DATA COMPRESSION FOR EARTH RESOURCES SATELLITES

C. L. MAY and D. J. SPENCER
TRW Systems
Redondo Beach, California

Summary This paper concerns an information-preserving data technique which is applicable to multi-spectral imagery such as that obtained by earth-resources satellites. The requirements for data compression in such missions is discussed and a rationale is presented for the use of distortion-free information-preserving compression.

The selected compression technique involves the use of the spectralspatial-delta-interleave (SSDI) algorithm, a form of DPCM, to eliminate gross spectral and spatial redundancies. This reduced data is then coded for transmission using either the Huffman or the Rice coding algorithms.

The coding algorithms have been simulated using a portion of frame 3698 taken during the Apollo S065 experiment and,the results are presented. A parametric study presents the compression achieved by the SSDI-Rice algorithm as a function of block size a nd split-pixel mode used. Implementation considerations are also given.

I. The Need for Data Compression in Earth Observation Experiments

The multispectral imaging sensors to be flown on the Earth Resources Technology Satellite (ERTS) will generate approximately 25 billion bits of data daily. In future earth observation missions this figure will likely multiply by several orders of magnitude. Such volumes of data and the implied data rates impose severe problems in onboard data storage, in communication links, in ground data processing and in ground data archiving. One technique which holds promise for easing this data problem is data compression.

The use of data compression in telemetry has been the subject of much research and a large number of papers. In spite of the progress made in this field, few missions have been able to justify the expense of onboard data compression with its attendant risks in reliability, possible alteration of data, and its cost in terms of size, weight, and power. Another factor limiting acceptability has been the fact that the scientific investigators usually design their own experiments, data formats, and ground data processing schemes.

The situation has changed recently. Digital logic has become smaller, more reliable, less expensive, and requires much less power for a given amount of processing. In addition, earth observation experiments are characterized by general purpose sensors which will be used by hundreds of investigators for many different purposes.[]] Such a variety of users demands that the designer achieve the greatest possible utility with his system.

The success of the earth resources survey programs will be measured to a large extent by the satisfaction it provides to the users of the data. That satisfaction in turn will depend on how well the data, as formatted and processed by the system, helps the individual accomplish tasks that are significant to him. In addition, any processing on the spacecraft should be accomplished without sacrificing the information fidelity requirement of the user. Imaging data, as produced by optical sensors, presents a particular attraction for reduction in the amount of data which must be handled at all stages in a data-gathering-to-ultimate-user chain. For example, high resolution multispectral scanners may produce data at 100 Mbps or higher rates. The data rates from such sophisticated sensors are set by the following factors:

1) The geographical coverage
2) The instantaneous field of view
3) Ground resolution element size
4) The spectral bandwidth

Table I indicates the results of a survey performed by TRW Systems of fugure requirements in many disciplines using earth observations data. The actual feasibility of including an experiment may thus be threatened by its large data rate and the accompanying data handling and communication problems that arise.

The following features are sought in any data compression scheme for Earth Resources imagery data:

- Significant compression
- Error within tolerances established by the application
- Inherently simple implementation capable of operation at high data rates
- Near-real time operation
- Non-overloading (increase in data rate due to unexpected data)

Many images may consist of large areas of relatively uniform texture interrupted by edges and high detail areas. Intuitively, it appears that a higher information rate should be used for signals from the second area than for the first. A source coding scheme should thus also be capable of recognizing pertinent statistical properties of the signal and adjusting its operation to them.

II. Redundancies Inherent in Multispectral Data

Two major types of redundancy are inherent in the multispectral image data. The first type is spatial redundancy due to properties of the scene being scanned and the scanning mechanism itself. The second type of redundancy is that existing between the spectral components of a scene. To illustrate these redundancies, assume that a plains area is being scanned. A relatively high spatial redundancy occurs and the ratio of spectral band intensities per picture element (pixel) remains relatively constant. Suppose that a cloud shadow is cast over a segment of the plains. When the shadowed region is encountered by the scanner, intensities in all spectral bands should simultaneously decrease and remain at some lower value until the show boundary is crossed. If the scanning system is such that continguous pixels represent the intensities of overlapping ground elements, then even higher pixel correlation exists in the spatial sense.

