1-1-1972

Multispectral Data Compression through Transform Coding and Block Quantization

P. J. Ready

P. A. Wintz

Information Note 050572

Multispectral Data Compression Through Transform Coding and Block Quantization

P. J. Ready

P. A. Wintz

The Laboratory for Applications of Remote Sensing

Purdue University

LARS Information Note 050572

MULTISPECTRAL DATA COMPRESSION THROUGH TRANSFORM CODING AND BLOCK QUANTIZATION[1]

Patrick J. Ready
Paul A. Wintz
TR-EE 72-2
May, 1972

Published by the
Laboratory for Applications of Remote Sensing (LARS)
and the
School of Electrical Engineering
Purdue University
Lafayette, Indiana 47907

[1]This work was supported by the National Aeronautics and Space Administration under Grant No. NGL 15-005-112.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure

Figure Page

Figure

ABSTRACT

Transform coding and block quantization techniques are applied to multispectral data for data compression purposes. Two types of multispectral data are considered, (1) aircraft scanner data, and (2) digitized satellite imagery. The multispectral source is defined and an appropriate mathematical model proposed.

Two error criteria are used to evaluate the performance of the transform encoder. The first is the mean square error between the original and reconstructed data sets. The second is the performance of a computer implemented classification algorithm over the reconstructed data set. The total mean square error for the multispectral vector source is shown to be the sum of the sampling (truncation) and quantization error.

The Karhunen-Loéve, Fourier, and Hadamard encoders are considered and are compared to the rate distortion function for the equivalent gaussian source and to the performance of the single sample PCM encoder.

The K-dimensional linear transformation is shown to be representable by a single equivalent matrix multiplication

of the re-ordered source output tensor. Consequences of this result relative to the K-dimensional Fourier and Hadamard transformations are presented.

Minimization of the total encoder system error over the number of retained transform coefficients and corresponding bit distribution for a fixed data rate and block size is considered and an approximate solution proposed. Minimization of the sampling error over the data block size for the continuous source is also considered.

The results of the total encoder system error problem are applied to both an artifically generated Markov source and to the actual multispectral data sets.

The Karhunen-Loéve transformation is applied to the spectral dimension of the multispectral source and the resulting principal components are evaluated as feature vectors for use in data classification.

Experimental results using the transform encoder and several different (i.e., one, two, and three dimensional) data blocks are presented for both the satellite and aircraft data sets. Performances of the encoders over the three test regions within the satellite data are evaluated and compared.

CHAPTER I

INTRODUCTION

Remote sensing of the environment is rapidly becoming a major area of research for engineers and scientists throughout the world [1,2]. Satellites and high altitude aircraft provide ideal platforms from which earth resources data may be gathered. The data is gathered in the form of several spectral images of a particular area of the earth under observation. Each image represents the spatial distribution of electromagnetic energy as seen through a given spectral window. Information about a particular area is obtained through the study of the spatial and spectral characteristics of the data for that area. Temporal characteristics are also useful, but are not considered in this study. A source producing data in the above manner is defined to be a multispectral source [3,4,5].

1.1. The Need for Data Compression

Due to the extremely large volumes of data generated by a multispectral source three major problems require attention. The first is the potentially wide bandwidth required to transmit data from a remote sensor to a data collection center, as in satellite transmission to a

ground station. As the quantity of data transmitted in a given amount of time increases so does the required bandwidth.

The second problem is the actual physical storage of multispectral data. The value of gathered data never vanishes since one cannot always predict with certainty the future applications of the data. Data libraries soon become unreasonably large as the quantity of stored data increases.

The third problem is the increasingly large blocks of time required for man/machine analysis of multispectral data. The three dimensional nature of the data makes it quite difficult, especially for the human analyst, to efficiently use the large amounts of information provided.

The application of appropriate data compression techniques to the data can significantly reduce the severity of the above three problems. Data compression reduces the quantity of data to be transmitted in a given amount of time and thereby decreases the required transmission bandwidth. Secondly, multispectral data can be stored in the compressed state and reconstructed upon user request, resulting in more efficient data storage.

Thirdly, analysis can sometimes be performed on the compressed data, thereby reducing the amount of data the human or machine analyst must process.

Since it is not possible to anticipate the needs of future users of the data, it is important that any data compression technique be information preserving. That is, the technique should not destroy more than what has been determined to be the maximum acceptable information loss. In addition, the data compression technique should be capable of efficient compression of the several different kinds of data (i.e., vegetation, desert, mountains, etc.) that a multispectral sensor might encounter over changing terrain.

1.2. Transform Coding and Block Quantization

The data compression technique analyzed in this study is that of an appropriate orthogonal transformation of each realization of the multispectral source output, followed by quantization of some subset of the transform coefficients. The most frequently reported application of this technique is in the area of one dimensional speech processing and two dimensional image processing [6], and has been investigated by several authors, most notably Kramer and Mathews [7], Huang and Schultheiss [8], Pratt et al [9], and Habbi and Wintz [10]. A general description of past and present advances in image processing is given by Huang, et al in [11] and by Wintz in [59]. Extensive bibliographies on data compression and bandwidth reduction have been compiled by Wilkins and Wintz [12], Pratt [13], and Rosenfeld [14].

Applications of transform coding and block quantization to the three dimensional multispectral source have yet to be reported by any authors, although Haralick and Dinstein [15], and Crawford et al [16] used a one dimensional spectral transformation for data clustering purposes. A non-transform coding approach (i.e., data omission, low pass filtering with mixed highs, etc.) applied to multispectral data is reported in [17], and some work with two dimensional transform coding of multispectral data by Silverman is given in [18] and Haralick et al in [48].

1.3. The Two Data Sources

The two principal sources of multispectral data, satellite and aircraft mounted scanners, are studied separately. This is done in order to provide results more representative of each, in that enough significant statistical difference may exist in their outputs to warrant the separation. One major source of the difference is the vastly different altitudes from which the data is gathered. Aircraft data is typically gathered at an altitude of from one half to eight miles while satellite data is obtained at altutudes of approximately 100 miles and more. The satellite data thus (1) may not have the spatial resolution of the aircraft data, (2) may contain more variations in ground structure and cover within a given time and spatial period than the localized aircraft data.

For example, Figures 1.1 and 1.2 are images from the two data sets on which most of the experimental results of this study are based. The satellite data contains vegetated areas, desert, and mountainous regions. In contrast, the aircraft data contains a large percentage of vegetation, some bare soil, and a rather small percentage of roads and buildings. (The panchromatic photograph shown in Figure 1.1 was taken in late May. The aircraft scanner data used in this study is from the same physical area, but gathered approximately one month later.)

1.4. Statistical Description of the Multispectral Source

The assumption is made that both the aircraft and satellite sensors may be adequately modeled by the source shown in Figure 1.3. The source is time discrete and amplitude continuous. Each realization of the source output is the ensemble of real numbers $x_1, x_2, \ldots, x_N$. The output is thus an N-dimensional vector $\underline{X}$ in Euclidean N-space.

The elements of $\underline{X}$ may be the spectral intensity of an area on the ground as measured through N different wavelength bands (channels), or $\underline{X}$ may represent the spatial variation of reflected energy within a given wavelength band, or $\underline{X}$ may be some combination of the above two interpretations. This generality in the definition of the source output vector $\underline{X}$ is useful in the analysis which follows.

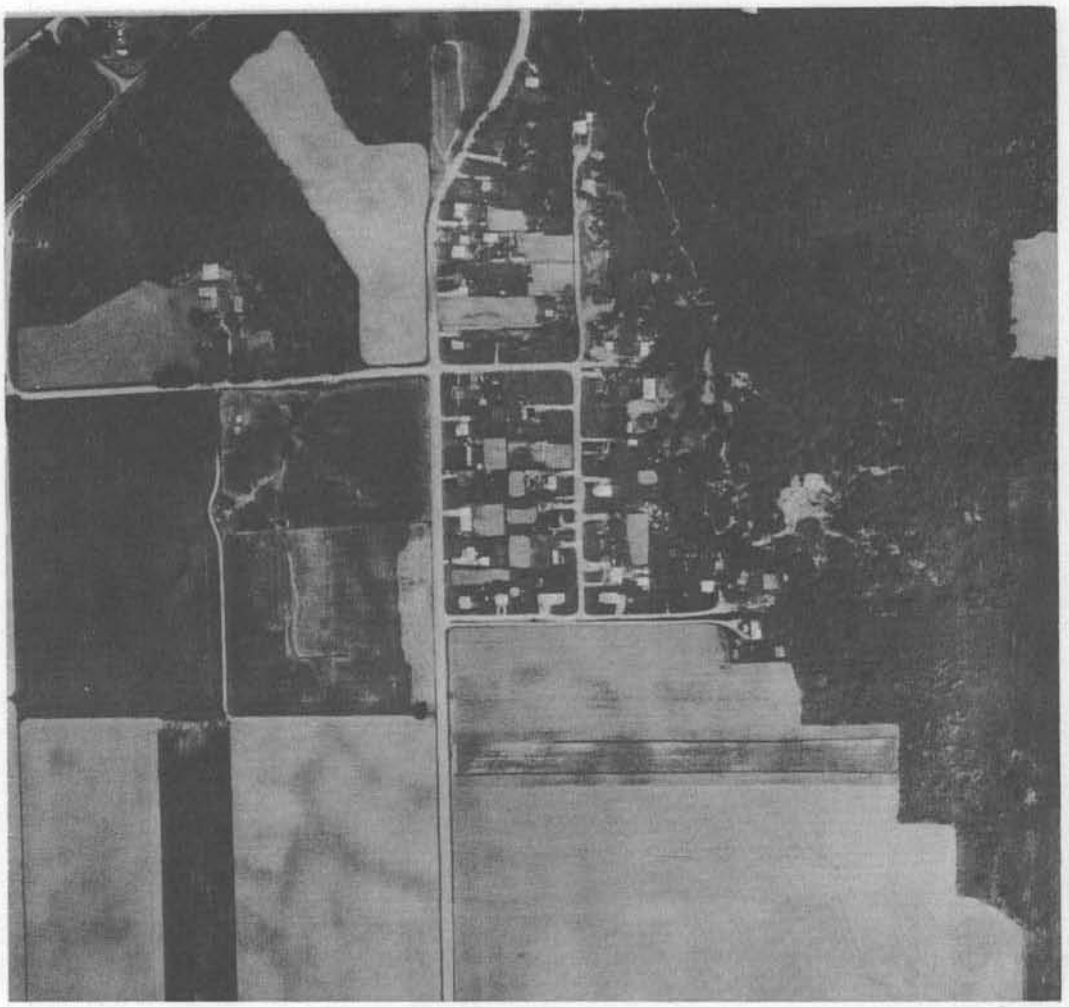


Figure 1.1. Panchromatic Photograph of the Aircraft Scanner Data Set.

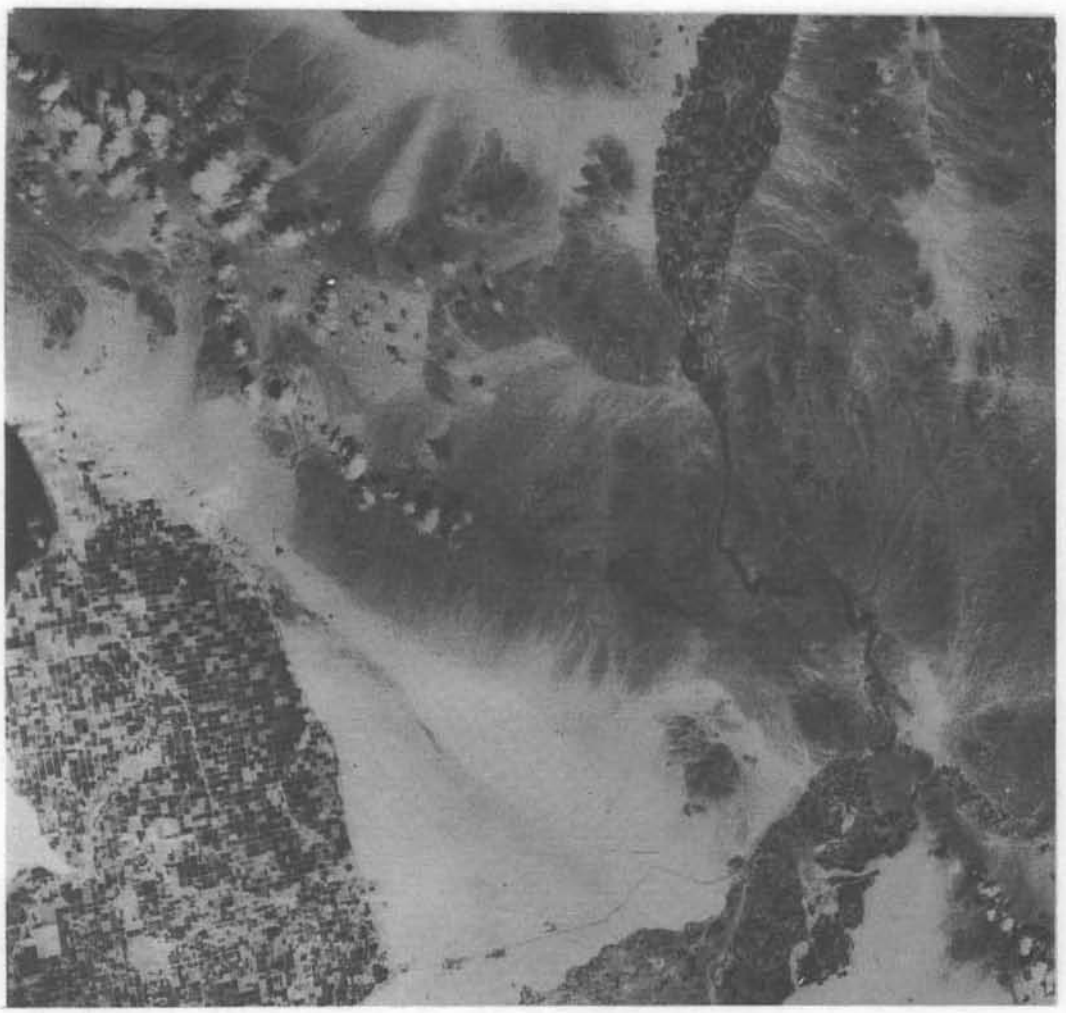


Figure 1.2. The Satellite Data Set (.59-.715μm).

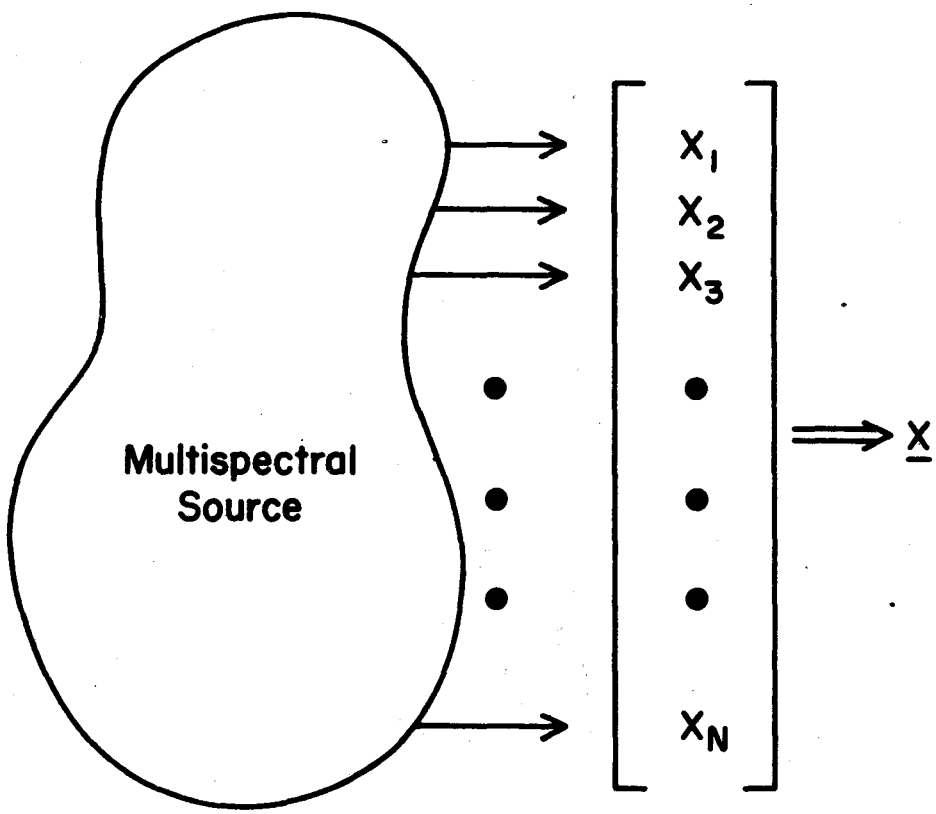


Figure 1.3. The Multispectral Vector Source.

In any case, $\underline{X}$ is assumed to be a random vector having mean $\underline{U}$

$$\underline{U} = E\{\underline{X}\} \tag{1.1}$$

and NxN real, symetric covariance matrix $\underset{\sim}{C}$

$$\underset{\sim}{C} = E\{(\underline{X}-\underline{U})(\underline{X}-\underline{U})^t\} \tag{1.2}$$

The variances of the $\{x_i\}_{i=1}^N$ are given by the diagonal elements of $\underset{\sim}{C}$. Thus

$$E\{(x_i-u_i)^2\} = \sigma_{x_i}^2, \quad i = 1, 2, \ldots, N \tag{1.3}$$

The average source output energy is then

$$E_s = E\{||\underline{X}-\underline{U}||^2\} \tag{1.4a}$$

$$= \sum_{i=1}^{N} \sigma_{x_i}^2 \tag{1.4b}$$

It is further assumed that $\underline{U} = \underline{0}$ for the multispectral source. This represents no restrictions on the source since $\underline{X}$ can be forced to have zero mean by subtracting $\underline{U}$.

No assumptions are made regarding the joint probability density function for the elements of $\underline{X}$, although it is reported in [19] that the jointly Gaussian assumption is a

reasonable approximation for several applications in pattern recognition.

The elements of $\underline{X}$ are, in general, correlated and successive realizations of $\underline{X}$ may also be correlated (i.e., the source has memory). It is this memory or source redundancy that is the motivation for transform coding of the source output $\underline{X}$.

1.5. Basic System Structure

The system diagram of the transform coding block quantization technique used in this study is presented in Figure 1.4.

The output of the transform encoder is the n-dimensional random vector ($n \leq N$) of real numbers $\underline{Y}$. This output is obtained by pre-multiplying the source output vector $\underline{X}$ by the nxN transform matrix $\underset{\sim}{T}$. The matrix transform $\underset{\sim}{T}$ is chosen based on its ability to pack a large percentage of the total output variance E_s into as few elements of $\underline{Y}$ as possible. This property of the transformation is discussed in Chapter 2.

Since (1) the time-bandwidth product of any physically realizable means of data transmission is finite and (2) the storage capacity of data libraries is limited, the n-vector $\underline{Y}$ must be quantized to a finite number of levels per vector element. The output of the bank of n quantizers $\underline{Y}^*$ is therefore a distorted version of the input vector $\underline{Y}$.

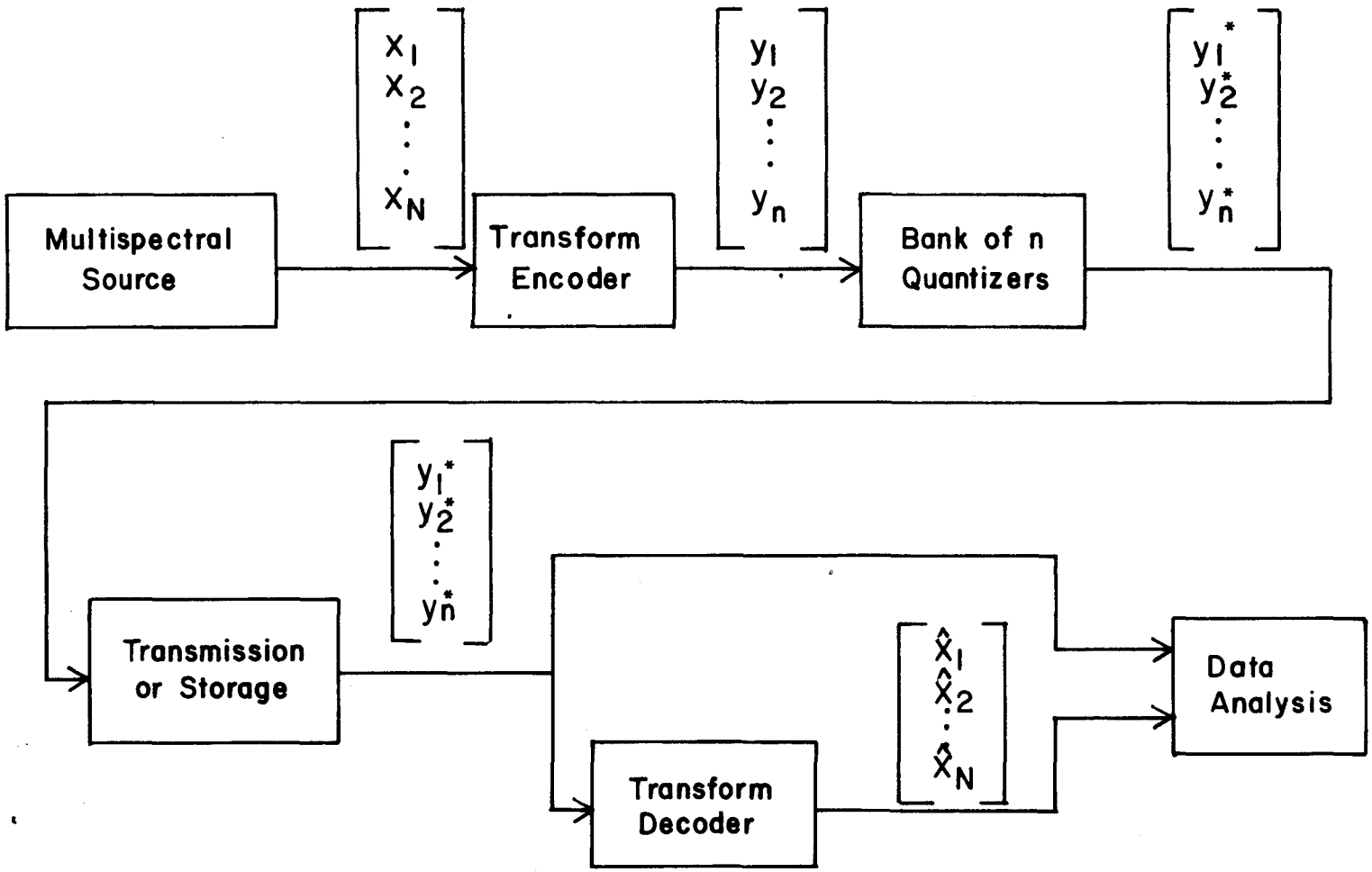


Figure 1.4. System Diagram of the Transform Coding-Block Quantization Data Compression Technique.

