of computation, it was assumed that the scanner had 10 channels. The times required to process the data through the LARSYSAA spectral-pattern recognition program, the digital table look-up classification program, the ELLTAB program, Univac 1108 and the Environmental Research Institute of Michigan (ERIM) multivariate interactive digital analysis system (MIDAS) (refs. 6 and 7) are given in table VI. The MIDAS is a very high speed multispectral analysis system, which is currently being developed by the ERIM. All the times listed in table VI are for classification only; it is assumed that additional time is used for tape reformatting, training-sample selection, spectral-signature development, output-data formatting, map-base conversion, film recording, tabulation of areas/class of materials, etc. Most, if not all, of the functions are required to produce usable multispectral scanner data products.

The MIDAS is a hardwired parallel processor that is now capable of processing m-bands × n-classes where $m \times n = 64$. The current prototype MIDAS can be extended to process more data bands and classes of materials. Processing times for an expanded system are unavailable at this time.

The times given in table VI for the LARSYSAA, digital table look-up, and ELLTAB programs are for identification of 24 classes of material, using as many as 10 input data bands. The time given for the MIDAS was for 16 classes of material and for 4 data bands.

TABLE VI.- COMPUTATION TIMES FOR SPECTRAL-PATTERN-RECOGNITION CLASSIFICATION FOR SELECTED SCANNER CONFIGURATIONS

IFOV, m	Swath width, km (n. mi.)	Area covered, km^2 (n. mi^2)	Univac 1108 CPU time, hr:min			System time, min:sec
			LARSYSAA	Digital table look-up	ELLTAB	MIDAS ERIM
10	185.2 (100)	34 299 (10 000)	1002:27	121:29	56:35	28:36
10	138.9 (75)	25 724 (7 500)	751:50	91:07	42:47	21:30
10	92.6 (50)	17 149 (5 000)	501:14	60:45	28:18	14:18
10	46.3 (25)	8 575 (2 500)	250:37	30:23	14:09	7:12
15	185.2 (100)	34 299 (10 000)	445:32	52:48	24:57	12:48
15	138.9 (75)	25 724 (7 500)	334:09	39:36	18:43	9:36
15	92.6 (50)	17 149 (5 000)	222:46	26:24	12:29	6:24
15	46.3 (25)	8 575 (2 500)	111:23	13:12	5:15	3:12
20	185.2 (100)	34 299 (10 000)	250:38	29:42	12:24	7:12
20	138.9 (75)	25 724 (7 500)	187:59	22:16	10:48	5:24
20	92.6 (50)	17 149 (5 000)	125:19	14:51	7:12	3:36
20	46.3 (25)	8 575 (2 500)	62:40	7:26	3:42	1:48
30	185.2 (100)	34 299 (10 000)	111:23	13:12	6:14	3:12
30	138.9 (75)	25 724 (7 500)	83:32	9:54	4:41	2:24
30	92.6 (50)	17 149 (5 000)	55:42	6:36	3:07	1:36
30	46.3 (25)	8 575 (2 500)	27:51	3:18	1:34	:48

The ERIM estimates that the speed of the MIDAS can be improved by a factor of 5 by increasing the efficiency of input/output devices. The MIDAS is an expensive, highly specialized system, but its processing speed may justify its cost.

System for Analysis of Compressed Data

Multispectral scanner data that have been compressed as described in the preceding section may be decompressed in a ground data-processing system and transformed to the original basis. Each original data band may be recovered with a known statistical error. The decompressed data may be analyzed by the techniques currently used.

It is also possible to analyze the compressed data without transforming them to the original basis. This analysis is possible because the supervised spectral-pattern-recognition analysis technique is a relative solution; e.g., if signatures are developed from compressed data, the unknown data points to be classified may be recognized in the compressed data format. If, however, compressed data are used for spectral-pattern-recognition analysis, a random sampling of the data should be decompressed to determine whether the quality of the original data bands is acceptable. The effect of including a noisy or otherwise defective data band in the compressed data set is not known.

RECOMMENDATIONS

Many areas requiring additional research and development have been identified. These areas are described in the following discussion and recommended in preparation for developing advanced multispectral scanner systems.

Quantitative Band Selection Development

A study should be initiated to develop a quantitative band selection procedure. More scanner data must be obtained, and the necessary theory must be developed. It is probable that such a development will take 3 to 5 years.

Tape Recorders

Tape-recorder technology should be advanced to accommodate approximately 400 to 500 megabits/sec during the 1980's. Encoding techniques should be developed to permit higher bit-packing densities while holding the bit error rate (1 bit error in 10^8) at current levels.

Data Compression

Scanner technology has been developed to a level exceeding current tape-recording capabilities. Because it is not necessary to continue recording the data without compression, tape-recorder-development requirements can be relaxed somewhat as the scanner data rates continue to increase. Some research has been done in data compression using data from visible-data bands. Additional work is necessary using a mixture of data bands throughout those portions of the electromagnetic spectrum that are used by multispectral scanners. The limits of compression must be established.