III. Redundancies Inherent in Multispectral Data

A strict information preserving data compression operation is one that, when applied to the output of a data source, reduces the amount of energy required to transmit characteristics of the source, without reducing the fidelity of the source output. In other words, with these techniques the reconstructed digital information is identical to the digital information entering the compressor. Generally, in SIP techniques use is made of the statistical relationship between nearby samples.

Two sequential operations are necessary for the strictly-information-pre serving coding of data. The first step involves a transformation algorithm, S, operating on the input data to decrease the variance of the level distribution, with sensor noise setting a lower bound on the variance achievable. Finally, an adaptive coding of this transformed data is performed by taking advantage of the transformed data statistics. Huffman or Shannon-Fano coding could be performed by measuring the varying statistics, but implementation considerations serve to prevent the use of such algorithms on board a satellite. The quasi-optimal Rice algorithm [2] is a practical alternative for coding the data.

For strictly-information-preserving systems, the inverse transformation, S^{-1} exists and is unique so that the original data can be derived by the ground processor. The performance of such a system can be measured in terms of the average number of bits per pixel (bippel) produced at the output of the compressor.

IV. The Proposed System for Consideration

The data compression technique which we have simulated consists of two parts, the SSDI algorithm and the Rice algorithm. The SSDI Algorithm operates on the raw digitized data pixel by pixel to reduce gross spectral and spatial redundancy with a DPCM-like transformation. The Rice Algorithm codes this data in a near-optimal fashion for transmission to the ground. Both of these techniques are inherently strictly information preserving (no distortion) but can be readily extended to permit a small amount of distortion to gain a higher degree of compression.

An algorithm developed at TRW to remove a maximum amount of redundancy subject to the constraints of minimal complexity and maximal operating speed is called the Spectral-Spatial-Delta-Interleave (SSDI). This method of compression first operates on the spatial redundancy existing in each spectral band and then uses this information to reduce spectral redundancies between adjacent bands. In the following discussion, 3 spectral bands are used and it is assumed that each intensity, 1, is quantized to 8 bits.

The SSDI algorithm is based on the transmission of the differential information between adjacent pixels in the scan direction. The basic algorithm is pictured in Figure 1. First, within each spectral band (α, β, γ are assumed to be the bands in this example) each pixel intensity is subtracted from the intensity of the pixel preceding it in the preferred scan direction. This technique is essentially DPCM, treating each spectral band by itself. Call a triple of these differentials ($\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$).

Next, these deltas are subtracted to yield the differences between the deltas in adjacent spectral bands. Call these second differences $d_A = \Delta\beta\text{-}\Delta\alpha$ and $d_B = \Delta\gamma\text{-}\Delta\beta$. If the spectral bands are truly correlated it should be true that on the average $|d_A| + |d_B| < |\Delta\beta| + |\Delta\gamma|$. Thus, the d differentials are clustered closer to the origin than the differentials. In fact, such is the case for the test multispectral data on which the SSDI algorithm was used.

These deltas are transmitted in such a fashion that the original PCM sensor data can be recovered exactly from the coded sequence. For each triple corresponding to a particular spatial pixel, the following sequence is transmitted, (Δ_α, d_A, d_B). Given the preceding pixel intensities ($l_\alpha^{i\text{-}l}$, $l_\beta^{i\text{-}l}$, $l_\gamma^{i\text{-}l}$), the current pixel intensities can be uniquely decoded as $i_\alpha^{\,i} = l_\alpha^{i\text{-}l} + \Delta_\alpha$, $l_\beta^{\,i} = l_\beta^{i\text{-}l} + \Delta_\alpha + d_A$, $l_\gamma^{i\text{-}l} + \Delta_\alpha + d_A + d_B$.