The vector $\underline{Y}^*$ is transmitted or stored, depending on the application of the system. In either case it is assumed that no error occurs in transmission or storage retrieval. Thus the input to the decoder is again $\underline{Y}^*$.

In some cases, to be discussed later, user analysis of the multispectral data may be carried out using $\underline{Y}^*$. In other cases the original source output $\underline{X}$ is of interest. The vector $\hat{\underline{X}}$ is an estimate of $\underline{X}$ obtained from $\underline{Y}^*$ by the transform decoder. This is accomplished by premultiplying $\underline{Y}^*$ by the transpose of $\underline{\underline{T}}$

$$\hat{\underline{X}} = \underline{\underline{T}}^t \underline{Y}^* \qquad 1.5$$

A useful system produces an estimate $\hat{\underline{X}}$ that is a reasonable approximation to the original source output $\underline{X}$ when measured by some error criteria meaningful to the various users of the data. At the same time the useful system must produce a significant reduction in transmission bandwidth or data storage requirements.

CHAPTER II

ERROR CRITERIA, SAMPLING AND QUANTIZATION

2.1. Definition of the Error Criteria

Two criteria are used in this study as a measure of the fidelity of reproduction of the original source output $\underline{X}$. The first is the mean square error between $\underline{X}$ and $\hat{\underline{X}}$ denoted by d. It is the expected value of the square of the Euclidean distance between $\underline{X}$ and $\hat{\underline{X}}$.

$$d = E\{||\underline{X}-\hat{\underline{X}}||^2\} \qquad 2.1$$

The percent mean square error is also used

$$d_p = (d/E_s)\cdot 100\% \qquad 2.2$$

where E_s has been previously defined in equation 1.4 as the average vector source output energy.

The second error criteria is percent classification accuracy. This measure reflects the ability of the data user (man or machine) to distinguish between pre-selected classes within the reconstructed data set. It is defined to be the number of data points correctly classified into a preselected number of categories, divided by the total number of data points considered (X100%). The classification results presented in this study are the result of

automatic point by point classification by computer [20]. It may be, however, that the results are also indicative of the performance of human "classifiers", such as photo-interpreters, etc. Thus an attempt is made to measure the usefulness of $\hat{\underline{X}}$ as compared to the original $\underline{X}$ in the analysis of multispectrally imaged areas. The ability of a data user to gain information about a particular area from each spectral band (image) transmitted or stored is critical. A particular spectral band having a relatively low distortion as measured by the mean square error criteria d, but distorted in such a way as to be useless for classification purposes is not an acceptable situation. It is therefore suggested in this study that a consideration of both error criteria, mean square error and classification accuracy, gives a reasonable indication of the effects of multispectral data compression on data quality.

2.2. The Rate Distortion Function

Using the mean square error criteria described in Section 2.1 an information rate may be associated with the multispectral source based on the concept of average mutual information [21, 22]. The average mutual information $I(\underline{X}, \hat{\underline{X}})$ between $\underline{X}$ and $\hat{\underline{X}}$ is defined to be the average information provided about the event

$$\text{Source output} = \underline{X} \qquad 2.3$$

by the occurrence of the event

$$\text{Transform decoder output} = \hat{\underline{X}} \tag{2.4}$$

In terms of the expectation operator, $I(\underline{X}, \hat{\underline{X}})$ is defined to be

$$I(\underline{X}, \hat{\underline{X}}) = E\{\log_2 \left(\frac{P(\underline{X}|\hat{\underline{X}})}{P(\underline{X})}\right)\} \text{ bits} \tag{2.5}$$

where $P(\underline{X})$ is the joint probability density function for $\underline{X}$, and $P(X|\hat{X})$ is the conditional probability density function for $\underline{X}$ conditioned on the event 2.4.

That $I(\underline{X}, \hat{\underline{X}})$ is a measure of the mutual information between $\underline{X}$ and $\hat{\underline{X}}$ is evident since (1) if $\underline{X}$ and $\hat{\underline{X}}$ are statistically independent (i.e., knowledge of $\hat{\underline{X}}$ gives no information about $\underline{X}$) then $P(\underline{X}|\hat{\underline{X}}) = P(\underline{X})$ and $I(\underline{X}, \hat{\underline{X}}) = 0$, (2) if the event $\hat{\underline{X}}$ implies $\underline{X}$ (i.e., knowledge of $\hat{\underline{X}}$ completely specifies $\underline{X}$) then $P(\underline{X}|\hat{\underline{X}}) = 1$, and

$$I(\underline{X}, \hat{\underline{X}}) = E\{\log_2 \left(\frac{1}{P(\underline{X})}\right)\} \quad \text{bits} \tag{2.6a}$$

$$= H(\underline{X}) \quad \text{bits} \tag{2.6b}$$

where $H(\underline{X})$ is the entropy or self information of the source.

The rate distortion function for the multispectral source $R(d)$ is defined by

$$R(d) = \inf_{P(\underline{X}|\hat{\underline{X}})\in S_d} \{I(\underline{X}, \hat{\underline{X}})\} \text{ bits} \qquad 2.7a$$

where S_d is set of all conditional probability density functions $P(\underline{X},\hat{\underline{X}})$ such that

$$d \geq \iint P(\underline{X},\hat{\underline{X}})||\underline{X}-\hat{\underline{X}}||^2 d\underline{X}d\hat{\underline{X}} \qquad 2.7b$$

as in 2.1. Since the infimum is over all conditional densities in S_d, the rate of the source is a function of the source statistics and not the transmission channel or the storage technique.

The rate distortion function as defined in 2.7 is derived for the Gaussian vector source with memory by Berger [22] Chapter 4, by Bunin in [23], and by Kolmogorov (for the time and amplitude continuous source) in [61]. R(d) is given parametrically by the equations

$$R(d) = \frac{1}{2N} \sum_{i=1}^{m(\varepsilon)} \log_2 \left(\frac{\lambda_i}{\varepsilon}\right) \text{ bits/vector element} \qquad 2.8a$$

$$d = \sum_{i=m(\varepsilon)+1}^{N} \lambda_i + \varepsilon m(\varepsilon) \qquad 2.8b$$

where the λ_i, $i=1, 2, \ldots, N$ are the ordered $(\lambda_1>\lambda_2>\cdots>\lambda_N)$ eigenvalues of the NxN source covariance matrix $\underset{\sim}{C}$ and

$m(\varepsilon)$ is an integer such that

$$\lambda_1 > \lambda_2 > \cdots > \lambda_{m(\varepsilon)} > \varepsilon \geq \lambda_{m(\varepsilon)+1} > \cdots > \lambda_N \qquad 2.8c$$

If $\varepsilon \geq \lambda_1$, then $R(d) = 0$.

The rate distortion function defined in 2.8 is a lower bound on the rate realizable by the block quantization scheme of Figure 1.4 (assuming a Gaussian input) since it is shown in [22] to be the lower bound for any coding scheme applied to the stationary Gaussian vector source with memory. Since no assumptions are made regarding the source density function, 2.8 becomes an upper bound on the rate distortion function for the multispectral source. It nevertheless serves as a measure of the efficiency of the block quantization technique. It is recognized however that in reality the multispectral source, whether aircraft or satellite, is not a Gaussian source (although it does have memory), and as such has a rate distortion function uniformly less than that of 2.8.

2.3. Sources of Error

The two sources of error in the system of Figure 1.4 are (1) the transform encoder (for $n<N$), and (2) the quantizer. That this is true may be shown in the following manner for the mean square error criteria. First, define the mean square quantization error d_q

$$d_q = E\{||\underline{Y}-\underline{Y}^*||^2\} \qquad 2.9a$$

$$= \sum_{i=1}^{n} E(y_i - y_i^*)^2 \qquad n \leq N \qquad 2.9b$$

and the "sampling" error due to the transform encoder

$$d_s = E\{||\underline{X}-\underset{\sim}{T}^t\underline{Y}||^2\} \qquad 2.10a$$

$$= E\{(\underline{X}-\underset{\sim}{T}^t\underline{Y})^t(\underline{X}-\underset{\sim}{T}^t\underline{Y})\} \qquad 2.10b$$

Bringing the transpose inside the parenthesis gives

$$d_s = E\{(\underline{X}^t-\underline{Y}^t\underset{\sim}{T})(\underline{X}-\underset{\sim}{T}^t\underline{Y})\} \qquad 2.11$$

or

$$d_s = E\{\underline{X}^t\underline{X}-\underline{X}^t\underset{\sim}{T}^t\underline{Y}-\underline{Y}^t\underset{\sim}{T}\underline{X}+\underline{Y}^t\underset{\sim}{T}\underset{\sim}{T}^t\underline{Y}\} \qquad 2.12$$

However, $\underset{\sim}{T}\underline{X} = \underline{Y}$, $\underline{X}^t\underset{\sim}{T}^t = \underline{Y}^t$, and $\underset{\sim}{T}\underset{\sim}{T}^t = \underset{\sim}{I}_n$, where $\underset{\sim}{I}_n$ is the nxn identity matrix. Thus 2.12 becomes

$$d_s = E\{\underline{X}^t\underline{X}-\underline{Y}^t\underline{Y}\} \qquad 2.13a$$

$$= \sum_{i=1}^{N} E(x_i^2) - \sum_{i=1}^{n} E(y_i^2) \qquad 2.13b$$

Define $\sigma^2_{y_i} = E(y_i^2)$, the variance of the random transform coefficient y_i. Then 2.13 becomes

$$d_s = \sum_{i=1}^{N} \sigma^2_{x_i} - \sum_{i=1}^{n} \sigma^2_{y_i} \qquad 2.14$$

It is easily shown (see [24]) that for the class of transformations considered in this study (i.e., orthogonal transformations)

$$\sum_{i=1}^{N} \sigma^2_{x_i} = \sum_{i=1}^{N} \sigma^2_{y_i} \qquad 2.15$$

Thus 2.14 may be written

$$d_s = \sum_{i=n+1}^{N} \sigma^2_{y_i} \qquad 2.16$$

Therefore the error due to the transform encoder is equal to the sum of the variances of the N-n discarded transform coefficients y_i, i=n+1, n+2, ..., N.

It is next shown that the total error d as defined in 2.1 is the sum of the sampling and quantization error. That is

$$d = d_s + d_q \qquad 2.17a$$

$$= \sum_{i=n+1}^{N} \sigma^2_{y_i} + \sum_{i=1}^{n} E\{(y_i - y_i^*)^2\} \qquad 2.17b$$

Expanding 2.1 gives

$$d = E\{(\underset{\sim}{X}-\underset{\sim}{T}^t\underset{\sim}{Y}^*)^t(\underset{\sim}{X}-\underset{\sim}{T}^t\underset{\sim}{Y}^*)\} \qquad 2.18a$$

$$= E\{\underset{\sim}{X}^t\underset{\sim}{X}-\underset{\sim}{Y}^{*t}\underset{\sim}{T}\underset{\sim}{X}-\underset{\sim}{X}^t\underset{\sim}{T}^t\underset{\sim}{Y}^*+\underset{\sim}{Y}^{*t}\underset{\sim}{T}\underset{\sim}{T}^t\underset{\sim}{Y}^*\} \qquad 2.18b$$

Recall that $\underset{\sim}{T}\underset{\sim}{X} = \underset{\sim}{Y}$ and $\underset{\sim}{X}^t\underset{\sim}{T}^t = \underset{\sim}{Y}^t$, and $\underset{\sim}{Y}^{*t}\underset{\sim}{Y} = \underset{\sim}{Y}^t\underset{\sim}{Y}^*$. Thus 2.18 becomes

$$d = E\{\underset{\sim}{X}^t\underset{\sim}{X}-2\underset{\sim}{Y}^{*t}\underset{\sim}{Y}+\underset{\sim}{Y}^{*t}\underset{\sim}{T}\underset{\sim}{T}^t\underset{\sim}{Y}^*\} \qquad 2.19$$

Changing from vector notation to summation of the vector elements gives

$$d = E\{\sum_{i=1}^{N} x_i^2 - 2\sum_{i=1}^{n} y_i^* y_i + \sum_{i=1}^{n} y_i^{*2}\} \qquad 2.20$$

That $\sum_{i=1}^{n} y_i^{*2} = \underset{\sim}{Y}^{*t}\underset{\sim}{T}\underset{\sim}{T}^t\underset{\sim}{Y}^*$ is true since $\underset{\sim}{T}\underset{\sim}{T}^t = \underset{\sim}{I}_n$. Rearranging terms in 2.20 gives

$$d = E\{\sum_{i=1}^{N} x_i^2 + \sum_{i=1}^{n} (y_i-y_i^*)^2 - \sum_{i=1}^{n} y_i^{*2} - \sum_{i=1}^{n} y_i^2 + \sum_{i=1}^{n} y_i^{*2}\} \qquad 2.21a$$

$$= E\{\sum_{i=1}^{N} x_i^2 - \sum_{i=1}^{n} y_i^2 + \sum_{i=1}^{n} (y_i-y_i^*)^2\} \qquad 2.21b$$

Bringing the expectation inside the brackets gives

$$d = \sum_{i=1}^{N} E\{x_i^2\} - \sum_{i=1}^{n} E\{y_i^2\} + \sum_{i=1}^{n} E\{(y_i-y_i^*)^2\} \qquad 2.22$$

Using the results of 2.14 and 2.15, 2.22 becomes

$$d = \sum_{i=n+1}^{N} \sigma_{y_i}^2 + \sum_{i=1}^{n} E\{(y_i - y_i^*)^2\} \qquad 2.23a$$

$$= d_s + d_q \qquad 2.23b$$

which is the desired result.

2.4. The Karhunen - Loéve Transformation

A discussion of the three transformations used in the data compression system outlined in Figure 1.4 is presented in this and the following section.

Several authors have shown that the linear transformation which minimizes the sampling error d_s for the correlated vector source with positive definite covariance matrix is the Karhunen - Loéve (K-L) transformation (see Kramer and Mathews [7], Palermo, et al., Appendix I [25], Wintz and Kurtenbach [26], Koschmann [56], or Loeve [29]). It is further shown in [8] to possess the added benefit of minimizing the quantization error d_q. The K-L transformation is also referred to as the eigenvector transformation, principal component transformation ([27] and Chapter II of [28]), or simply as the optimum transform.

The K-L transformation is an orthogonal transformation determined by the second order moments of the source. In this sense it is source dependent. It is this adaptation

to the source that gives the K-L transformation its unique ability to minimize the total error d. The transformation itself is a rotation of the source output $\underline{X}$ in N-space to a more favorable orientation with respect to the measurement co-ordinate system. This more favorable orientation is one in which the average energy E_s of the source is redistributed over the co-ordinates such that a larger percentage of E_s is distributed over fewer co-ordinates. Figure 2.1 demonstrates this rotation for the two dimensional case (N=2). It is evident from Figure 2.1 that the K-L coefficient y_1 has larger variance than either x_1 or x_2. If variance is considered to be a measure of information content, as it is in mean square error rate distortion analysis, then knowledge of y_1 conveys more information than does knowledge of either x_1 or x_2. For larger values of N it is often the case that a rather small number of transform coefficients contain a large percentage of the total source variance (energy) E_s. This packing of the source variance provides a means of reducing the number of coordinates required to reconstruct the original source output vector within a given error. Those transform coefficients having small variance are neglected (replaced by their mean values, which are constrained to be zero).

The K-L transformation is the unique orthogonal transformation that diagonalizes the source covariance matrix $\underset{\sim}{C}$.

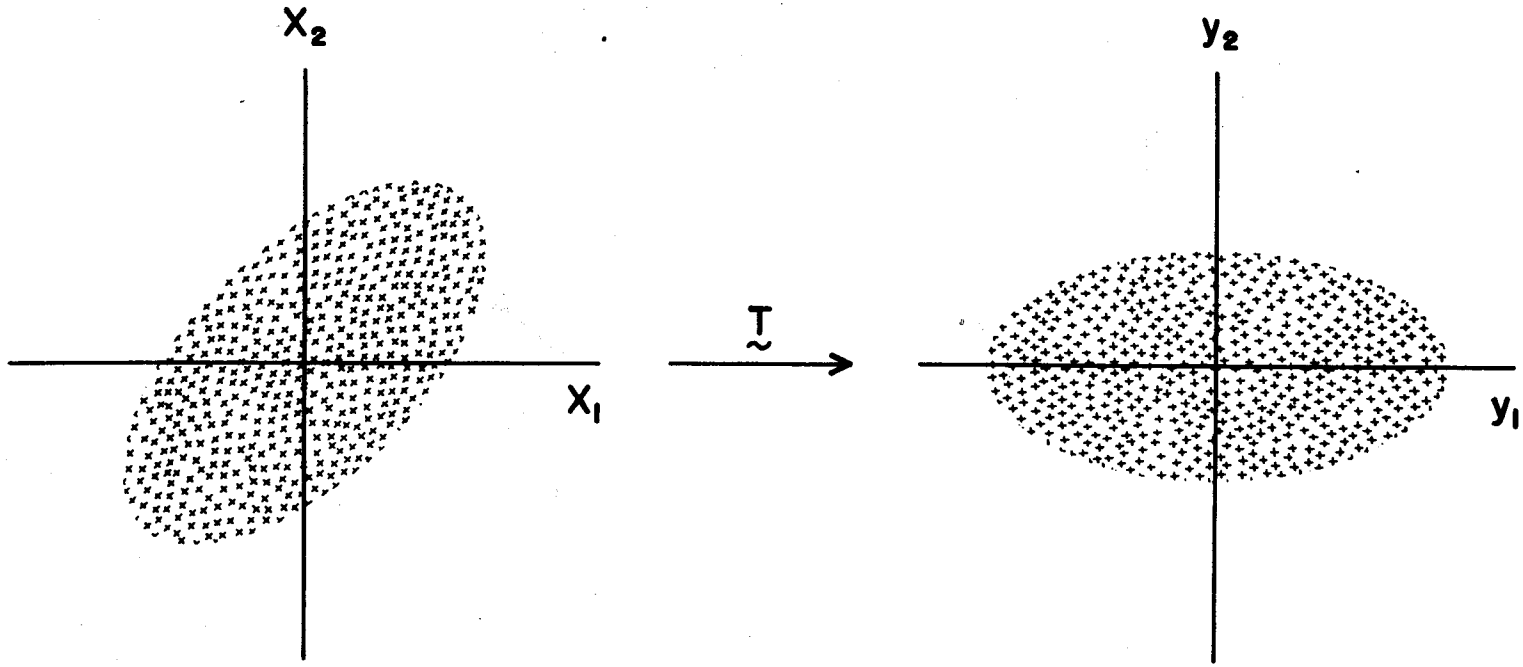


Figure 2.1. The Karhunen - Loéve Transformation with N=2.

$$\underset{\sim}{T}\underset{\sim}{C}\underset{\sim}{T}^{t} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{bmatrix} \qquad 2.24$$

where the λ_i, $i=1, 2, \ldots, N$ are the eigenvalues of $\underset{\sim}{C}$ and their order is chosen such that

$$\lambda_1 \geq \lambda_2 \cdots \geq \lambda_N \qquad 2.25$$

From 2.24 it is evident that the rows of $\underset{\sim}{T}$ are the N normalized solutions to the characteristic equation

$$\underset{\sim}{C}\underline{v}_i = \lambda_i \underline{v}_i \qquad i = 1, 2, \ldots, N \qquad 2.26$$

The covariance matrix for the transform coefficients is

$$E\{\underline{Y}\underline{Y}^{t}\} = E\{\underset{\sim}{T}\underline{X}(\underset{\sim}{T}\underline{X})^{t}\} \qquad 2.27a$$

$$= E\{\underset{\sim}{T}\underline{X}\underline{X}^{t}\underset{\sim}{T}^{t}\} \qquad 2.27b$$

$$= \underset{\sim}{T}\underset{\sim}{C}\underset{\sim}{T}^{t} \qquad 2.27c$$

Since $\underset{\sim}{T}\underset{\sim}{C}\underset{\sim}{T}^{t}$ is the diagonal matrix of 2.24 the transform coefficients are uncorrelated and their respective variances

are given by the ordered eigenvalues of $\underset{\sim}{C}$. The application of the vector $\underline{Y}$ as a feature vector in pattern recognition is discussed in [30] and is considered in this study in section 3.2.

2.5. The Fourier and Hadamard Transformations

The two additional transforms considered are the Fourier and Hadamard transforms. Both are non source dependent in that the set of N orthonomal basis vectors is fixed regardless of the source characteristics. Some variability in these transforms is possible through an appropriate definition of the source output dimensionality K. For example in image processing it is advantageous to define the source to be two dimensional and use the two dimensional Fourier or Hadamard Transforms. As described in later sections, the multispectral source is defined to be a three dimensional source and the three dimensional Fourier and Hadamard transforms are used. Consideration of the source dimensionality allows the encoder to take advantage of correlations existing between neighboring points in each dimension defined.

Several authors have applied the Fourier and Hadamard transforms in studies of two dimensional block quantization encoders. Landau and Slepian [31], Pratt, Kane, and Andrews [9], Kennedy, Clark, and Parkyn [32], Huang and Woods [33], and Habbi and Wintz [10] have reported results.

The rows of the Fourier Transformation matrix are the

sampled, harmonically related sine and cosine functions. The rows of the Hadamard matrix are the discrete version of an Nth order set of orthogonal Walsh functions [34,35, 36].