Swath-Width Selectability

Current scanner designs are limited to fixed swath widths. Areas of dynamic change are frequently much less than 185.2 kilometers (100 nautical miles) in width. Because advanced high-resolution multispectral scanners produce such great quantities of data, it may be desirable to have a choice of swath widths to limit data collection to only the required survey area. It is also conceivable that the very wide swath widths may be desirable because of operational considerations.

Data-Processing/Analysis Systems

Development of data-processing and data-analysis systems should continue. The main effort should be directed toward simplifying the procedure for converting the raw data to usable products. Advanced data-processing systems should have the capability of isolating and extracting specific areas from the original data set so that users may request desired portions of the data instead of receiving a 185.2-kilometer-square (100 nautical mile square) area.

CONCLUSIONS

It is likely that multispectral scanners with instantaneous fields of view as small as 10 meters and with 185.2-kilometer (100 nautical mile) or larger swath widths may be required during 1980-90. Such multispectral scanners produce a very high data bit rate that exceeds current tape-recording and data-transmission capabilities.

The state of the art in tape-recorder technology is not likely to advance to the level required for direct recording of data for the worst-case scanner configurations; but, with data compression before recording, it is likely that the data can be adequately recorded.

Considerable modification of the existing ground receiving stations is required (at considerable expense) before the noncompressed multispectral scanner data can be received. A few of the potential scanner configurations are compatible with the existing receiving stations (the space tracking and data network and the tracking and data relay satellite system), if the data are compressed before being transmitted from the spacecraft.

The current systems for analyzing large volumes of data are inadequate for advanced multispectral scanner systems. However, if detailed surface classifications are desired for subsets of the data, current approaches, such as the ELLTAB and LARSYSAA programs, are adequate. For large-scale surveys, such as nationwide or worldwide land use mapping and updating, systems such as the MIDAS will be required.

Lyndon B. Johnson Space Center
National Aeronautics and Space Administration
Houston, Texas, December 20, 1976
177-52-83-00-72

REFERENCES

1. Coggeshall, M. E.; and Hoffer, R. M.: Basic Forest Cover Mapping Using Digitized Remote Sensor Data and Automatic Data Processing Techniques. LARS Information Note 030573, Purdue Univ. (West Lafayette, Ind.), 1973.

2. Crawford, C. L.; Dye, R. H.; and Hanson, D. S.: Signature Data Processing Study. Final report, NASA contract 9-9848, vol. 1, Aerospace Systems Division of Bendix Corp., (Ann Arbor, Mich.), 1970.

3. Ready, P. J.; Wintz, P. A.; Whitsitt, S. J.; and Landgrebe, D. A.: Effects of Compression and Random Noise on Multispectral Data. Proceedings of the Seventh International Symposium on Remote Sensing of Environment (Environmental Research Institute of Mich., Ann Arbor, Mich.), vol. II, 1971, pp. 1321-1356.

4. Landgrebe, D. A.: LARSYSAA, Processing System for Airborne Earth Resources Data. LARS Information Note 091968, Purdue Univ. (West Lafayette, Ind.), 1968.

5. Jones, C.: Implementation of an Advanced Table Look-up Classifier for Large Area Land Use Classification. Proceedings of the Ninth International Symposium on Remote Sensing of Environment (Environmental Research Institute of Mich., Ann Arbor, Mich.), vol. II, 1974, pp. 813-824.

6. Kriegler, F.; Marshall, R.; et al.: Multivariate Interactive Digital Analysis Station (MIDAS): A New Fast Multispectral Recognition System. Proceedings of the Conference on Machine Processing of Remotely Sensed Data (Purdue Univ., West Lafayette, Ind.), 1973, pp. 46-51 to 46-64.

7. Kriegler, F.; Gordon, M.; et al.: The MIDAS Processor. Proceedings of the Tenth International Symposium on Remote Sensing of Environment (Environmental Research Institute of Mich., Ann Arbor, Mich.), vol. II, 1975, pp. 757-767.

NASA-Langley, 1977 S-472

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

SPECIAL FOURTH-CLASS RATE
BOOK

POSTAGE AND FEES PAID
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION
451

U.S.MAIL

322 001 C1 U D 770429 S00903DS
DEPT OF THE AIR FORCE
AF WEAPONS LABORATORY
ATTN: TECHNICAL LIBRARY (SUL)
KIRTLAND AFB NM 87117

POSTMASTER: If Undeliverable (Section 158 Postal Manual) Do Not Return

> *"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."*
>
> —NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons. Also includes conference proceedings with either limited or unlimited distribution.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include final reports of major projects, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C. 20546