As a final observation, the SSDI (and modifications) is based on physical characteristics of multispectral data. As can be inferred from Figure 2 the triples (Δ_α, Δ_β, Δ_α) which occur per pixel tend to be clustered in an ellipsoid whose major axis is in the vector direction ı l, l, l ı. Thus, on the average, the differences d_A and d_B tend to be small. Neglecting additive noise, these triples become more concentrated about this vector as the spectral correlation of the data increases.

The algorithm described is the SSDI in its most elemental form. Many modifications of the SSDI have been used such as basing each A value not only on the current set of pixels but on a weighted set of pixels previously transmitted in order to realize a reduction in sensor noise and to diminish the Δ magnitude arising due to a large step in spatial intensity, as along a boundary of an object. This procedure, termed the SSDIA, bases each differential level, Δ, on the difference between the current sample in a band and the average of the amplitudes of the previous sample in that scan line and the three samples in the previous scan line which are contiguous to the present sample. Note that the four samples which are averaged have been previously transmitted and are available at the ground for unambiguous reconstruction. (A similar algorithm is described in [3]).

Following application of the SSDI algorithm, the differentials are further encoded by taking advantage of their statistical occurrence. Given the probability of the occurrence of each differential level, a Huffman code can be assigned to each level. Note that this code is based on the statistics of the entire scan. At the cost of greater complexity, the Huffman coding could be performed adoptively with the code changing as the scene statistics change.

An alternative to the use of Huffman encoding of SSDI data involves use of the Rice machine, developed by R. F. Rice at JPL[2]. This information preserving algorithm uses several concatenated codes to adapt the system to the time-varying data activity of the source. The resultant coding system produces output rates within .3 bits/pixel of the one-dimensional difference entropy for entropy values ranging from 0 to 8 bippels.

The appeal of the Rice Machine lies in its adaptivity to the source activity by provision for a multiplicity of coding schemes. Rice has developed efficient algorithms for selecting the proper coding scheme. The decoder must have knowledge of the coding scheme used at all times to permit proper data reconstruction. Any error which occurs in the coded bit stream or in the decoder decision as to mode used propagate to the end of a scan line.

V. Implementation Considerations

An important property of the SSDI algorithms is the low complexity of implementation associated with them. In the basic algorithm, only subtraction operations are associated with encoding and only additions are required for decoding, permitting a high speed system. High speed and low complexity are very desirable for the propose usage.

A study was performed by T. A. Zimmerman of TRW Systems[4] concerning the implementation of the data handling subsystem for a solid state multispectral imaging sensor. This processor performs the SSDI and the Rice algorithms on digitized input data. A block diagram of the system is given in Figure 3. The implementation assumed a resolution of 100 feet, five Megasamples per second, and 8-bits per sample. Processing at this bit rate of 40 Mbs requires a total of 233 flat packs, occupying a volume of 40 cubic inches and dissipating 43 watts.

Important considerations influence the implementation of the other algorithms described in this paper. For example, the preceding estimate does not include buffering which can be significant for data rate fluctuations. The block length statistics described below are a first attempt to judge this requirement. Also, averaging algorithms such as the SSDIA and onedimensional averaging DPCM require storage of the previous line of data (when working with scanned data). This may require an impractically large storage in some cases. Additional considerations such as random noise cause error propagation from line to line with these algorithms as discussed in [3].

VI. Simulation of the SSDI-Rice Algorithm on Satellite Imagery

The achievable compression using the SSDI technique is entirely dependent upon the data source. To validate its usefulness, requires testing on a variety of imagery.