2.6. Equivalent Matrix Transformations

The usual technique used to perform a K-dimensional discrete linear transformation (Fourier, Hadamard, etc.) is the performance of K one-dimensional transformations - one in each dimension defined for the source. For example the two dimensional Fourier transform (K=2) would be obtained by pre and post multiplying the $N_1 x N_2$ source output "matrix" $[X]_{N_1,N_2}$ by the $N_1 x N_1$ matrix $\underset{\sim}{T}_1$ and the $N_2 x N_2$ matrix $\underset{\sim}{T}_2$ respectively

$$[Y]_{N_1,N_2} = \underset{\sim}{T}_1 [X]_{N_1,N_2} \underset{\sim}{T}_2^t \qquad 2.28$$

It is now shown that a K-dimensional transformation may be obtained by a single equivalent matrix multiplication of the re-defined source output.

First the source output tensor $[X]_{N_1,N_2,\ \ldots,\ N_K}$ is re-arranged into a N-dimensional ($N = N_1 N_2 \cdots N_K$) vector $\underline{X}$ by sequentially ordering the elements of $[X]_{N_1,\ \ldots,\ N_K}$ by varying the outermost indices first. For example in two dimensions (K=2) with $N_1 = N_2 = 2$

$$[X]_{2,2} = \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{bmatrix} \qquad 2.29$$

and $\underline{X}$ is therefore

$$\underline{X} = \begin{bmatrix} x_{1,1} \\ x_{1,2} \\ x_{2,1} \\ x_{2,2} \end{bmatrix} \qquad 2.30$$

For the K dimensional source an element of the transform coefficient tensor $[Y]_{N_1,N_2,\ \ldots,\ N_n}$ is given by

$$y_{i_1,i_2,\ \ldots,i_n} = \sum_{l_1=1}^{N_1} \sum_{l_2=1}^{N_2} \cdots \sum_{l_K=1}^{N_K} x_{l_1,l_2,\ldots,l_K} t_{i_1,l_1} t_{i_2,l_2} \cdots t_{i_K l_K} \qquad 2.31$$

where the t_{i_j,l_j} are elements of the transform matrix $\underset{\sim}{T}_j$

$$\underset{\sim}{T}_j = \begin{bmatrix} t_{1_j,1_j} & t_{1_j,2_j} & \cdots & t_{1_j,N_j} \\ t_{2_j,1_j} & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ t_{N_j,1_j} & & \cdots & t_{N_j,N_j} \end{bmatrix} \quad j=1,2,\ldots,K \qquad 2.32$$

If the elements of $[X]_{N_1,\ N_2,\ \ldots,\ N_K}$ are re-arranged according to the outermost index rule given above, then the transform coefficient tensor $[Y]_{N_1,\ N_2,\ \ldots,\ N_K}$ becomes an N element vector ($N = N_1 \cdot N_2 \cdot \cdot \cdot N_K$) with elements

$$y_m = \sum_{n=1}^{N} x_n t'_{m,n} \qquad m=1,\ 2,\ \ldots,\ N \qquad 2.33$$

where

$$m = \ell_1 + (\ell_2 - 1)N_2 + (\ell_3 - 1)N_3 + \ldots + (\ell_K - 1)N_K \qquad 2.34a$$

$$n = i_1 + (i_2 - 1)N_2 + (i_3 - 1)N_3 + \ldots + (i_K - 1)N_K \qquad 2.34b$$

and

$$t'_{m,n} = t_{i_1,l_1} t_{i_2,l_2} \cdots t_{i_K,l_K} \qquad 2.35$$

It is important to note that for each value of m, m=1, 2, ..., N there exists one and only one possible set of integer values for the l_i, i=1, 2, ..., K. The same is true for n and the i_j, j=1, 2, ..., K.

Rewriting 2.33 gives

$$\underline{Y}' = \underline{T}'\underline{X} \qquad 2.36$$

where vectors $\underline{Y}$ and $\underline{X}$ are the re-arranged transform and output tensors respectively, and $\underline{T}'$ is the new equivalent transformation matrix with elements $t_{m,n}$, m = 1, 2, ..., N; n = 1, 2, ..., N; $N = N_1 N_2 \cdots N_K$.

The equivalent matrix $\underline{T}'$ is now shown to be an orthogonal matrix. From 2.34 and 2.35

$$\sum_{m=1}^{N} t'_{m,n} t'^{*}_{m,n} = \sum_{l_1=1}^{N_1} \sum_{l_2=1}^{N_2} \cdots \sum_{l_K=1}^{N_K} t_{i_1,l_1} t^{*}_{j_1,l_1} t_{i_2,l_2}$$

$$t^{*}_{j_2,l_2} \cdots t_{i_K,l_K} t^{*}_{j_K,l_K} \qquad 2.37a$$

$$= \sum_{l_1=1}^{N_1} t_{i_1,l_1} t^{*}_{j_1,l_1} \sum_{l_2=1}^{N_2} t_{i_2,l_2} t^{*}_{j_2,l_2} \cdots$$

$$\sum_{l_K=1}^{N} t_{i_K,l_K} t^{*}_{j_K l_K} \qquad 2.37b$$

where * indicates the complex conjugate.

The orthogonality of the original transformation matrices $\underset{\sim}{T}_j$, j=1, 2, ..., K requires that

$$\sum_{l_p=1}^{N_p} t_{i_p,l_p} t^*_{j_p,l_p} = \delta_{i_p,j_p} \qquad p=1, 2, \ldots K \qquad 2.38$$

where δ_{i_p,j_p} is the Kronecker delta. Thus 2.37 becomes

$$\sum_{m=1}^{N} t'_{m,n} t'^*_{m,j} = \delta_{m,j} \qquad 2.39a$$

Similarly it may be shown that

$$\sum_{n=1}^{N} t'_{m,n} t'^*_{j,n} = \delta_{m,j} \qquad 2.39b$$

From 2.39 it is evident that $\underset{\sim}{T}'$ is an orthogonal matrix, which is the desired result.

2.7. Application to the Hadamard and Fourier Transforms

An interesting consequence of the results of section 2.6 is that the equivalent matrix for the K-dimensional Hadamard transform is again a Hadamard matrix - the NxN, ($N=N_1 N_2 \cdots N_K$) Hadamard matrix. That this is true may be seen from equations 2.35 and 2.39. The entries of $\underset{\sim}{T}'$ must

be $\pm (1/N)^{K/2}$ from 2.35. This and the orthogonality of $\underset{\sim}{T}'$ (2.39) combine to require $\underset{\sim}{T}'$ to be the Nth order Hadamard Matrix.

The equivalent matrix for the K-dimensional Fourier Transform, though an orthogonal matrix, is not the Nth order Fourier transform. This is shown by an example.

Consider the two-dimensional Fourier transform (K=2) of equation 2.28. The two transforms, one for each dimension of the source, are given by 2.32 with

$$t_{i_j,l_j} = \frac{1}{\sqrt{N_j}} \exp\left[\frac{2\pi}{\sqrt{N_j}}(i_j-1)(l_j-1)\right] \qquad \begin{matrix} i_j=1,2,\ldots,N_j \\ l_j=1,2,\ldots,N_j \\ j=1,2 \end{matrix} \qquad 2.40$$

and from 2.35

$$t'_{m,n} = t_{i_1,l_1} t_{i_2,l_2} \qquad \begin{matrix} m = 1,2,\ldots,N \\ n = 1,2,\ldots,N \end{matrix} \qquad 2.41$$

The following relationships from equation 2.34 are true

$$\begin{array}{ll} m = 1 \Rightarrow l_1=1,\ l_2=1 & n = 1 \Rightarrow i_1=1,\ i_2=1 \\ \quad 2 \Rightarrow l_1=2,\ l_2=1 & \quad 2 \Rightarrow i_1=2,\ i_2=1 \\ \quad \vdots & \quad \vdots \\ N_1 \Rightarrow l_1=N_1,\ l_2=1 & N_1 \Rightarrow i_1=N_1,\ l_2=1 \end{array}$$

$$m = N_1+1 \Rightarrow l_1=1,\ l_2=2 \qquad n = N_1+1 \Rightarrow i_1=1,\ i_2=2$$

$$\vdots \qquad\qquad \vdots$$

$$N \Rightarrow l_1=N_1,\ l_2=N_2 \qquad N \Rightarrow i_1=N_1,\ i_2=N_2$$

2.42

Using the above relationships between m, n, i_1, i_2, l_1, l_2 and equations 2.40 and 2.41 gives the entries in the equivalent Fourier matrix with K=2

$$t'_{m,n} = \frac{1}{\sqrt{N_1}} \exp\left[\frac{2\pi}{N_1}(i_1-1)(l_1-1)\right] \cdot$$

$$\frac{1}{\sqrt{N_2}} \exp\left[\frac{2\pi}{N_2}(i_2-1)(l_2-1)\right] \qquad 2.43a$$

$$= \frac{1}{\sqrt{N_1 N_2}} \exp\left[\frac{2\pi}{N_1}(i_1-1)(l_1-1)+\frac{2\pi}{N_2}(i_2-1)(l_2-1)\right]$$

2.43b

where the relationships between the indices is provided by 2.42.

2.8. Quantization

The remaining source of error in the data compression system of Figure 1.4 is the quantizer. Its function is to map the infinite set of possible n-vectors $\underline{Y}$ into the finite set of n-vectors $\underline{Y}^*$, and thereby allow a finite time - bandwidth product or finite storage requirements.

The quantizer used in this study is the optimum non-uniformly spaced output level quantizer described by Max in [37] and Lloyd in [38], and generally referred to as the Lloyd-Max quantizer. This quantizer is optimum in the sense that the output levels are chosen such that the mean square quantization error (d_q of equation 2.9) is minimized, assuming a normally distributed input. The input to the quantizer is the vector $\underline{Y}$. Quantization is achieved through a bank on n quantizers, one for each y_i, i=1, 2, ..., n. The output of the quantizer is coded for digital transmission or storage by assigning a binary code word to each output level. The natural code is used in this study, although other codes have been investigated by Hayes and Bobilin [39], and Wintz and Kurtenbach in Chapter 3 of [40].

The ith quantizer is referred to as an m_i - bit quantizer, reflecting the fact that it has 2^{m_i} possible output levels, and that each output level is coded with m_i bits. Figure 2.2 is an example of the ith m_i bit quantizer with m_i=2. The quantizer assigns the output value $y_i^*=v_j$ when $u_{j-1}<y_i\leq u_j$. In practice the y_i are normalized to have unit variance and the u_j and v_j given in [37] for $m_i \leq 5$ and in [40] for $6 \leq m_i \leq 9$ are used.

Results presented in [8] and [26] show that the quantization error for the ith quantizer with a Gaussian input may be closely approximated by

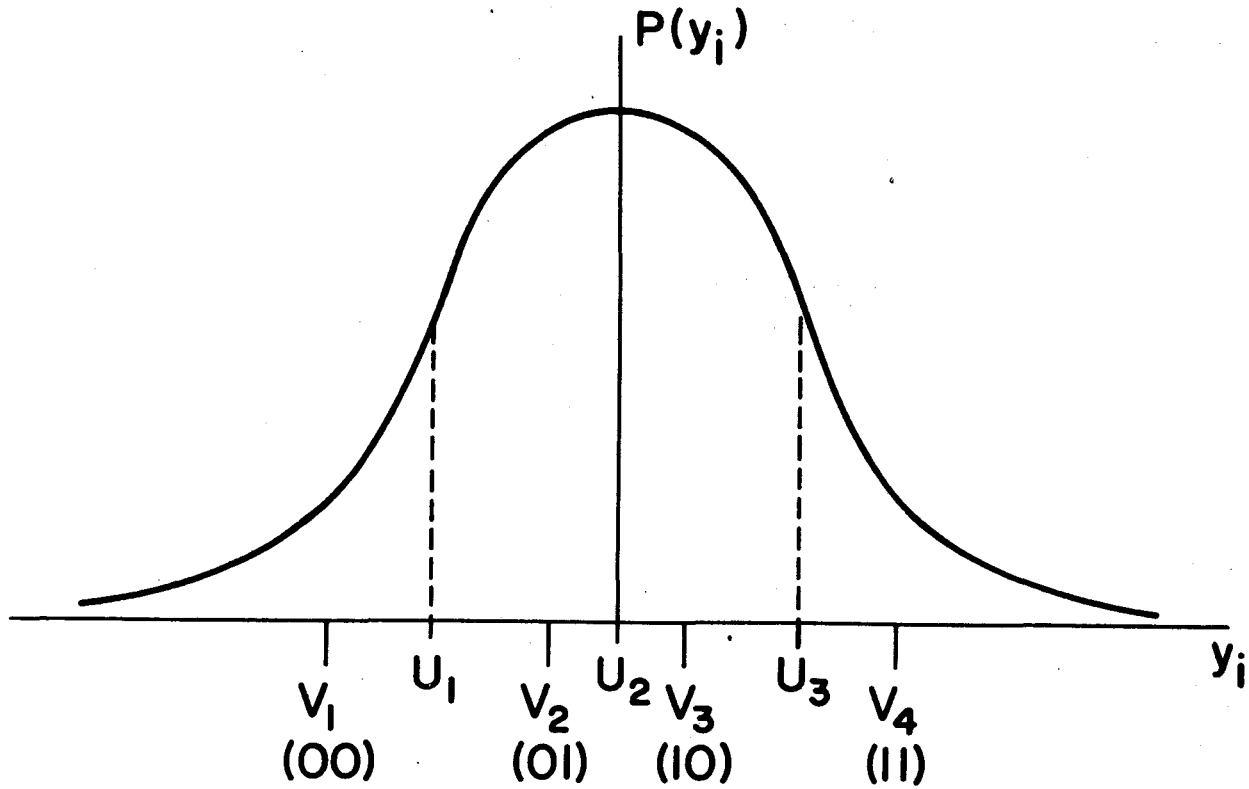


Figure 2.2. The i^{th} Lloyd - Max Non-Uniform Quantizer with $m_i=2$.

$$d_q\Big|_{i\text{th quantizer}} = \sigma_{y_i}^2 \, C^{-2m_i} \qquad 2.44$$

where $\sigma_{y_i}^2$ is the variance of the transform coefficient y_i and C is a constant equal to $10^{1/4}$, or approximately 1.78.

2.9. Minimization of the Total System Error

The results of section 2.3 show that the total system error d is the sum of the quantization error d_q and the sampling error d_s where, as in 2.14

$$d_s = E_s - \sum_{i=1}^{n} \sigma_{y_i}^2 \qquad 2.45$$

and the total quantization error d_q is the sum of the error due to each of the n quantizers

$$d_q = \sum_{i=1}^{n} \sigma_{y_i}^2 \, C^{-2m_i} \qquad 2.46$$

Combining 2.45 and 2.46 gives the total error

$$d = E_s - \sum_{i-1}^{n} \sigma_{y_i}^2 + \sum_{i=1}^{n} \sigma_{y_i}^2 \, C^{-2m_i} \qquad 2.47a$$

$$= E_s - \sum_{i=1}^{n} \sigma_{y_i}^2 \, (1-C^{-2m_i}) \qquad 2.47b$$

It is evident from 2.47 that the total error is a function of (1) n- the number of transform coefficients retained, and (2) m_i- the number of bits assigned to the ith quantizer. The sampling error decreases with increasing n while the quantization error increases with n. It is desirable to find that value of n and resulting bit distribution $\{m_i\}_{i=1}^{n}$ which minimizes the total error d. The following minimization problem is stated and an approximate solution is presented.

Problem: Given that a finite number of bits, say M_b, are to be distributed in some manner to n Lloyd-Max quantizers, i.e.,

$$\sum_{i=1}^{n} m_i = M_b \qquad 2.48$$

determine the optimum number of samples n (ordered according to decreasing variance) and corresponding bit distribution $\{m_i\}_{i=1}^{n}$ that minimizes the total system error $d=d_q+d_s$ of equation 2.47

The problem may be re-stated consisely as

$$\min_{(n,m_i)} \left\{ E_s - \sum_{i=1}^{n} \sigma_i^2 (1-C^{-2m_i}) \right\} \qquad 2.49$$

with the constraint

$$\sum_{i=1}^{n} m_i = M_b \qquad 2.50$$

where the subscript on the coefficient variance $\sigma^2_{y_i}$ has been omitted. By definition (equation 1.4) E_s is a constant and

$$E_s \geq \sum_{i=1}^{n} \sigma_i^2 (1-C^{-2m_i}) \qquad 2.51$$

with the equality when n=N and $M_b \to \infty$. An equivalent problem is then

$$\max_{(n,m_i)} \left\{ \sum_{i=1}^{n} \sigma_i^2 (1-C^{-2m_i}) \right\} \qquad 2.52$$

subject to

$$\sum_{i=1}^{n} m_i = M_b \qquad 2.53$$

Treating the m_i and n as continuous variables and introducing the undetermined Lagrange multiplier λ gives the following set of equations

(1) $$\left\{ \frac{\partial}{\partial m_j} \left[\sum_{i=1}^{n} \sigma_i^2 (1-C^{-2m_i}) + \lambda \sum_{i=1}^{n} m_i \right] = 0 \right\}_{j=1}^{n} \qquad 2.54a$$

$$(2) \qquad \frac{\partial}{\partial n}\left\{\sum_{i=1}^{n}\sigma_i^2(1-C^{-2m_i}) + \lambda\sum_{i=1}^{n}m_i\right\} = 0 \qquad 2.54b$$

$$(3) \qquad \sum_{i=1}^{n}m_i = M_b \qquad 2.54c$$

Equation 2.54a represents n equations while 2.54b and c represent one equation each. There are then n+2 equations and n+2 unknowns $(n, \lambda, \{m_i\}_{i=1}^{n})$.

Carrying out the partial differentiation indicated in 2.54a gives

$$\sigma_j^2(2\ln C)C^{-2m_j} + \lambda = 0 \qquad 2.55$$

Similarly 2.54b becomes

$$\sigma_n^2(1-C^{-2m_n}) + \lambda m_n = 0 \qquad 2.56$$

by approximating 2.54b with

$$\frac{\partial}{\partial n}\left\{\int_1^n \sigma^2(i)(1-C^{-2m(i)})di + \lambda\int_1^n m(i)di\right\} = 0 \qquad 2.57$$

to give 2.56.

Solving for m_j in 2.55 gives

$$m_j = \frac{1}{2}\log_C\left[\frac{-2A\sigma_j^2}{\lambda}\right] \qquad 2.58$$

where A=ln(C). Substituting 2.58 into 2.54c

$$M_b = \frac{1}{2}\log_C\left\{\left(\frac{-2A}{\lambda}\right)^n \prod_{j=1}^{n} \sigma_j^2\right\} \qquad 2.59$$

Solving 2.59 for λ and substituting the result into 2.58 gives

$$m_j = \frac{1}{2}\log_C\left\{\sigma_j^2\left[\frac{C^{2M_b}}{\prod_{i=1}^{n}\sigma_i^2}\right]^{\frac{1}{n}}\right\} \qquad 2.60a$$

$$= \frac{1}{2}\log_C\sigma_j^2 + \frac{M_b}{n} - \frac{1}{2n}\sum_{i=1}^{n}\log_C\sigma_i^2 \qquad 2.60b$$

Evaluating 2.60 at j=n and substituting into 2.56

$$\sigma_n^2 - \left[\frac{\prod_{i=1}^{n}\sigma_i^2}{C^{2M_b}}\right]^{\frac{1}{n}}\left[1+\ln(\sigma_n^2)+\frac{2AM_b}{n} - \frac{1}{n}\sum_{i=1}^{n}\ln(\sigma_i^2)\right]=0 \qquad 2.61$$

The value of n satisfying 2.61 and the corresponding bit distribution given by 2.60 represent the approximate solution to the minimization problem described earlier. The solution is approximate since the $\{m_i\}_{i=1}^{n}$ and n are assumed to be continuous variables while in reality they may take on only integer values. The approximate solution is achieved in the following manner: Given the NxN transformation $\underset{\sim}{T}$ the transform vector $\underline{Y}$ is created by the transform encoder. The

number of bits per vector element (i.e., the rate R) is specified. Thus

$$M_b = N \cdot R \tag{2.62}$$

Equation 2.61 is then used to determine the value of n which most closely satisfies the equality. That is

$$n_{opt} = \min_{1 \le n \le N} \left\{ \left| \sigma_n^2 - \left(\frac{\prod_{i=1}^{n} \sigma_i^2}{C^{2Mb}} \right)^{\frac{1}{n}} \left(1 + \ln(\sigma_n^2) + \frac{2AM_b}{n} - \frac{1}{n} \sum_{i=1}^{n} \ln(\sigma_i^2) \right) \right| \right\} \tag{2.63}$$

The n_{opt} determined by 2.63 is then used in 2.60 to compute the resulting bit distribution over the $n=n_{opt}$ Lloyd-Max quantizers. The values of the m_i generated by 2.60 are not, in general, integers. The integer values must therefore be chosen according to some rule. The rule used in this study is the following: (1) round-off each m_i to the nearest integer, (2) if the resulting rate R_I exceeds or is less than the specified rate R then either: (a) remove bits, one at a time from all possible y_i and choose those which result in the smallest increase in d, or (b) add bits, one at a time

to all possible y_i and choose those which result in the largest decrease in d.

It is noted that equation 2.60 for the bit distribution is also the solution to the less restrictive minimization problem considered in [8] and [40]. In that problem M_b and n are fixed, and the best set of m_i is the desired result. The solution is again 2.60.

2.10. The Markov Source - An Example

As an example of the sampling and bit distribution procedure described in the last section the time discrete, zero mean, one dimensional (K=1), unit variances, stationary, Gaussian Markov vector source (N=100) is considered. The elements of the source covariance matrix $\underset{\sim}{C}$ are then

$$c_{i,j} = \exp(-\alpha|i-j|) \qquad 2.64$$

Since $\sigma^2_{x_i} = 1$, the $c_{i,j}$ are also the correlations between the x_i.

The transformation $\underset{\sim}{T}$ is chosen to be the 100x100 K-L transformation, implying that the variances of the transform coefficients are the eigenvalues of $\underset{\sim}{C}$.