Ideally, such images should be perfectly registered low noise multispectral images taken from a satellite and available in a readily usable digitized format. We were fortunate in obtaining from the Laboratory for Applications of Remote Sensing (LARS) at Purdue University, digital tapes containing two such scenes. These were taken on the Apollo 9 mission during the S065 experiment, the first attempt to obtain multispectral photography from space. The S065 experiment obtained pictures of the earth from space using four different film/filter combinations, simultaneously, by means of a four-camerq assembly mounted in the spacecraft hatch window. Shutters of the cameras were electrically operated in synchronism producing approximately registered photographs. Intensity sampling, digitization and digital registration between three spectral bands (approximately red, green and blue) was performed by LARS.

Frame number 3698 of the Salton Sea, California area was used for the simulation. This picture was taken at an altitude of 105 n.m., sun elevation angle 53°, with a 5 percent cloud cover. Each image has been scanned into approximately 2000 x 2000 pixels, with 8,bits per pixel, but only a smaller 100 x 100 pixel subscene has been used to date. This particular region included both agricultural and desert areas as shown in Figure 4. More extensive simulation of both images is planned.

Several statistical measurements were taken using the chosen data. The probability density function of the intensity values is measured by simply counting the probability of occurrence of (quantized) intensity values over the scene.

The pdf of first order intensity differences is obtained by first subtracting the intensities of two pixels, separated by k-1 intervening pixels, at a time over the entire scene. For each pixel pair, the difference is computed and the occurrence of each difference value is counted over the scene, divided by the number of pixel pairs used, and displayed graphically. The pdf measures the inter-pixel spatial correlation and the smaller the variance of this distribution about zero, the higher the correlation. Figure 5 illustrates this distribution.

The spectral correlation of the data is obtained by measuring the joint pdf of spatial differences Vectors over the scene. Each pixel intensity is subtracted from the intensity of the previous pixel in each of the M spectral bands and arrayed as an M vector of differentials for each pixel pair.

These differential vectors are grouped, counted, converted to a percentage, and graphed as a multi-dimensional joint pdf. If the data were noiseless and completely correlated spectrally, each vector of M differentials would lie along the vector $\lfloor \underline{l, l, ..., l} \rfloor$ in probability space. Therefore, the higher the spectral correlation, the closer the clustering of probabilities about this vector. Figure 2 gives a two-dimensional view of this clustering effect for our data and illustrates a moderately high spectral correlation.

The cross spectral-spatial correlation is obtained by forming the normalized dot products of the intensity vectors, l_i and l_{i+k}, corresponding to pairs of pixels spatially separated by k-1 intervening pixels. Normalization removes the effects of scene illumination and the dot product gives the cosine of the angle between these vector intensities. The closer this normalized dot product is to unity, the high the correlation between the pairs. Figure 6 illustrates the variation of the correlation between pixel pairs as a function of their spacing, k, for the given scene..

Simulation of the SSDIA algorithm using four contagious pixels for averaging, with Huffman coding gives an average code length of 4.244 bippels, a compression of about 47 percent. A Huffman code is given in Table II for the SSDI distribution generated from the chosen S065 scene.

Simulation results of the SSDI/SSDIA - Rice algorithm on the same scene are presented in Table III with the best results obtained using a one-dimensional averaged DPC8-Rjce algorithm.[*] These various cases are compared parametrically for several Rice split-pixel modes. A spatial "blur," as defined in the footnote of Table III, was applied to the raw sensor data in an attempt to reduce mis-registration effects. Performance is increased by

---

[*] The one-dimensional averaging DPCM is performed for each element in each spectral band separately by the equation

$$\Delta_{i,j} = x(i,j) - 1/4[x(i,j-1) + x\ (i-1, j-1) + (i-1,j) + x(i-1,j-1)]$$

.5 to .8 bippels by use of the blur. To what extent this is due to averaging or to registration correction was not evaluated.

Many interesting comparisons may be made from Table Ill. The minimum bippels for the SSDIA (without blur) was obtained in the Rice (7, 1) mode. The results of Rice indicate that the first order source entropy should then be in the range of 4.5 bippels. This agrees with the entropy derived from the SSDIA distribution for the same data. Another interpretation of this result is that there is about I bippel rms. of sensor noise which, if ignored, requires only 3.75 bippels in the (7, 0) mode. This seems to be borne out by the investigations of Ready [5] (operating on the same data) which have shown an increase in classification accuracy with compression in the same range.