Figures 2.3, 2.4 and 2.5 summarize the results obtained with the above source definitions and $\alpha = 0.05$. Figure 2.3 shows the optimum number of samples determined by 2.63 as a function of the data rate. The lower curve represents the actual n_{opt} used after applying the bit distribution rule of

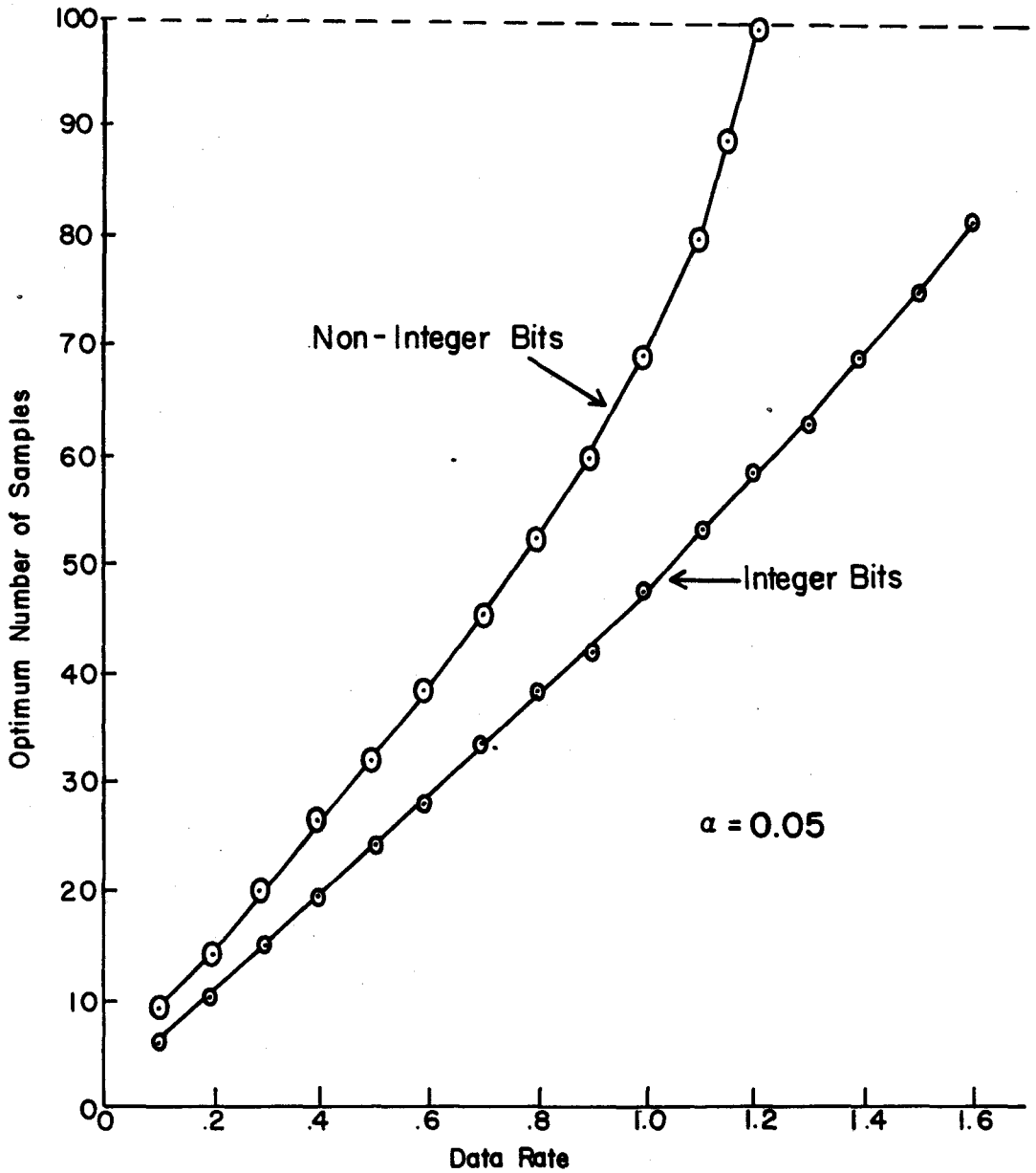


Figure 2.3. Optimum Number of Samples Versus Data Rate for the Markov Source.

section 2.9 since some m_i become zero when the integer constraint is imposed. Figure 2.4 is the resulting bit distribution over the n_{opt} Lloyd-Max quantizers as determined by 2.60 with data rate R = 1.0 bits/vector element. The step-wise curve represents the results of the integer bit distribution rule. As a result only 24 samples are retained as opposed to n_{opt} = 32 determined by 2.63. Figure 2.5 is a plot of the data rates achieved with the Markov source using the above results and the transform coding, block quantization scheme of Figure 1.4. The lower curve is the theoretical lower bound for the Gaussian Markov source (i.e., its Rate-Distortion function). The two curves are relatively close, the actual rate being higher by a factor of only approximately 1.25 over the distortion range considered. Also included are the data rates achieved using single sample [58] quantization of the original source output, referred to as standard PCM.

2.11. Selection of the Optimum Block Size

In section 2.9 the total system error d is minimized over n and the m_i, with M_b and N fixed. The problem of selecting the optimum block size for the K-dimensional source is now considered.

If the source is defined to be K-dimensional and if the maximum number of elements in $\underline{X}$ is fixed at N (due to hardware constraints, processing time limitations, etc.), then

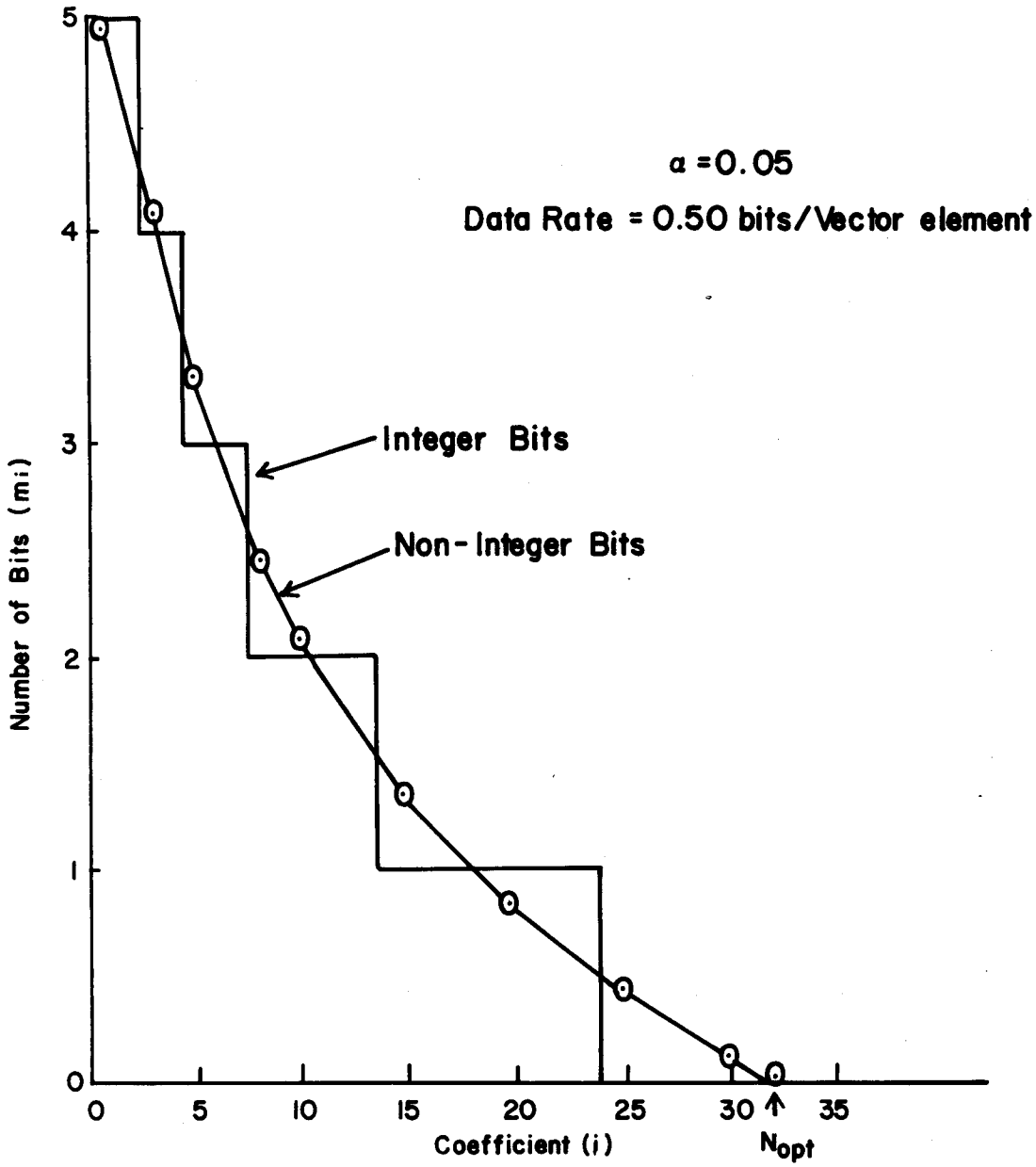


Figure 2.4. Bit Distribution over the n_{opt} Quantizers for the Markov Source.

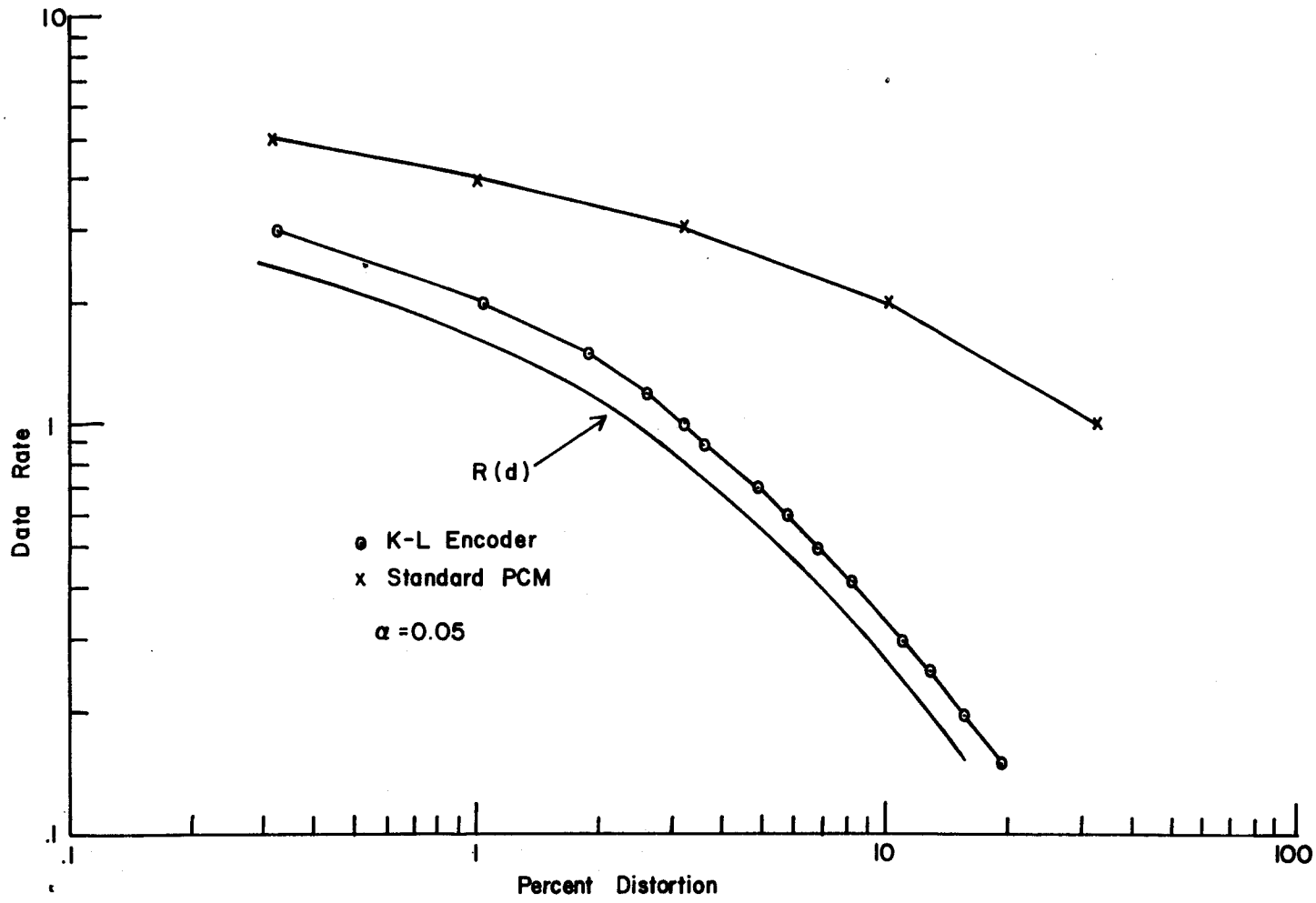


Figure 2.5. Data Rate Versus Percent Distortion for the Markov Source.

the question arises as to how many elements in each of the K-dimensions of the source should be selected. In other words, choose $N_1, N_2, \ldots, N_K$ to minimize the sampling error d_s with the constraint $N_1 \cdot N_2 \cdots N_K = N$. The problem may be re-stated using 2.14

$$\min_{\{N_i\}_{i=1}^{K}} \left\{ \sum_{i=1}^{N} \sigma_x^2 - \sum_{i=1}^{n} \sigma_{y_i}^2 \right\} \qquad 2.65$$

subject to

$$\prod_{i=1}^{K} N_i = N \qquad 2.66$$

An approximate solution to the above problem is presented in Appendix B under the following assumptions

(1) The source is time and amplitude continuous, (i.e., $\underline{X}$ becomes the continuous process $x(t_1, t_2, \ldots, t_K)$.

(2) The source statistics are stationary.

(3) The K-L sampler is used, i.e.,

$$y_i = \int_0^{N_1} \int_0^{N_2} \cdots \int_0^{N_K} x(\underline{t}) \phi_i(\underline{t}) d\underline{t} \qquad 2.67$$

where the $\phi_i(\underline{t})$ are eigenfunctions defined below.

The results of Appendix B show that the proper choice of the $\{N_i\}_{i=1}^{K}$ that minimize d_s must satisfy

$$N_i \oint_i \sum_{j=n+1}^{\infty} \lambda_j \phi_j^2(\underline{t}', N_i) d\underline{t}' = N_l \oint_l \sum_{j=n+1}^{\infty} \lambda_j \phi_j^2(\underline{t}', N_l) d\underline{t}' \quad 2.68$$

where $\oint_i$ and $\underline{t}'$ are defined in Appendix B and the $\phi_i(\cdot)$ are the eigenfunctions of the integral equations

$$\lambda_i \phi_i(\underline{t}) = \int_1^{N_1} \int_2^{N_2} \cdots \int_K^{N_K} R(\underline{t}-\underline{\tau}) \phi_i(\underline{\tau}) d\underline{\tau} \quad 2.69$$

and $R(\underline{t}-\underline{\tau})$ is the source autocorrelation function. Equation 2.68 may be simplified if the kernel $R(\underline{t}-\underline{\tau})$ is separable [41], i.e., if

$$R(\underline{t}-\underline{\tau}) = \prod_{i=1}^{K} R_i(t_i - \tau_i) \quad 2.70$$

Then 2.68 becomes

$$N_i \sum_{j=n+1}^{\infty} \lambda_j \theta_{j,i}^2(N_i) = N_l \sum_{j=n+1}^{\infty} \lambda_j \theta_{j,l}^2(N_l) \quad \forall i, l \quad 2.71$$

where

$$\prod_{i=1}^{K} \theta_{j,i}(t_i) = \phi_j(\underline{t}) \qquad 2.72$$

Although 2.68 and 2.71 are not explicit expressions for the N_i, a trial and error search could be used to select the best set. It is noted however, that for each new value of N_i the appropriate set of eigenvalues λ_i and eigenfunctions ϕ_i must be computed.

CHAPTER III

EXPERIMENTAL RESULTS PART I - AIRCRAFT SCANNER DATA

3.1. Introduction and Description of the Source

The results achieved using aircraft scanner data and the transform coding-block quantization system of Figure 1.4 with the sampling and quantization methods described in Chapter II are presented in this chapter. The data source is the digitized output of an airborne multispectral scanner flown over predominantly agricultural areas in the midwest.

The scanner itself [3,42] is an analog device sensitive to the electromagnetic energy emited or reflected from the particular area to which it is focused [52]. This sensitivity is restricted to several adjacent spectral bands or "channels". The following is a list of the spectral channels and their corresponding wavelength bands for the data set discussed in this chapter.

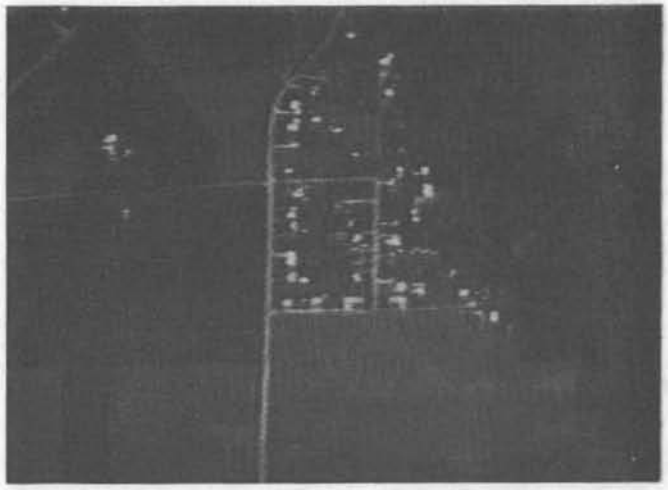
Table 3.1. Spectral Channels and Their Corresponding Wavelength Bands.

Channel	Wavelength Band (μm)
1	0.40 - 0.44
2	0.62 - 0.66
3	0.66 - 0.72
4	0.80 - 1.00
5	1.00 - 1.40
6	2.00 - 2.60

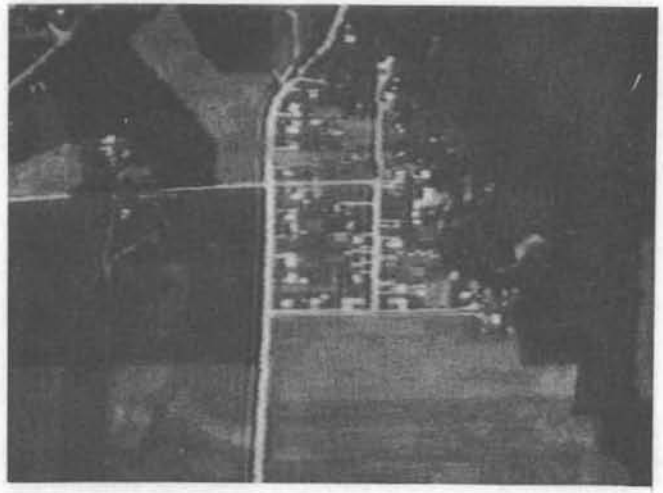
The altitude at which the scanner is normally flown ($\sim$ 3000 ft. A.G.L.) constrains the ground resolution to an approximately circular area having a diameter at nadir of $\sim$ 9 feet. The width of a typical data set (flightline) is approximately one mile. The sampling rate used in the digitization of the scanner data is such that the one mile width yields up to 444 samples per line. Flightline lengths are of course variable, but typically vary from one to twenty-five miles. One square mile of aircraft scanner data as described above contains on the order of 2×10^5 samples per channel.

The specific data set chosen for analysis is shown in Figure 3.1. This same area is shown in Figure 1.1 via panchromatic photograph. The data was gathered over Tippecanoe County, Indiana on the morning of June 30, 1969 at an altitude of 3000 ft. The entire flightline (PFL24) is 24 miles long while the section analyzed and shown in Figure 3.1 is approximately 0.9 miles by 0.7 miles (384 samples by 293 samples per channel). This section of the flightline is representative of the remaining 23 miles, and is quite typical of other aircraft scanner data sets taken over agricultural terrain. Areas of high and low spectral and spatial detail are represented as well as irregular field boundaries, roads, diagonal fields, etc.

Results obtained using the data compression scheme of Figure 1.4 are presented in the following sections. The



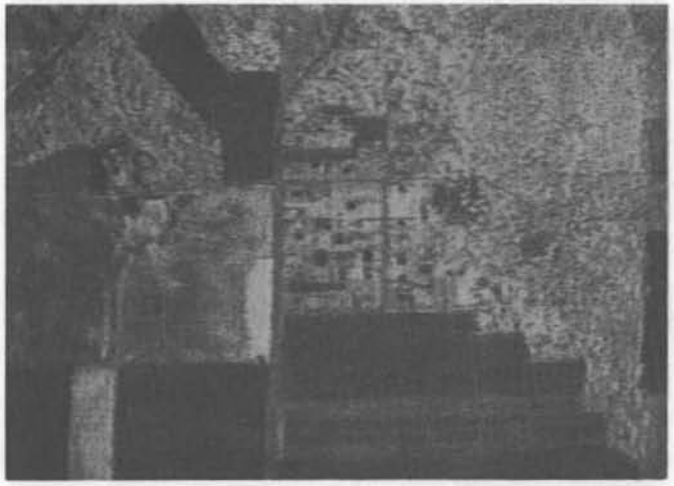
Channel 1



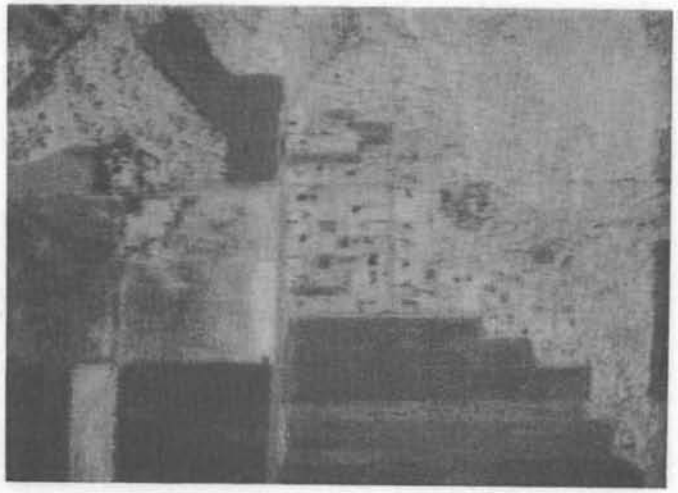
Channel 2

Channel 3

Figure 3.1. The Aircraft Scanner Data Set.