The improvement of nearly 1/2 bippel over the SSDIA with the simple one-dimensional averaging DPCM is explainable only by lack of correlation between the spectral bands. This is not unexpected for such widely differing spectral bands in pictures which have undergone many processing steps introducing geometric distortions. Comparisons on other data indicate the improvement achievable with the SSDI on perfectly registered data with closer band spacing.

The results of Table III represent the <u>average</u> data rates. Distributions of block sizes were measured for each case and runs were also made parametrically with block length. As seen in Figure 7, there is little change in compression with block length except that due to the fixed overhead bits required by the Rice algorithm. A block length of 8 or 12 seems best when balancing average code length and the desire for the greater noise immunity obtainable with shorter blocks. Similar results were found for the SSDIRice and SSDIA-Rice, leading to the choice of 12 for block length in Table III.

References

1. D. J. Spencer, V. K. Agarwal, R. K. Nisbett, P. S. Romo, “Data Handling for Earth Observatory Satellites,” Proc. of IEEE National Telemetering Conference, Washington, D. C.; April 14, 1971.

2. R. F. Rice, “TV Data Compression for a Grand Tour of the Outer Planets,” Proc. of Fourth Hawaii International Conference on System Sciences, pp. 595-597; January, 1971.

3. D. J. Connor, R. F. W. Pease, W. G. Scholes, “Television Coding Using Two-Dimensional Spatial Prediction,” BSTJ Vol. 50, No. 3, Pp. 10491061, March, 1971.

4. T. A. Zimmerman, "Multispectral Data Compression Using the Spectral Spatial Delta Interleave Algorithm and Rice Algorithm," TRW IOC 733.3-179, September 17, 1971.

5. P. J. Ready, P. A. Wintz, S. J. Whittsitt, and D. A. Landgrebe, "Effects of Compression and Random Noise on Multispectral Data," Proc. 7th Int. Symp. Remote Sensing of Environment, Univ. of Michigan, Ann Arbor; May, 1971.

Acknowledgement

Special appreciation is owed to the Laboratory for Applications of Remote Sensing (LARS) of Purdue University for Supplying the digital tapes of the Apollo 9 scenes used in this study. We also appreciate the help of T. A. Zimmerman and V. K. Agarwal of TRW on this project, and the helpful discussion of John E. Taber.

	SPECTRAL RESOLUTION	SPATIAL RESOLUTION	FIELD OF VIEW	OBSERVATIONAL FREQUENCY
Earth Resources	Med.	50-300 feet	200 n.mi - 400 n.mi.	90 days
Global Oceanography	Med.	.5 n.mi., 50 n.mi.	100 n.mi.-1500 n.mi.	1 day, 30 days
Coastal Oceanography	High	25-100 feet	1-100 n. mi.	1,3,7,10,14,30 days
Meterology	Low	5-20 n.mi.	1500 n.mi. - 5000 n.mi.	3-12 hours

MULTISPECTRAL IMAGING SENSOR EXAMPLE

Spatial Width	200 n. mi.
Spatial Resolution	100 feet
Spectral Resolution	12 bands, 0.4-1.2 µ
Data Rate Encoding	8 bits/sample = 248 Mbps

TABLE I. TYPICAL USER DATA REQUIREMENTS FOR EARTH OBSERVATIONS

DELTA	PROBABILITY	CODE
-8	.014	011111
-7	.021	001010
-6	.031	01110
-5	.042	00100
-4	.054	1001
-3	.071	0101
-2	.087	0000
-1	.097	111
0	.103	101
1	.098	110
2	.084	0001
3	.071	0100
4	.055	1000
5	.041	00101
6	.032	00111
7	.020	001011
8	.015	011110
S	.064	0110XXXX

Table II
A Huffman Code for the SSDIA for Frame 3698

Differential amplitudes from -8 to +8 are coded using from three to six bits. Differential amplitudes in the ranges [-16, -9] and [9, 16] are coded with four bits preceded by prefix 0110. The average code length is 4.244 bipels, implying a data reduction of 47 percent.