Channel 4



Channel 5

Channel 6

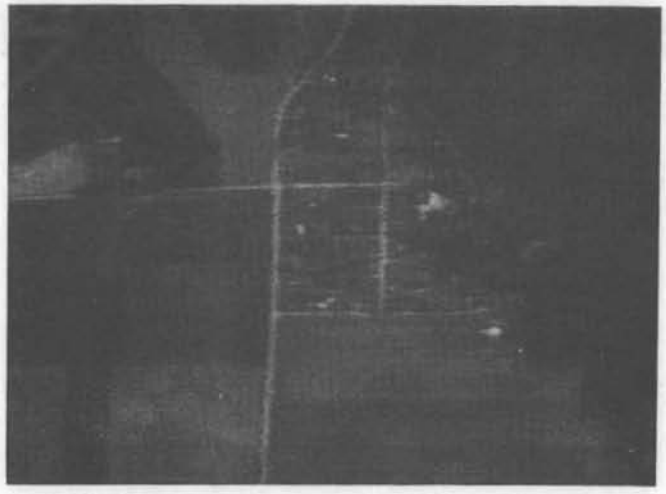


Figure 3.1. The Aircraft Scanner Data Set (continued).

multispectral source output is transformed using the Karhunen-Loéve (K-L), Fourier and Hadamard transformations in various combinations of the three dimensions of the source. In addition, the choice of which transform coefficients to retain and the corresponding bit distribution over the quantizers are optimized as described in section 2.9. The rate versus distortion curves achieved with each of the three transformations are compared to each other and also to (1) the rates achieved with standard single sample quantization (standard PCM) and (2) the theoretical minimum rates realizable assuming a gaussian source (i.e., the rate-distortion function described in section 2.2, equation 2.8).

The first and second order statistics of the data set are also presented in this section. The three dimensions of the source, the two spatial and one spectral, are considered for use with the one, two, and three dimensional transformations.

3.2. The Spectral Dimension and Principal Components

Each ground resolution point has an associated N-dimensional vector $\underline{X}$ whose elements are the N spectral channel intensities for that particular point. This concept is shown in Figure 3.2. The elements of $\underline{X}$ are, in general, correlated and each channel typically has a different mean and variance. This is evident in the six spectral images of Figure 3.1. The actual means and variances are given in Figure 3.3 along with the spectral correlation matrix.

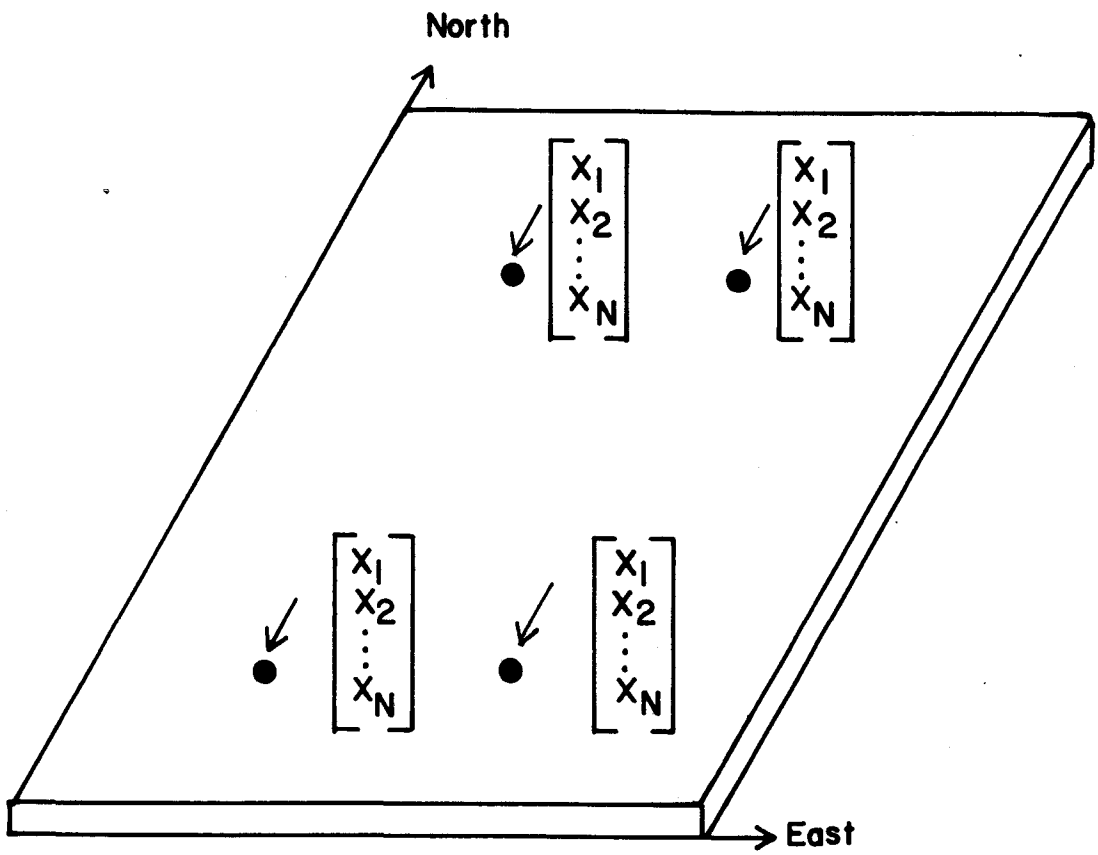


Figure 3.2. Ground Resolution Points and their Associated Channel Vectors.

Channel

	1	2	3	4	5	6
Mean	73.0	93.3	79.3	111.0	150.8	83.2
Variance	287	1136	1094	795	746	345

1.00					
0.90	1.00				
0.90	0.98	1.00			
-0.39	-0.45	-0.54	1.00		
-0.36	-0.41	-0.50	0.89	1.00	
0.73	-0.78	0.80	-0.46	-0.40	1.00

Correlation Matrix

Figure 3.3. First and Second Order Spectral Statistics.

The source output vector $\underline{X}$ is input to the transform encoder where it is pre-multiplied by the K-L, Fourier or Hadamard transformation matrix and then quantized. The output of the quantizer is then transmitted or stored. However with the spectral definition of the source output a special significance may be given to the quantizer output $\underline{Y}^*$ when the K-L transformation is used. The significance lies on the fact that $\underline{Y}^*$ may be used as a feature vector in the maximum likelihood classification of spatial characteristics (various types of vegetation, water, etc.) within the data set [30,43,44,45]. When used in this manner $\underline{Y}^*$ (or $\underline{Y}$) is commonly refered to as the vector of principal components.

The feature vector usually used in classification of multispectral data is the channel vector $\underline{X}$. However it has been found experimentally that a subset of the elements of $\underline{X}$ often gives classification results as acceptable as those obtained with the entire vector [20,46]. This is due to the fact that the correlations existing between channels reduces the number of channels having non-redundant information.

The choice of which subset of the elements of $\underline{X}$ to use as features is generally a difficult problem. The solution is usually based on a combination of (1) statistical inter-class distance measures resulting from an exhaustive search of all possible combinations of the elements of $\underline{X}$ to form a subset of given dimension [19,47], and (2) an intuitive

selection based on previous experience with perhaps similar data.

The principal component vector $\underline{Y}^*$ offers a possible solution to the feature selection problem. As discussed in section 2.4 the elements of $\underline{Y}^*$ are uncorrelated. In addition, the total variance of any ordered subset of the element of $\underline{Y}^*$ is greater than the total variance of the same number of elements of $\underline{X}$ [30]

$$\sum_{i=1}^{n} \sigma^2_{y_i} \geq \sum_{i} \sigma^2_{x_i} \qquad n \leq N \qquad 3.1$$

where the summation of the $\sigma^2_{x_i}$ is over any n values of i. The effects of the quantizer are ignored (i.e., $y_i = y_i^*$), since it is assumed that if $\underline{Y}^*$ is to be used as a feature vector the quantization will be so fine as to produce negligable error.

If variance is considered a measure of information content, then the feature set $\{y_i\}_{i=1}^{n}$ is at least as good as, and in general better than, any n-member subset of elements from $\underline{X}$ as indicated by 3.1.

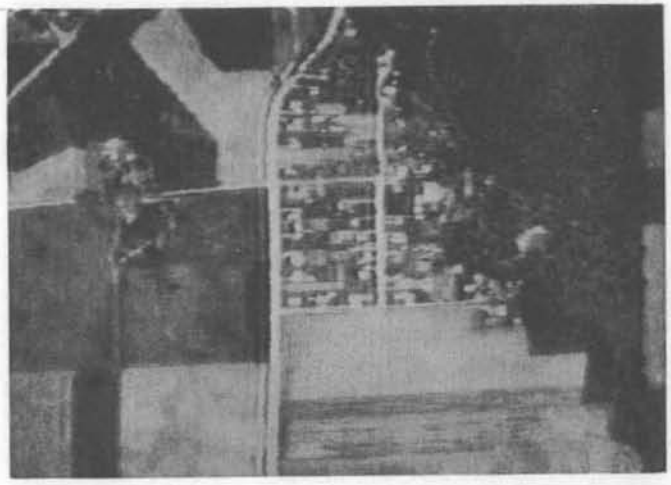
An additional favorable characteristic of $\underline{Y}$ is the fact that in the visual analysis of a multispectrally scanned area the analyst is limited to the observation of one spectral image at a time. However some areas may not produce significantly different responses in the particular

spectral channel chosen. A compromise must be made and a channel that is good on the average for the area of interest is selected. Some features will then be less distinguishable than if the best channel for these particular features had been chosen. The variance-packing property of the K-L transformation provides a new image in y_1 having variance greater than any of the original spectral channel images i.e.

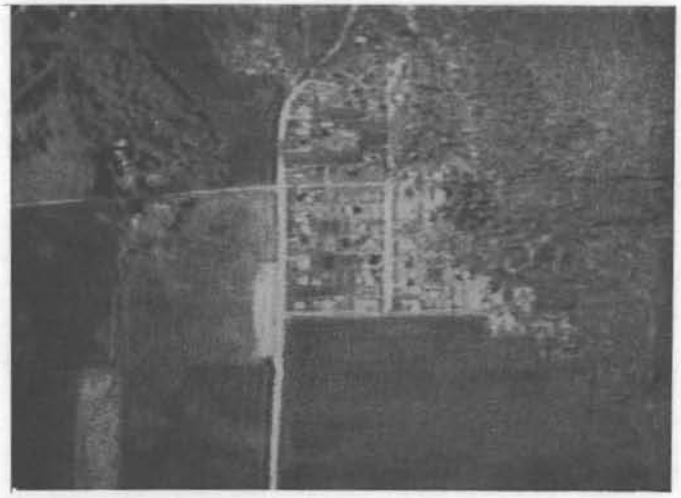
$$\sigma^2_{y_1} \geq \sigma^2_{x_i} \qquad i = 1,2, \ldots, N \qquad 3.2$$

This indicates that on the average more data variability (information) exists in the y_1 image of the area of interest and that visual analysis of the y_1 image may therefore be more productive.

The six (N=6) principal component images of the original data set of Figure 3.1 are presented in Figure 3.4 (i.e., y_1^*, y_2^*,..., y_6^* with $m_i = 8$ $\forall i$). As described in section 2.4 the K-L transformation is composed of the eigenvectors of the source covariance matrix. The 6x6 spectral covariance matrix for the aircraft scanner data set yields the following K-L transformation



Component 1



Component 2

Component 3

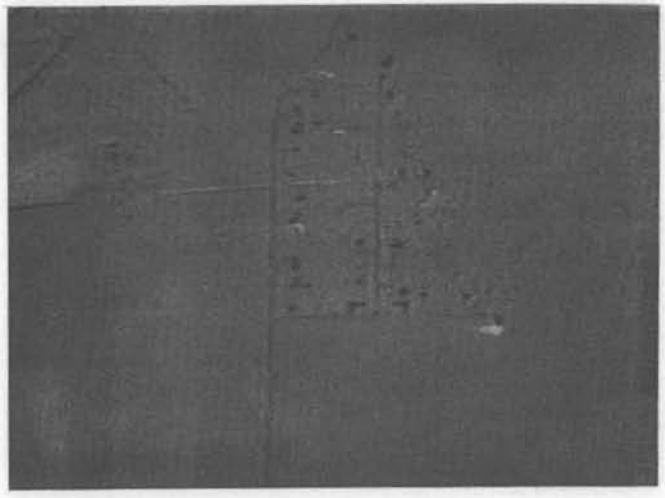
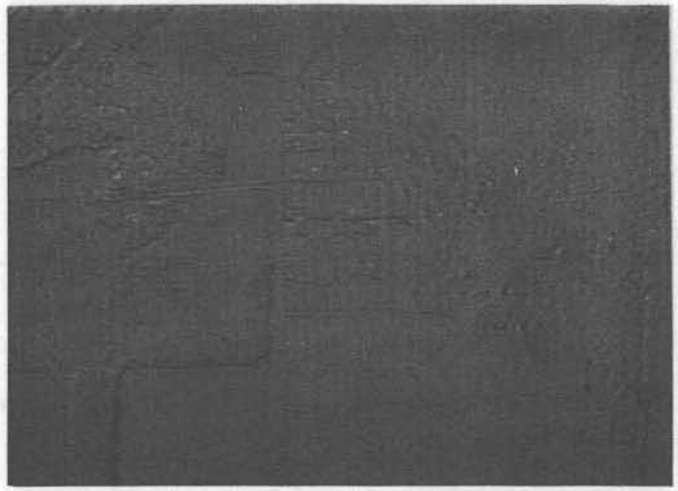
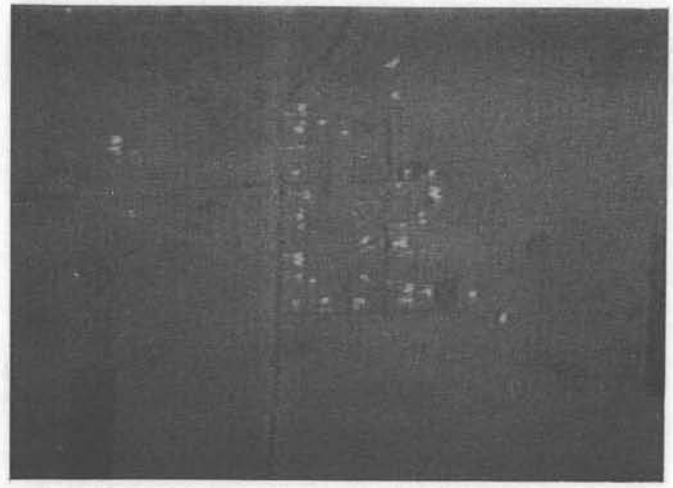


Figure 3.4. The Six Principal Component Images for the Aircraft Scanner Data Set.



Component 4



Component 5

Component 6

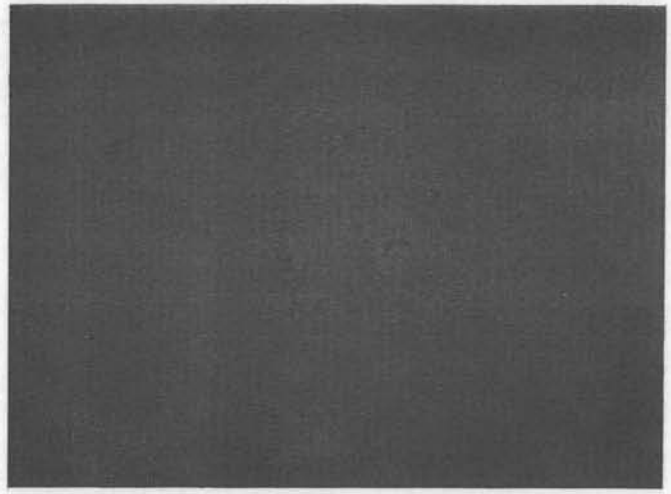


Figure 3.4. The Six Principal Component Images for the Aircraft Scanner Data Set (continued).

$$\underset{\sim}{T} = \begin{bmatrix} 0.25797 & -0.55725 & 0.56553 & -0.35458 & -0.32652 & 0.26599 \\ 0.18572 & 0.36388 & 0.24356 & 0.61212 & 0.62068 & 0.11763 \\ -0.08914 & -0.21988 & -0.15659 & -0.18867 & 0.23038 & 0.91133 \\ 0.01575 & 0.03032 & 0.06466 & -0.67578 & 0.67360 & -0.29022 \\ 0.93811 & -0.31709 & -0.13615 & -0.02113 & -0.01910 & -0.00768 \\ -0.10350 & -0.63817 & 0.75737 & 0.08281 & 0.03090 & -0.02462 \end{bmatrix} \quad 3.4$$

where the rows of $\underset{\sim}{T}$ are the eigenvectors of $\underset{\sim}{C}$ ordered as described by 2.25. The corresponding spectral eigenvalues are

Table 3.2. The Six Spectral Eigenvalues

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
3209.9	931.4	118.5	83.88	46.0	13.4

Each principal component (transform coefficient) y_i is a linear combination (weighted sum) of the original six spectral channels. The weights associated with each channel are given by the column entries of $\underset{\sim}{T}$. For example the first principal component image has the greatest weights associated with channel 2 (0.62-0.66μm) and channel 3 (0.66-0.72μm), two of the three channels in the visible wavelength band. The second principal component image favors channel 4 (0.80-1.00μm) and channel 5 (1.00-1.40μm), the two channels in the reflective infrared band. The third component image has greatest weight associated with channel 6 (2.00-2.60), the only channel in the thermal infrared region. These same trends are supported by a comparison of the principal

component images (Figure 3.4) with the original spectral images (Figure 3.1). It is also evident that the variances (eigenvalues of $\underset{\sim}{C}$) of the principal component images decrease rapidly as evidenced by the grey, low detail appearance of component images 3 through 6.

3.3 The Spectral Dimension and Fourier Components

The Fourier component images of the aircraft scanner data set are also of interest. They are obtained by transformation of the spectral channel vector $\underline{X}$ by the appropriate 6x6 Fourier matrix $\underset{\sim}{T}$ (The Hadamard matrix of order 6 does not exist[9] and thus prevents consideration of the Hadamard component images in this particular six-channel case. Hadamard images based on 12 channel data have been produced however, and are quite similar to the 12 Fourier component images).

The 6x6 Fourier transformation matrix is

$$\underset{\sim}{T} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 & -\frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ 1 & 1/2 & -1/2 & -1 & -1/2 & 1/2 \\ 1 & -1/2 & -1/2 & 1 & -1/2 & -1/2 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & 0 & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \\ 1/\sqrt{2} & -1/\sqrt{2} & 1/\sqrt{2} & -1/\sqrt{2} & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \qquad 3.5$$

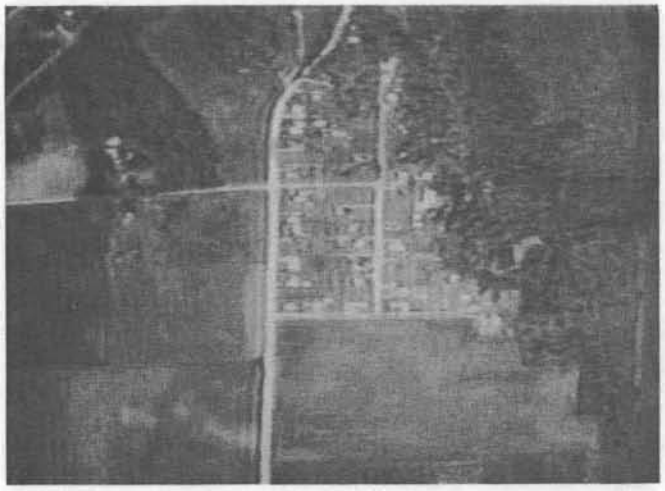
where the rows of $\underset{\sim}{T}$ are ordered by decreasing component variance as shown in Table 3.3.

Table 3.3. The Fourier Coefficient Variances

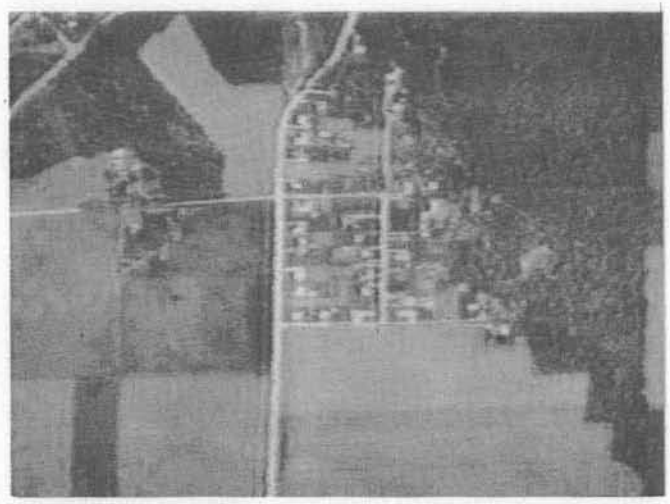
$\sigma^2_{y_1}$	$\sigma^2_{y_2}$	$\sigma^2_{y_3}$	$\sigma^2_{y_4}$	$\sigma^2_{y_5}$	$\sigma^2_{y_6}$
1218.8	1199.8	1017.2	486.8	421.1	59.3

The first Fourier image is proportional to the mean value of the total spectral response for the area. Each original spectral channel is weighted equally. Emphasis on specific wavelength bands begins with the second image. These weights, however, are non-source dependent and would be the same for any six wavelength bands. The six Fourier images are presented in Figure 3.5.

Although there is a packing of variance as with the principal component images, the Fourier image variances do not decrease as rapidly. In terms of the complex Fourier transform, the second and third images are the imaginary and real parts of the complex second harmonic coefficient respectively. Similarly the fourth and fifth images are the real and imaginary parts of the complex third harmonic coefficient. The sixth image is the real part (the imaginary is the negative of the imaginary part of the second harmonic, and is therefore redundant and not retained) of the complex fifth harmonic coefficient.



Component 1



Component 2

Component 3

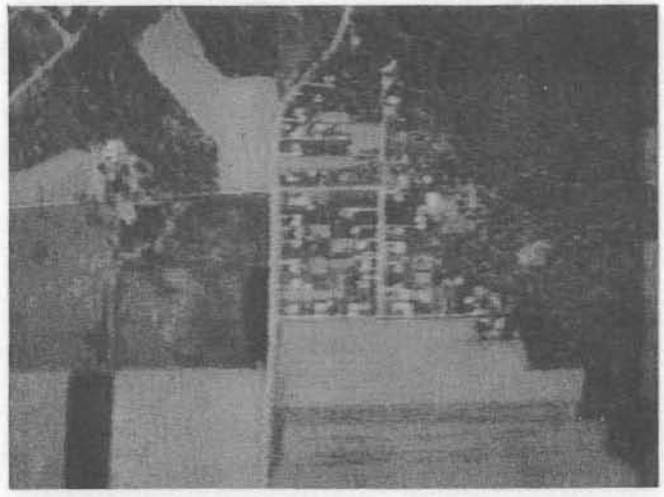
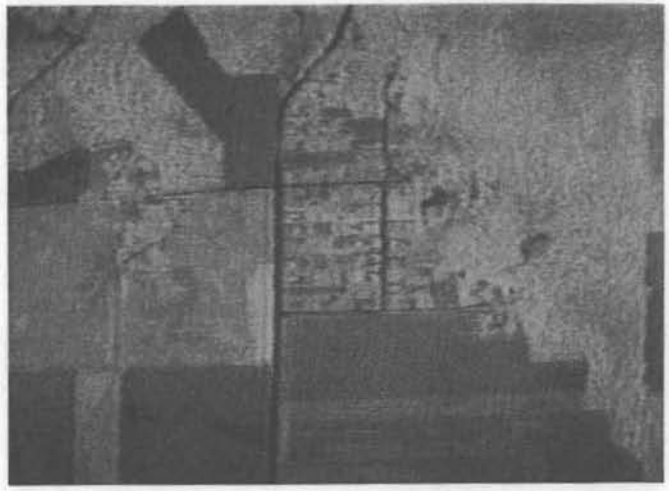
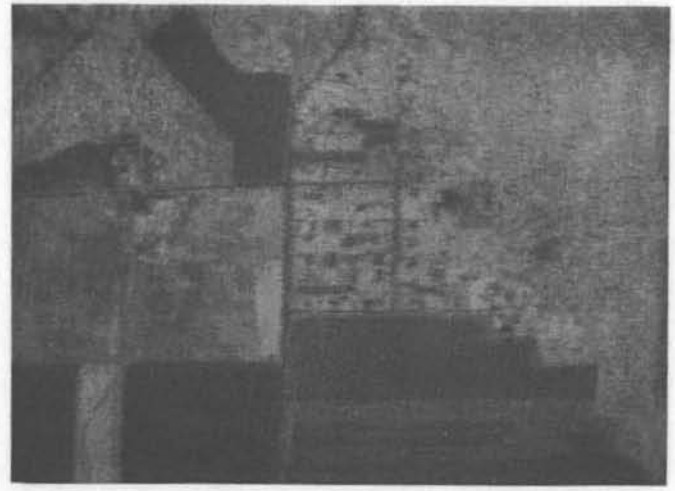


Figure 3.5. The Six Fourier Component Images for the Aircraft Scanner Data Set.



Component 4



Component 5

Component 6

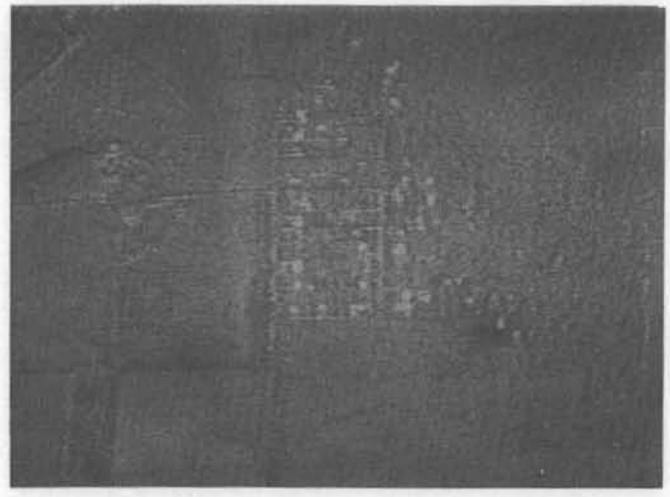


Figure 3.5. The Six Fourier Component Images for the Aircraft Scanner Data Set (continued).

Computer classification results using (1) the spectral channel vector $\underline{X}$, (2) the K-L coefficients or principal components, and (3) the Fourier coefficients are presented in Figure 3.6. A gaussian maximum likelihood decision rule is used to classify selected areas of the data set into one of six classes (various types of vegetation, roads, etc.). The sixth class is a null class into which all points having classification error probabilities greater than a specified threshold are placed. All points in the null class are considered errors and are used as such in computing classification accuracy. Training of the classifier is based on 1.3% of the total data set, and test results are based on classification of 16.2% of the total data set.

The abscissa in Figure 3.6 is the number of features (n) used in the classification. For each value of n an exhaustive search is conducted to determine the n-member set of features from $\underline{X}$ and $\underline{Y}$ giving the highest percent correct classification results. The features used are given below.

Table 3.4. Features Used for Classification

n	K-L	Channel	Fourier
1	1	2	3
2	1,2	3,5	1,3
3	1,2,3	2,5,6	1,2,3
4	1,2,3,4	1,2,5,6	1,2,3,4
5	1,2,3,4,5	1,2,3,5,6	1,2,3,4,6
6	1,2,3,4,5,6	1,2,3,4,5,6	1,2,3,4,5,6

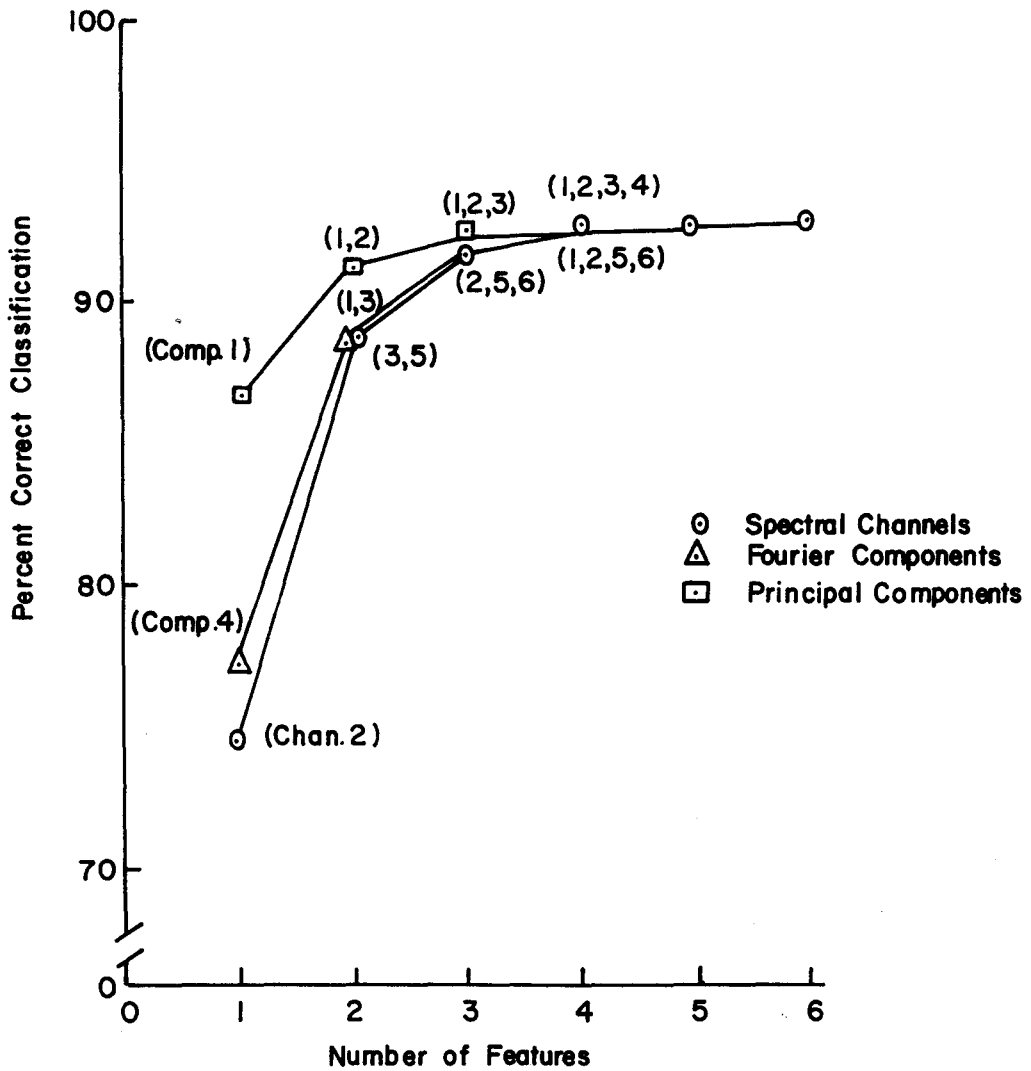


Figure 3.6. Classification Accuracy Versus Number of Features Using the Spectral, Fourier, and Principal Components.

The above list and Figure 3.6 point out the feature selection advantages inherent in the K-L or principal component transformation. For each value of n the first n principal components are the best selection. This is obviously not the case with the original spectral channels. The Fourier components are apparently not quite as consistant as the principal components, although (with the exception of n=1) the "first n" selection rule is close to being correct.

3.4. Data Compression in the Spectral Dimension

In the last section the vector $\underline{Y}$ or ($\underline{Y}^*$) is discussed in terms of pattern recognition and data interpretation. In this section it is but one step in the over-all (spectral) data compression system of Figure 1.4. It is the input to the bank of n Lloyd-Max quantizers. The choice of n and the number of bits m_i assigned to each quantizer is discussed in section 2.9. The application of those results to the channel vector $\underline{X}$ is presented in this section.

The data rates achieved using the K-L and Fourier encoders are shown in Figure 3.7. Also included are the theoretical lower bound R(d) for the gaussian source, and the actual data rates achieved using single sample encoding (standard PCM) of the source output vector $\underline{X}$. The number of transform coefficients retained and resulting m_i are listed in Appendix C. The K-L encoder lies below the Fourier encoder, while both are below standard PCM. However the

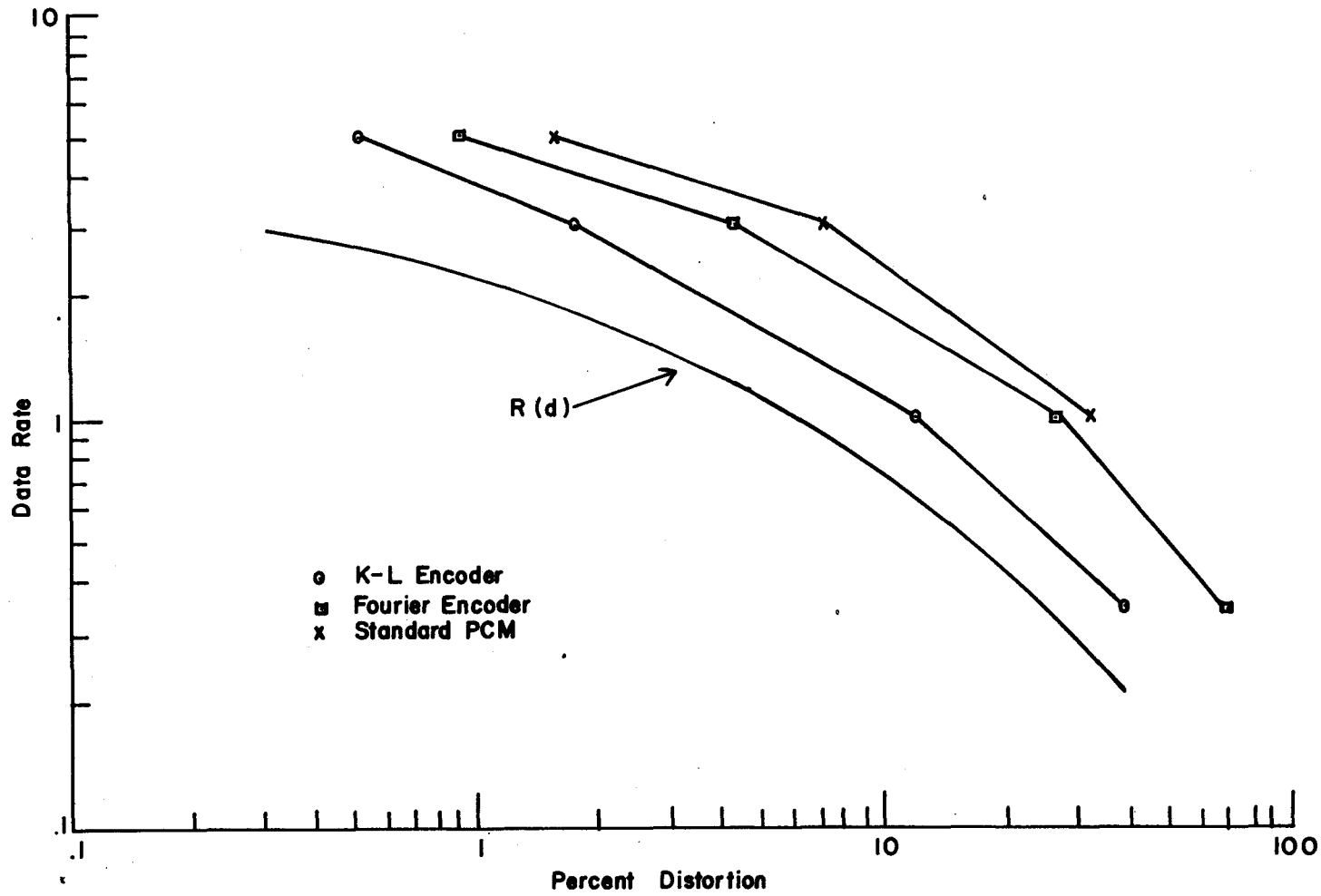
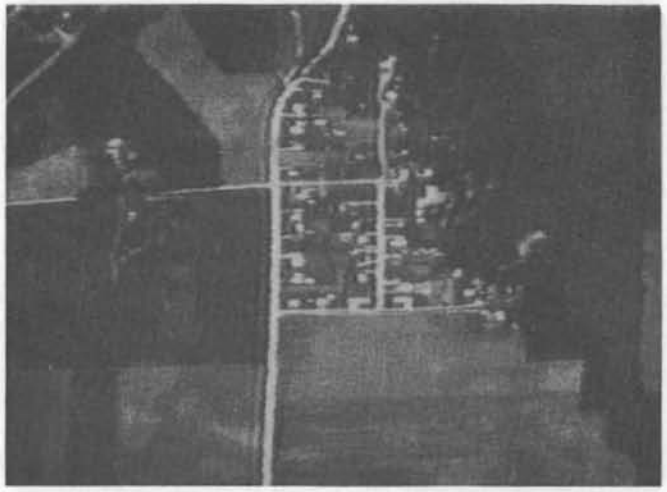


Figure 3.7. Data Rate Versus Percent Distortion Using the 1x1x6 Data Block.

Fourier encoder represents a relatively small improvement over standard PCM. A rate reduction by a factor of only 20% is realized for the distortion range considered (as compared to a factor of about 50% for the K-L encoder). This poor performance of the Fourier encoder is in general agreement with the results of the last section where the six Fourier images are presented (see Figure 3.5). It is evident from the images that the variances of the Fourier coefficients decrease slowly indicating that little variance packing has been accomplished.

Images representing the reconstructed output $\hat{\underline{X}}$ of the spectral data compression system are presented in Figure 3.8. An interesting effect is noticeable as the data rate decreases. For low data rates (i.e., around 0.5 bits/vector element) spatial features distinguishable in some of the original channels become visible in channels where they were originally not distinguishable. This effect can be seen in Figure 3.8 where a small field to the immediate left of the bottom road intersection is not visible in either the original channel 2 image or the reconstructed image at 2 bits/vector element, but does become visible in the 1.0 and 0.5 bit/vector element images. With low data rates the reconstructed source output vector $\hat{\underline{X}}$ tends to lie in the same direction in N-space as the first row $\underline{t}_1$ of the transformation matrix $\underset{\sim}{T}$. Any deviation the original $\underline{X}$ may have from that direction is lost as the data rate decreases. In



R = 2.0

R = 1.0

R = 0.5

Figure 3.8. Reconstructed Channel 2 Image Using the K-L Encoder at Three Different Data Rates (1x1x6 Data Blocks).

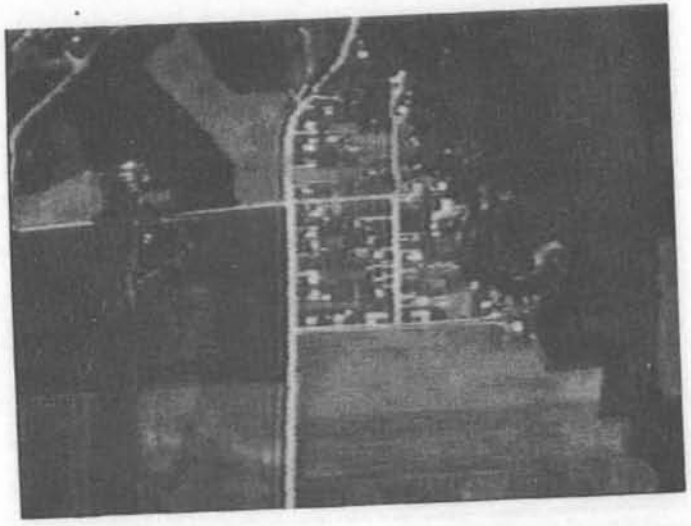
addition the magnitude of $\hat{\underline{X}}$ in the direction of $\underline{t}_1$ is determined by the projection of $\underline{X}$ onto $\underline{t}_1$. If one or more elements (channels) of $\underline{X}$ are relatively large then $\underline{X}\cdot\underline{t}_1$ may be dominated by these elements. The field discussed above is quite bright in channels 4 and 5 and thus becomes visible in the reconstructed channel 2 image ($\hat{x}_2$) at low data rates.

An "error image" between the 1 bit/vector element image and the original channel 2 image is shown in Figure 3.9. It is the square of the difference between $\underline{X}$ and $\hat{\underline{X}}$ for each resolution point. Black represents 0 error and white represents a squared error of 255 or greater. It is evident that most of the error occurs over the roads and buildings and relatively little error is present in the vegetated areas.

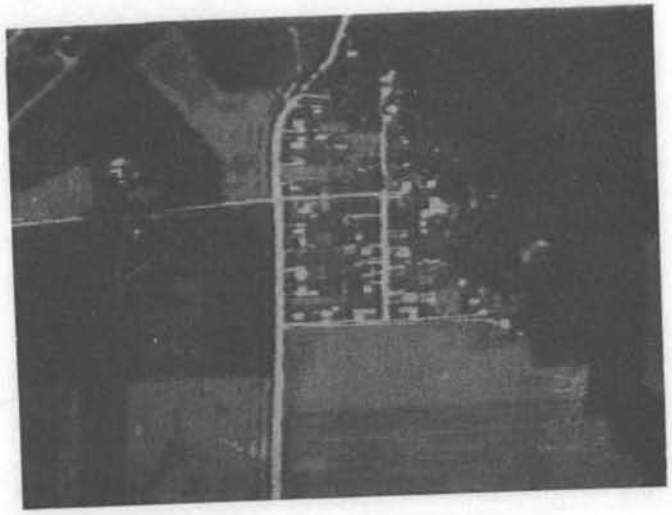
3.5. The Two Spatial Dimensions

In contrast to the definition of a spectral N-vector associated with each ground resolution point as described above, the two spatial dimensions of the multispectral source output suggest the concept of N-two dimensional images, one for each spectral channel. This idea is presented in Figure 3.10.

Correlations existing in the horizontal and vertical directions are compared in Figure 3.11 where the estimates of the normalized source autocorrelation functions ($R_x(\tau)$ and $R_y(\tau)$) are plotted. The $R_x(\tau)$ curve lies above the $R_y(\tau)$ curve for small τ. This reflects the fact that the aircraft



Original Data



R = 1.0

Error Image

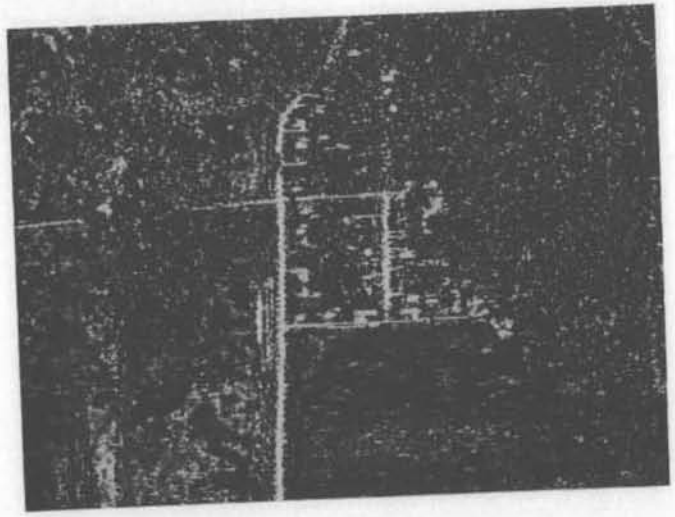


Figure 3.9. Error Image Between the Original and Reconstructed Channel 2 Image Using the K-L Encoder (1x1x6 Data Blocks).