TABLE III. Results of the SSDI-Rice Applied to a Subscene of Frame 3698

RICE MODE	SSDI		SSDIA		ONE DIM AVE	
	BLUR	NO BLUR	BLUR	NO BLUR	BLUR	NO BLUR
(8,0)	4.56	5.52	4.10	4.88	3.62	4.23
(7,1)	4.28	4.81	4.37	4.75	4.53	4.80
(6,2)	4.57	4.92	4.91	5.25	5.19	5.52

Notes:
(1) RICE algorithm block size in all cases was 12 pixels in each of 3 bands or 36 samples in all.
(2) (7,1) and (6,2) mode results were obtained by using (7,0) and (6,0) modes, then adding 1 or 2 bits, respectively, to the results.
(3) Blur equation for every pixel in each spectral band:

$$X(i, j) = .68\, x(i, j) + .28[x(i, j-1) + x(i, j+1) + x(i+1, j) + x(i-1, j)] + .04[x(i-1, j-1) + x(i-1, j+1) + x(i+1, j-1) + x(i+1, j+1)]$$

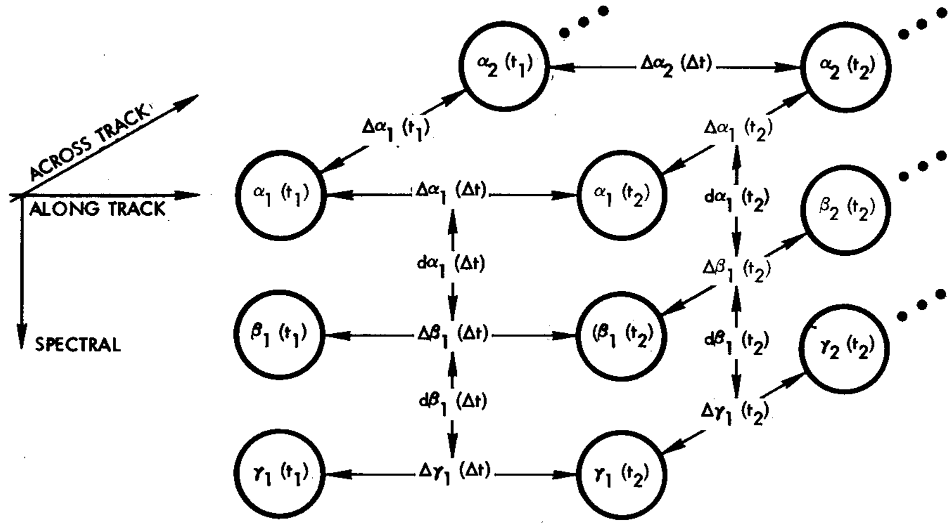

FIGURE 1. DEFINITION OF FIRST AND SECOND ORDER DIFFERENCES IN SSDI

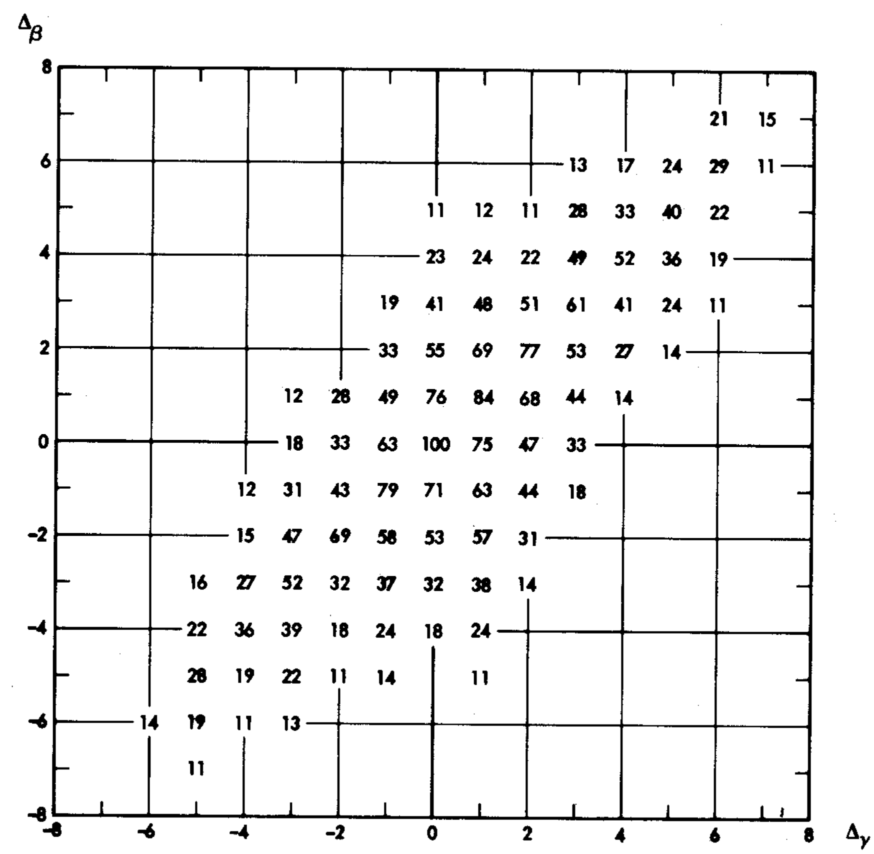

FIGURE 2. Projection of the Joint Probability Density, $P(\Delta_\alpha, \Delta_\beta, \Delta_\gamma)$, on the plane.**

** Defined by Δ_β and Δ_γ. To Illustrate Ellipsoidal Clustering About Vector $\lfloor 1, 1, 1 \rfloor$.
Percent probability is shown multiplied by ten. Only points having at least 0.5 percent probability are given