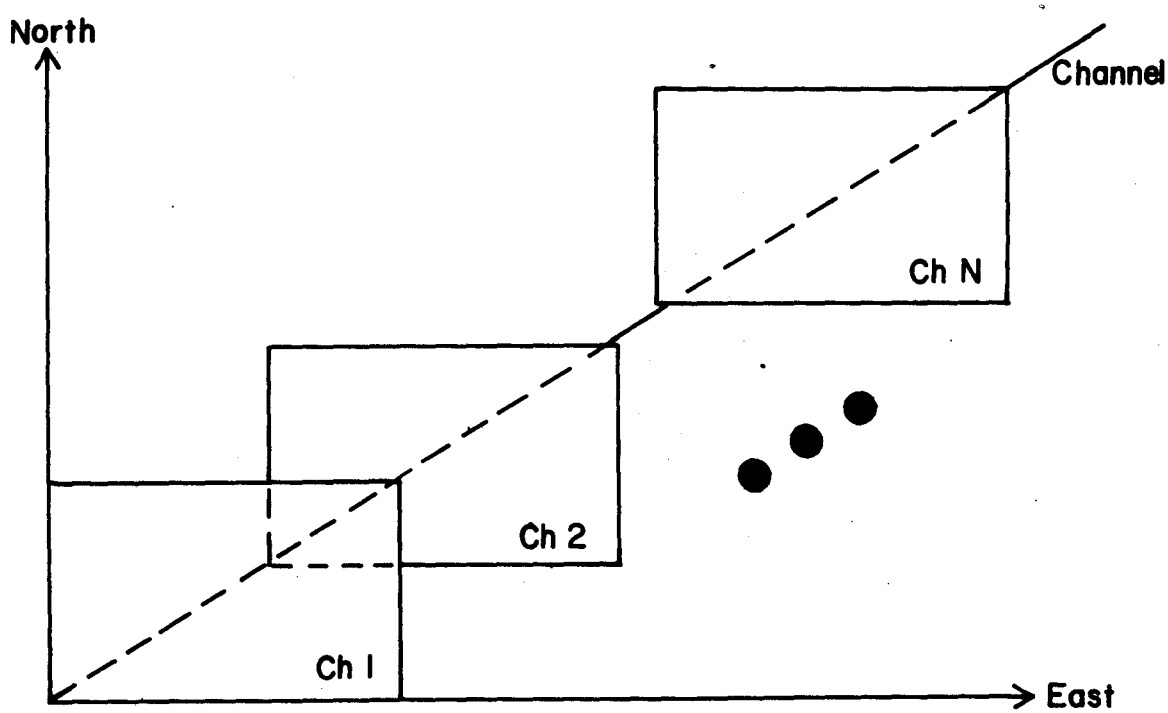


Figure 3.10. The N Two-Dimensional Spectral Images.

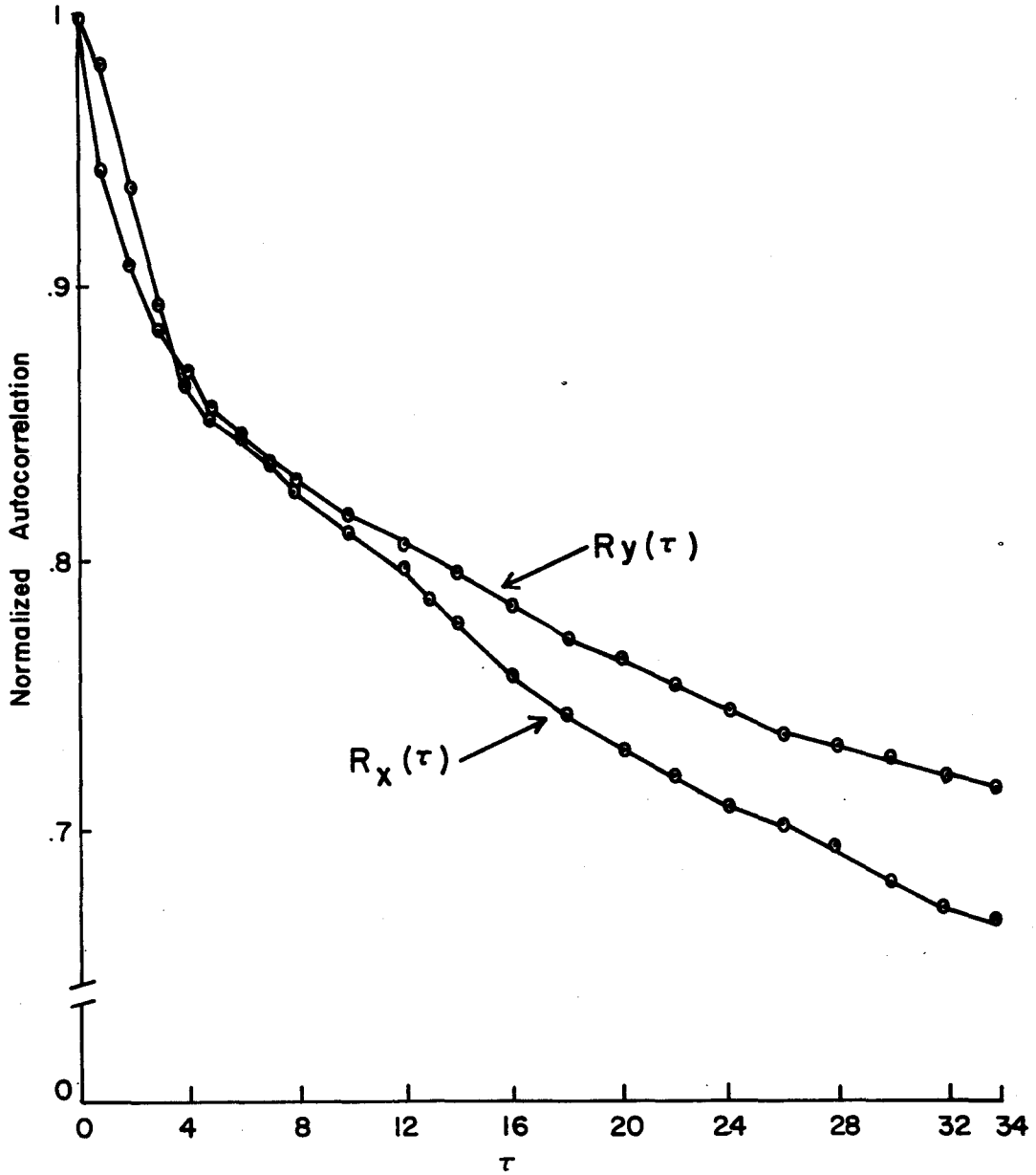


Figure 3.11. Normalized Inter-Line and Inter-Column Autocorrelation Functions.

scanner gathers data by sweeping out a single horizontal line at a time. The finite bandwidth of the scanner electronics tends to correlate adjacent points within a scan line more than vertically adjacent points.

The spatial transformations applied to the data set are, for purposes of comparison, one dimensional (horizontal) and two dimensional (horizontal and vertical). The data is taken in one and two dimensional blocks as shown in Figure 3.12. As discussed in section 2.6 the two dimensional blocks are re-arranged into column vectors. The two block-sizes considered are 1x64x1 (1 vertical point by 64 horizontal points by 1 spectral point) for the horizontal one dimensional spatial block, and 8x8x1 for the two dimensional spatial block. The NxN (N=64) covariance matrix is computed for both data blocks by averaging $\underline{X}\,\underline{X}^t$ over each of the six spectral images. The first column of the 8x8x1 data block covariance matrix is plotted in Figure 3.13.

Data rates achieved using the above two data blocks and the data compression system of Figure 1.4 are shown in Figures 3.14 and 3.15. The K-L, Fourier and Hadamard transformations are considered, with the gaussian source lower bound R(d) and the standard PCM encoder also included. The K-L encoder realizes the lowest data rates for both data blocks, while the Fourier and Hadamard encoders also achieve rates substantially below those for standard PCM. In both cases (1x64x1 and 8x8x1 data blocks) the Fourier encoder gives

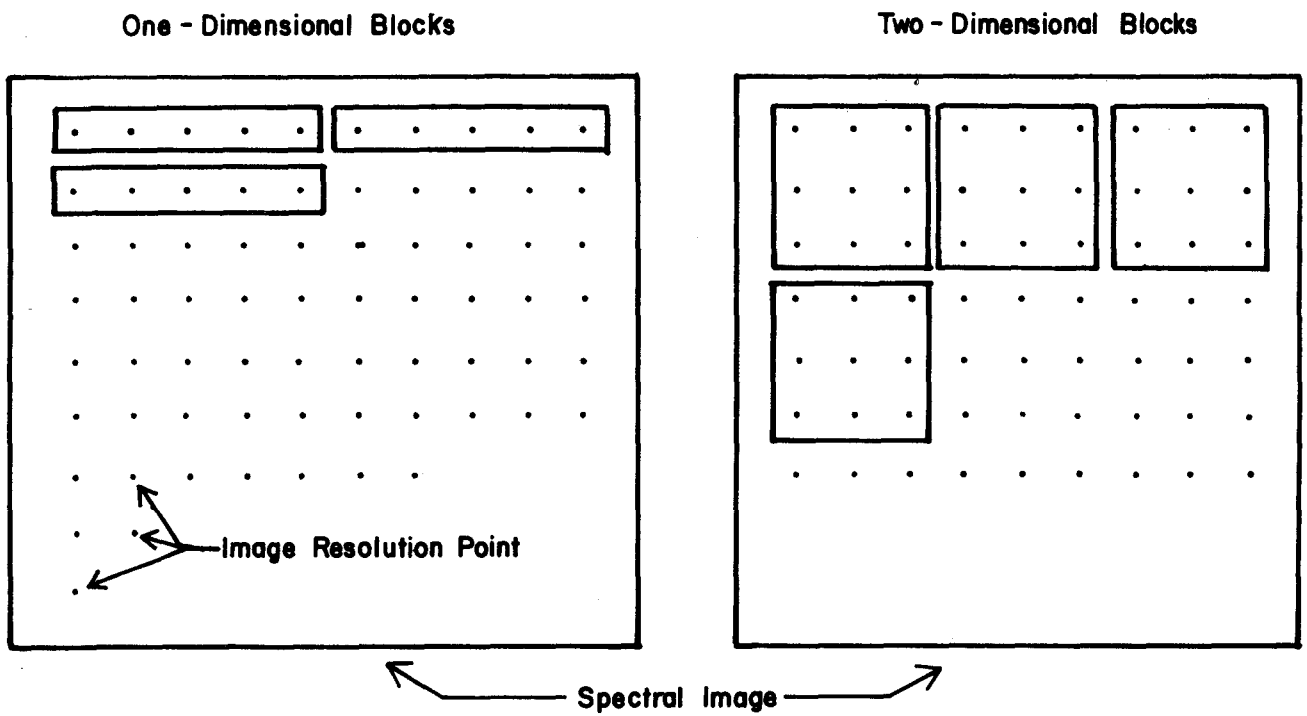


Figure 3.12. One and Two-Dimensional Spatial Data Blocks.

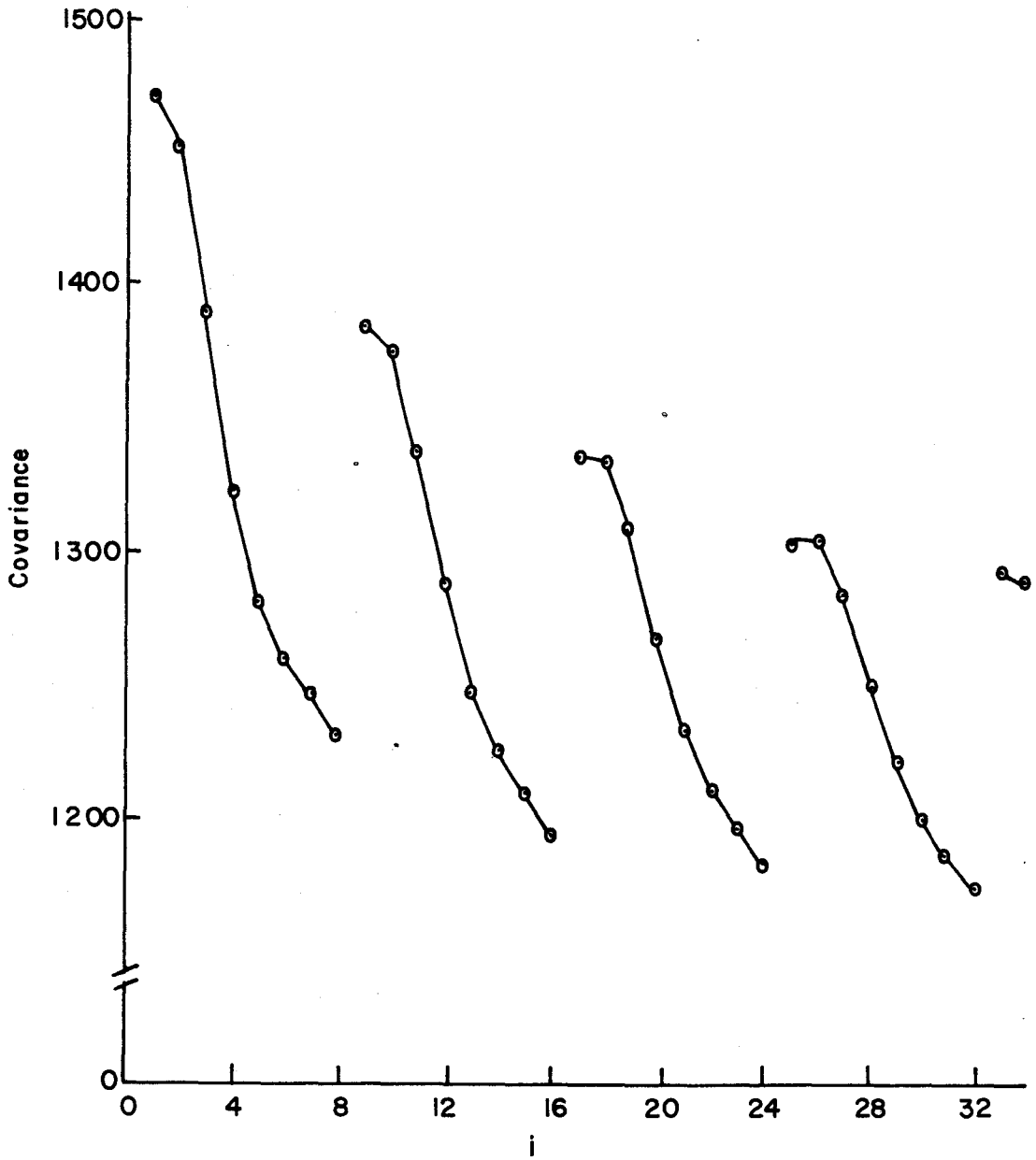


Figure 3.13. First Column of the 8x8x1 Data Block Covariance Matrix.

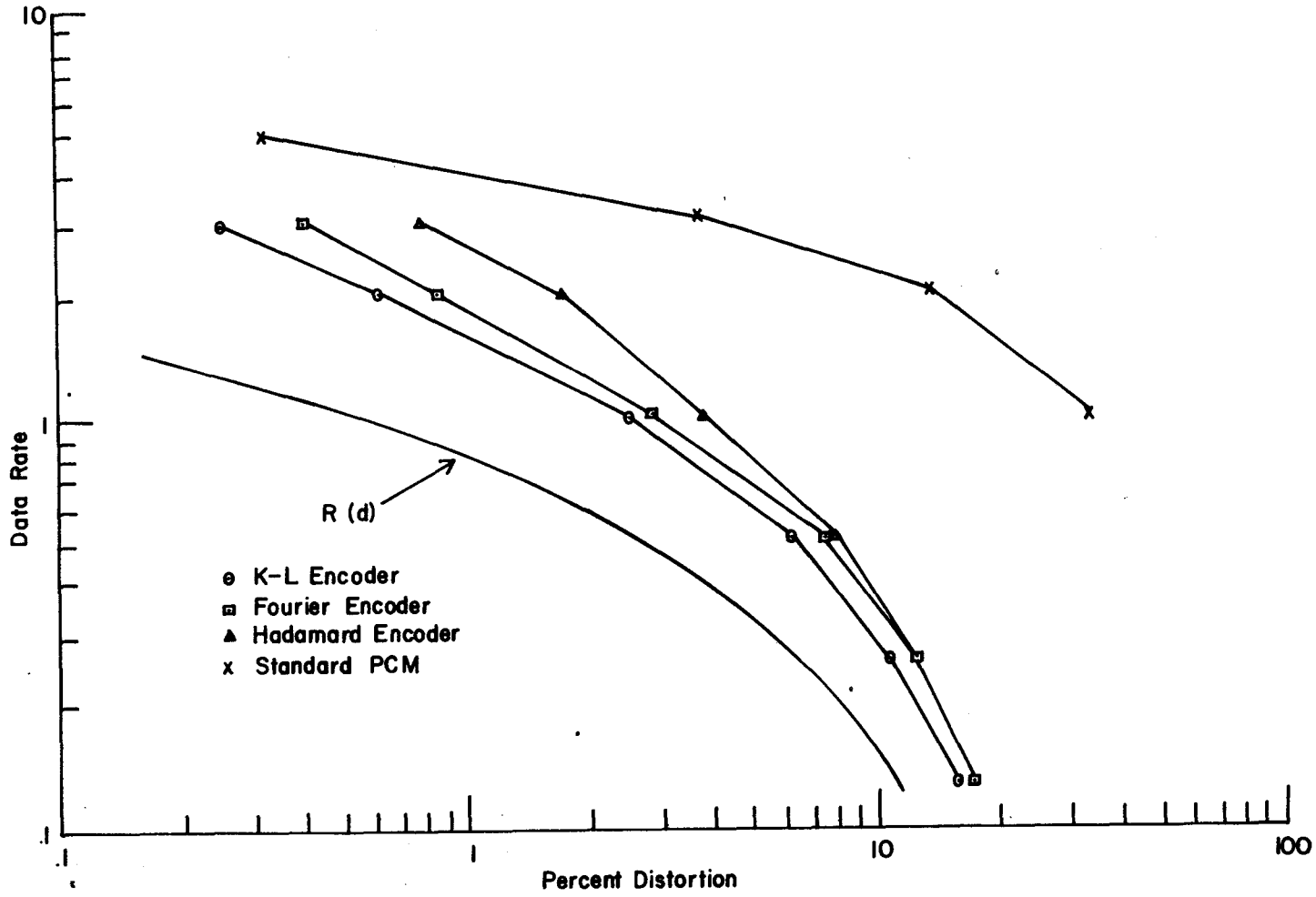


Figure 3.14. Data Rate Versus Percent Distortion Using the 1x64x1 Data Block.

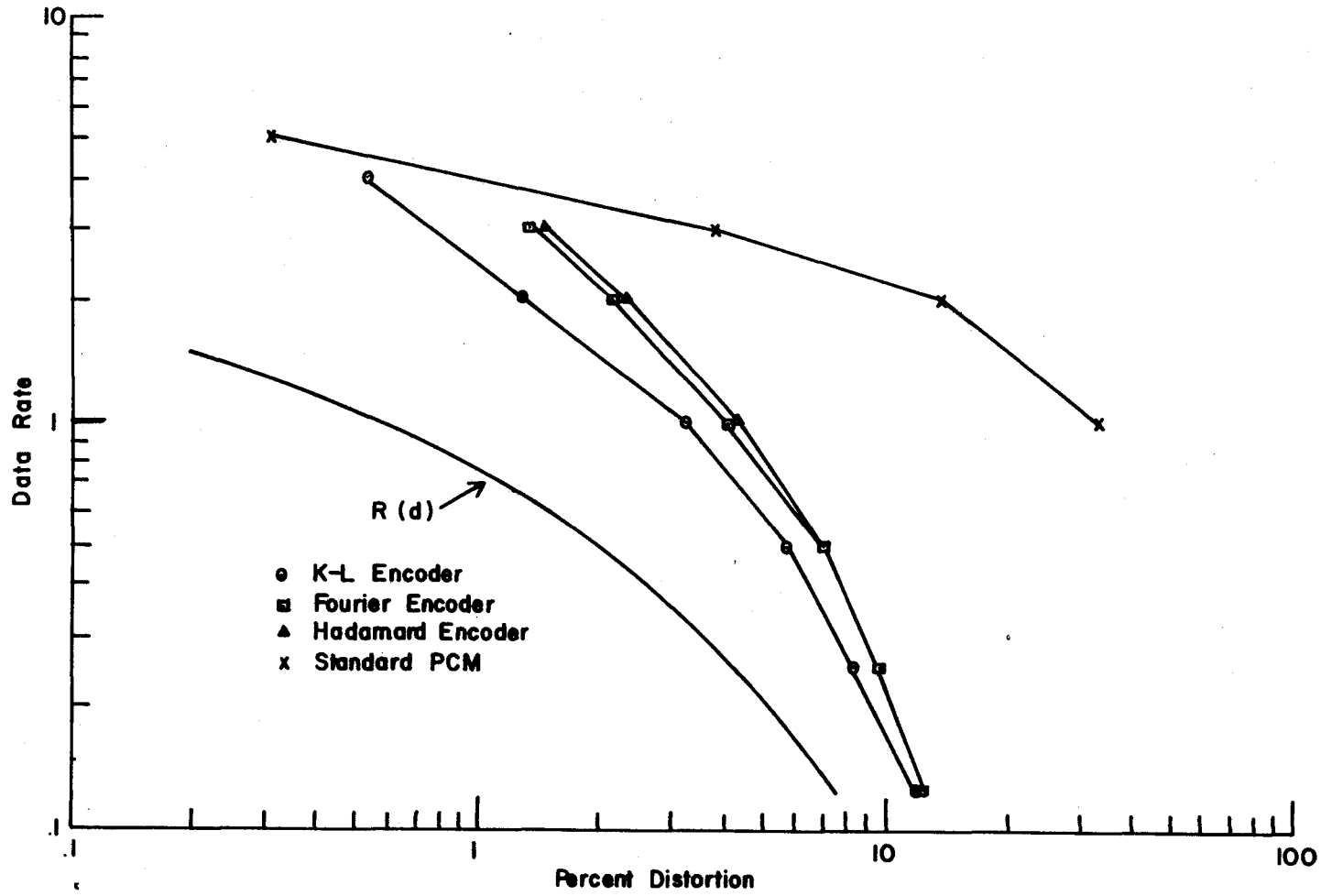
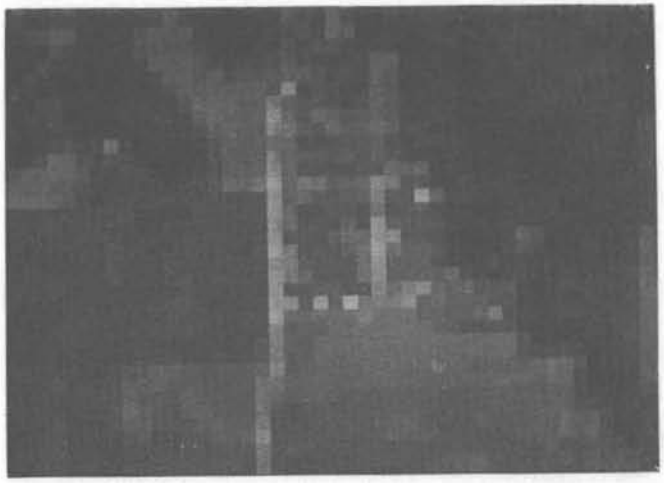


Figure 3.15. Data Rate Versus Percent Distortion Using the 8x8x1 Data Block.