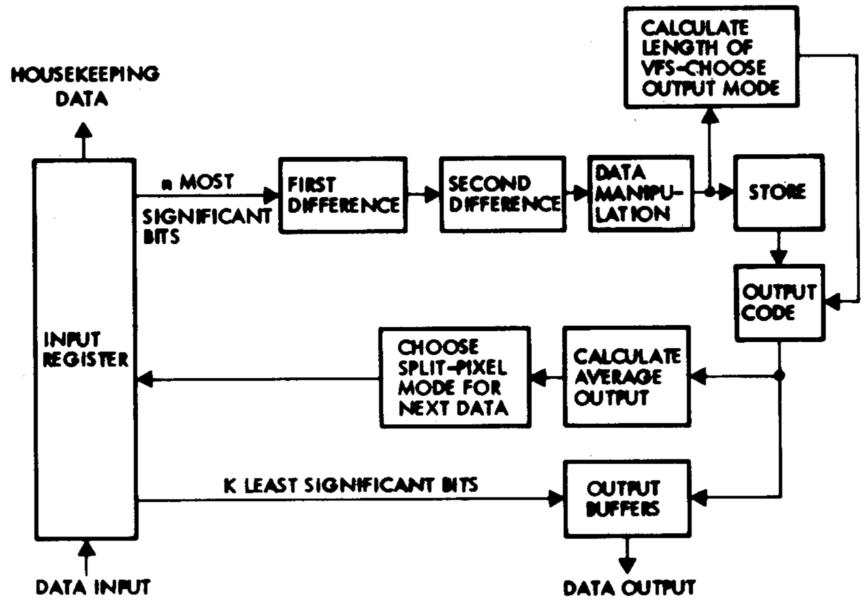

FIGURE 3. OVERALL BLOCK DIAGRAM OF THE SSDI - RICE ALGORITHM

FIGURE 4. FRAME 3698 OF THE IMPERIAL VALLEY FROM THE APOLLO 9 50-65 EXPERIMENT

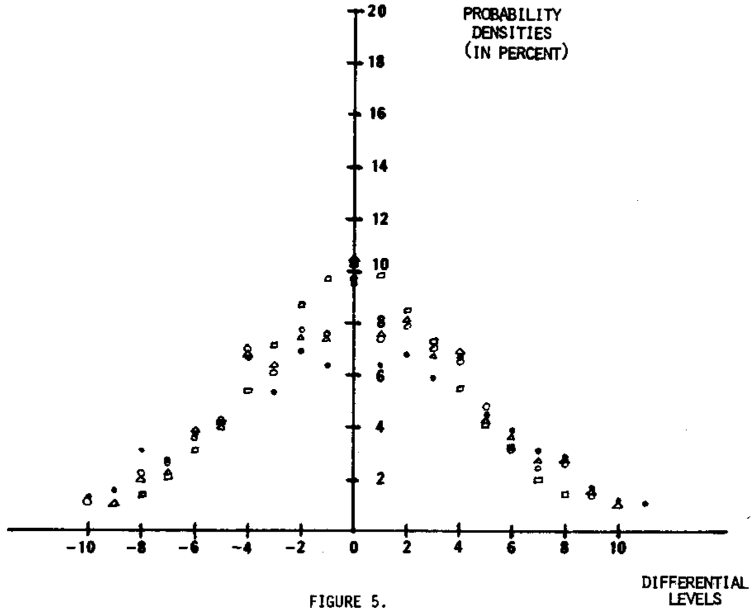

FIGURE 5.

The Probability of Occurrence of Differential Levels in the Chosen Scene

Legend: First-Order Spacial Differences (DPCM)

○ - band 1

● - band 2

△ - band 3

□ - Spectral-Spatial Delta Interleave (SSDIA)

Note: That only the differential levels occuring more than one percent of the time are given on the plot.

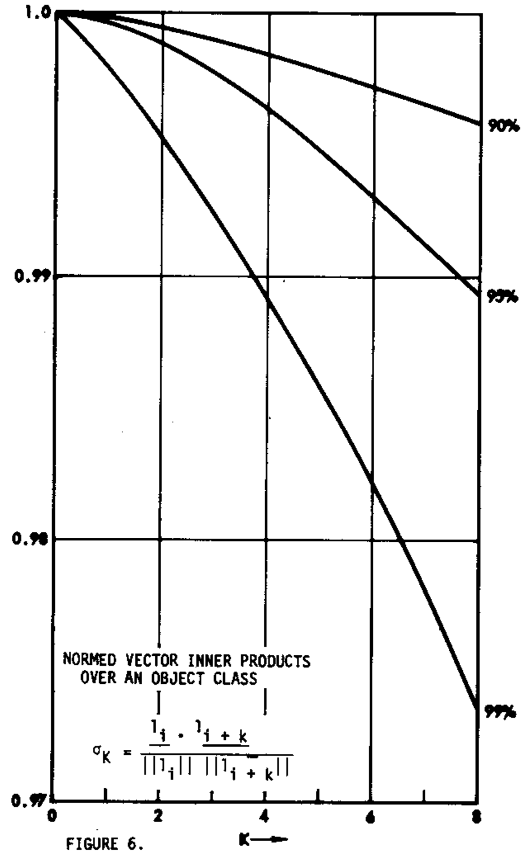

FIGURE 6.

Graphs of the Normalized Inter-Pixel Vector Intensity DOT Product, σ_K, Over an Object Class. The Curves Correspond to the Percentages of Pixels a Distance K Apart Having Normed DOT Products Greater Than σ_K.

FIGURE 7.

BIT RATE AS A FUNCTION OF BLOCK LENGTH FOR THE ONE-DIMENSIONAL AVERAGING DPCM