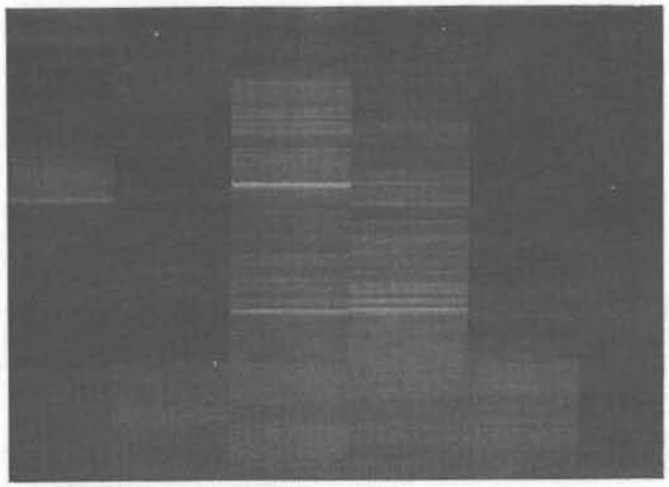
lower data rates than the Hadamard. The two dimensional block is more efficient with all three encoders for distortions near 7% and greater. Below 7% the one dimensional block is best. The number of transform coefficients retained and resulting m_i are listed in Appendix C for the K-L encoder and the one dimensional data block.

The relative sizes of the 8x8x1 and 1x64x1 data blocks in relation to the original spectral data set are shown in Figure 3.16. These images are the result of retaining only the first coefficient in the K-L encoder and quantizing that coefficient to 512 levels (m_i=9). This is not the manner in which the 0.141 bit/vector element data rate would be achieved.

Channel 2 images using the K-L encoder at rates of 2,1, and 0.5 bits/vector element for the 1x64x1 and 8x8x1 data blocks are presented in Figures 3.17 and 3.18 respectively. Figure 3.19 is the error image between the original channel 2 image and the reconstructed image at 1 bit/vector element. As with the spectral data blocks (1x1x6) most of the error is found near the roads and buildings. However in this case the error is due to spatial high frequencies (detail) as opposed to spectral high frequencies. In the low detail area the prime source of the error for the two dimensional encoder is the distortion near the edges of the 8x8x1 data blocks. (See Appendix A for a discussion of error within the data block.) Essentially no image degradation is



8x8x1

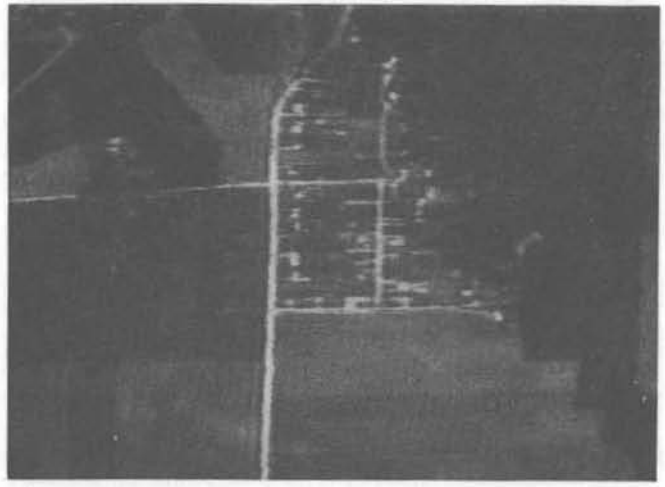


1x64x1

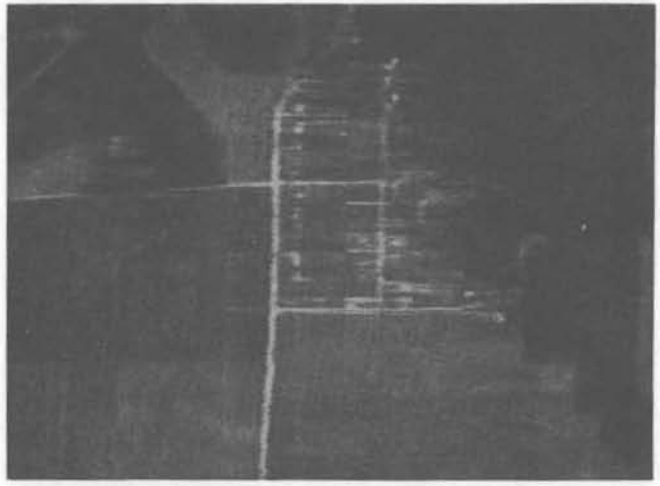
Original Image

Figure 3.16. Comparison of the One- and Two-Dimensional Data Blocks with the Original Spectral Image.

R = 2.0

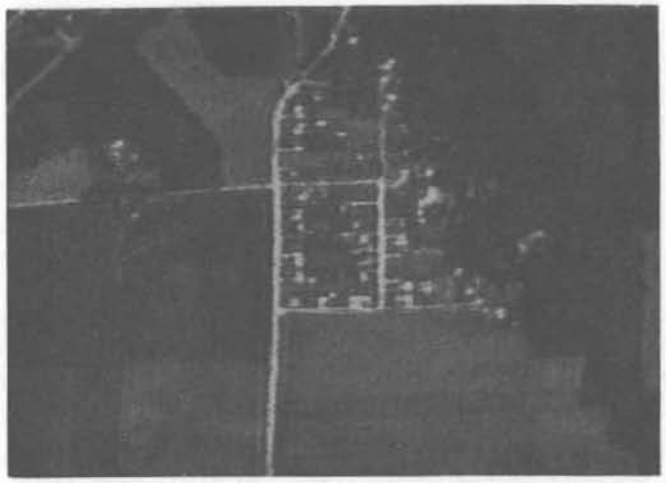


R = 1.0

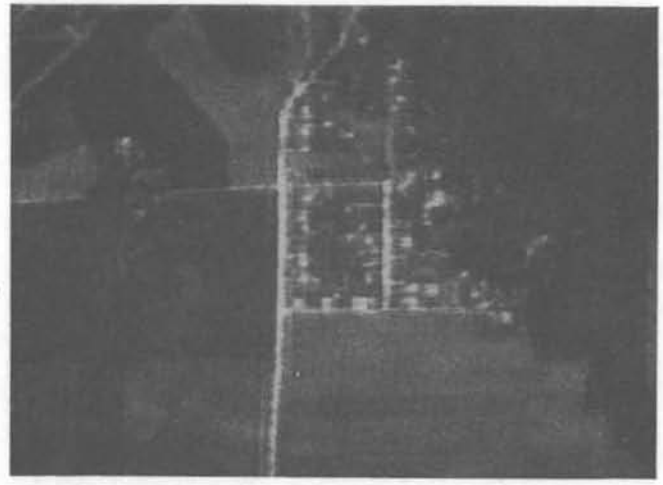


R = 0.5

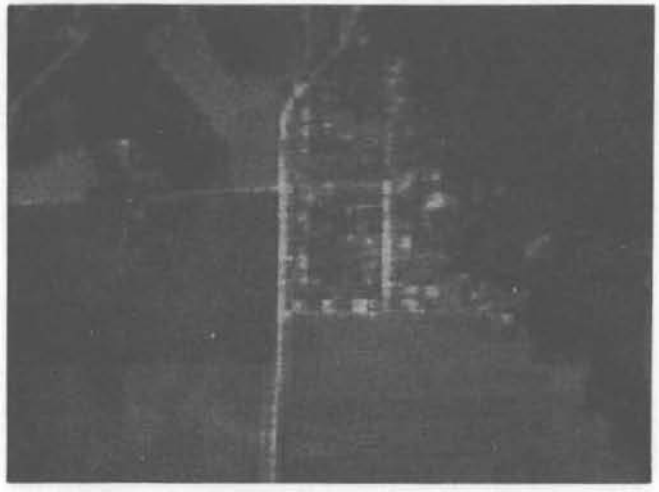
Figure 3.17. Reconstructed Channel 2 Image Using the K-L Encoder at Three Different Data Rates (1x64x1 Data Blocks).



R = 2.0

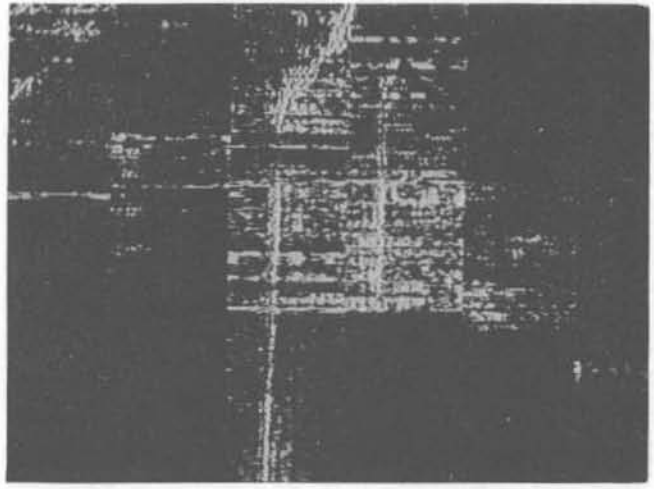


R = 1.0

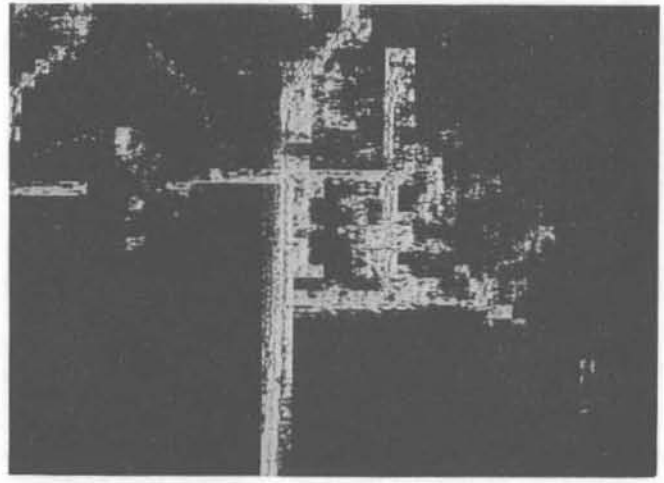


R = 0.5

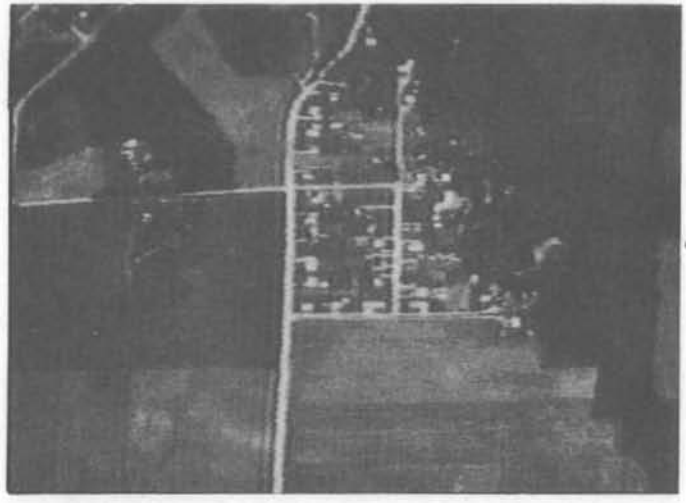
Figure 3.18. Reconstructed Channel 2 Image Using the K-L Encoder at Three Different Data Rates (8x8x1 Data Blocks).



1x64x1 Data Blocks



8x8x1 Data Blocks



Original Image

Figure 3.19. Error Images Between the Original and Reconstructed Channel 2 Image Using Two Different Data Blocks and the K-L Encoder.

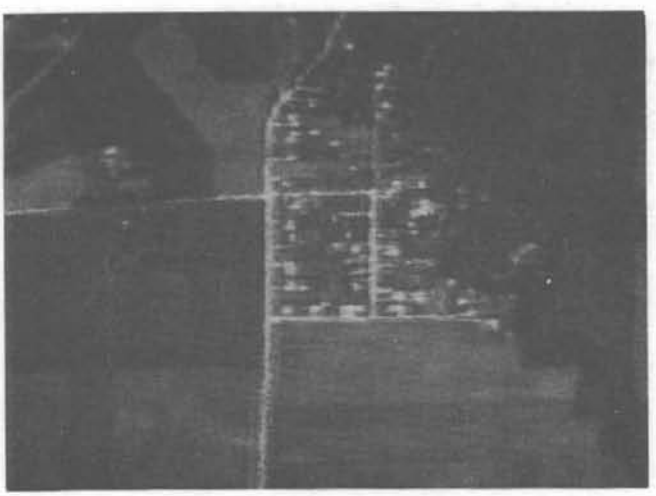
noticeable in the 2 bit/vector element image, while some distortion is evident in the 1 and 0.5 images for both data blocks. Channel 2 images using the Fourier and Hadamard encoders at 2 bit/vector element and 8x8x1 data blocks are shown in Figure 3.20.

The effects of data compression on percent classification accuracy are presented in Figure 3.21 for the K-L and standard PCM encoders using the 8x8x1 data blocks. The two curves represent the performance of the maximum likelihood classifier (described in section 3.2) over the reconstructed spectral images for several different data rates. The performance of the K-L encoder is far superior to that of the PCM encoder although both converge to 94% correct classification for large (> 4 bits/vector element) data rates. The apparent insensitivity of the K-L encoder to degradation in classification performance at low data rates is a consequence of the variance packing property discussed in section 2.4. At one bit/vector element for example, the PCM encoder allocates only 2 output levels to each element of $\underline{X}$. The K-L encoder however, distributes the 64 bits in the following manner

Table 3.5. Integer Bit Distribution Over the K-L Coefficients

i	1	2	3	4 ... 9	10 ... 19	20 ... 30	31 ... 64
m_i	7	4	4	3 ... 3	2 ... 2	1 ... 1	0 ... 0

Fourier Encoder



Hadamard Encoder

Figure 3.20. Reconstructed Channel 2 Image Using 8x8x1 Data Blocks with the Fourier and Hadamard Encoders at R = 2.0.

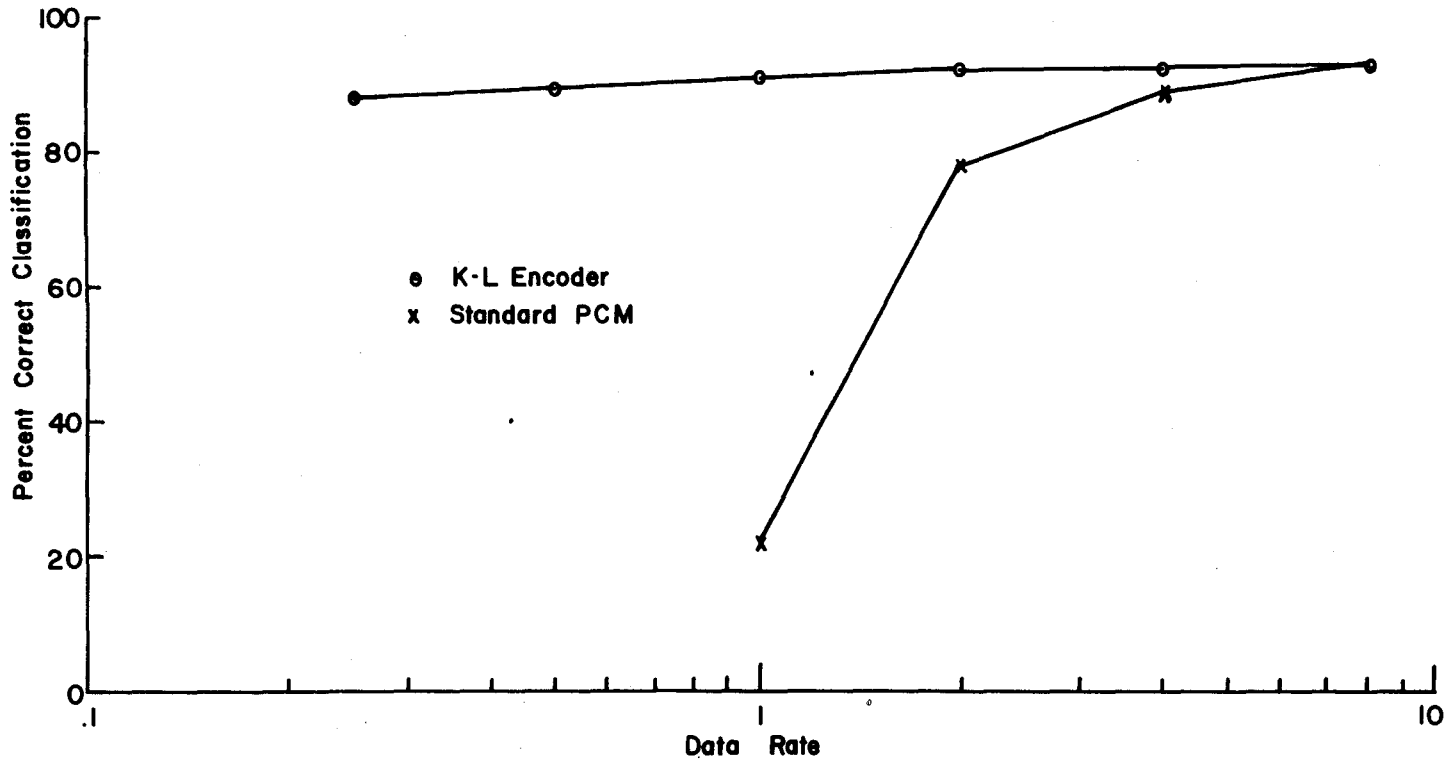


Figure 3.21. Percent Correct Classification Versus Data Rate Using 8x8x1 Data Blocks.

Thus the first quantized coefficient may take on one of 128 possible values, the second may be one of 16 values and so on. The increased number of possible values for y_i^* (and therefore $\hat{\underline{X}}$) allows greater class separation within the data.

3.6. The Three Dimensional Source

The multispectral scanner output is actually three dimensional - the two spatial dimensions and the spectral dimension. In order to take advantage of correlations existing in all three dimensions of the source a three dimensional data block is defined as shown in Figure 3.22. The results of sections 3.4 and 3.5 indicate that spatial correlations are higher than those found in the spectral dimension. For a given block size N, ($N=N_1 \cdot N_2 \cdot N_3$) it is therefore advantageous to include more spatial samples than spectral (i.e., $N_1, N_2 > N_3$). For this reason the three dimensional data block is chosen to be 8x8x2, 8 vertical samples by 8 horizontal sample by 2 spectral samples. As described in section 2.6 the data block is re-arranged into a column vector $\underline{X}$ having $N=8 \cdot 8 \cdot 2=128$ elements.

From the spectral correlation matrix of Figure 3.3 it is evident that some pair-wise combinations of channels are more highly correlated than others. Since the data block is 8x8x2 it is advantageous to arrange the channels into those pairs having the highest correlation. This is possible since

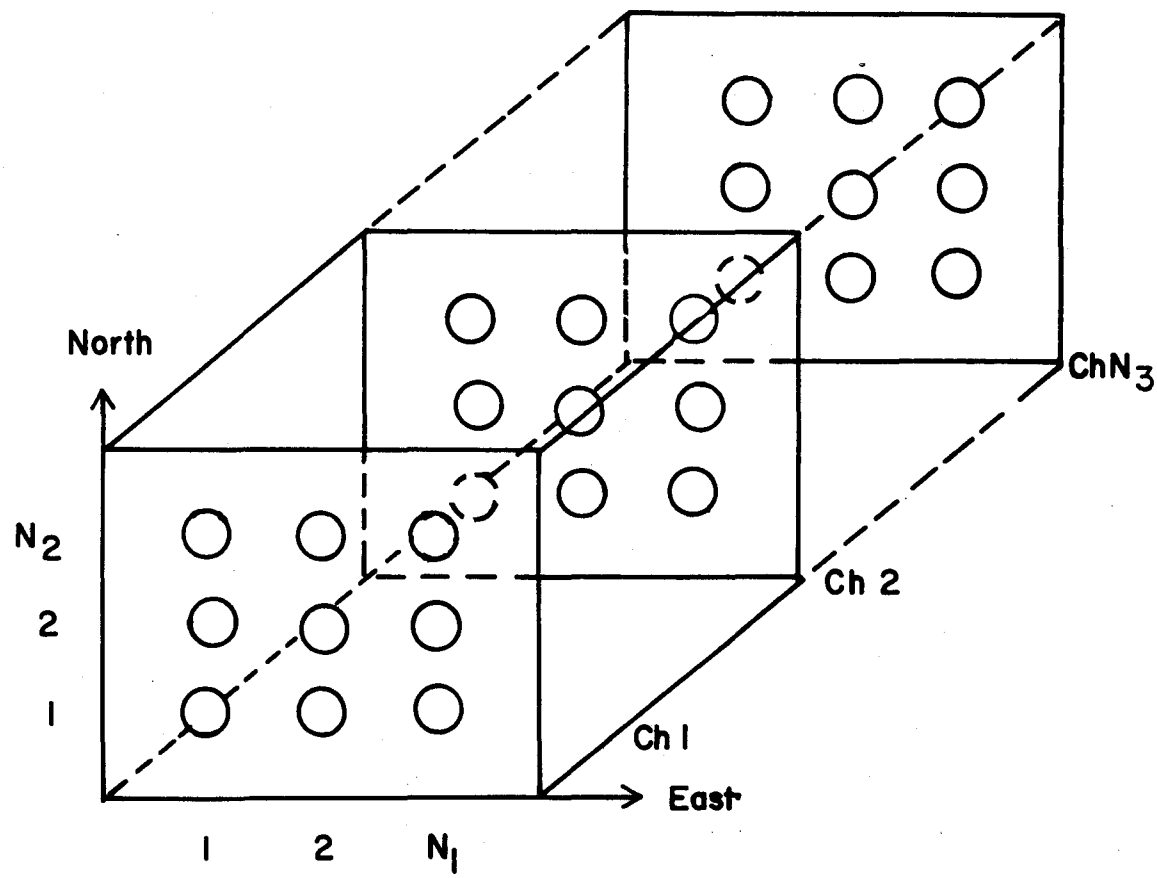


Figure 3.22. The Three-Dimensional Data Block.

all the channels of data are presented simultaneously at the output of the scanner, and the order in which they are selected is arbitrary. Based on the spectral correlation matrix the following channel sequence is selected

Table 3.6. Re-Ordered Spectral Channels

Channel	Wavelength Band (μm)
1	0.40 - 0.44
6	2.00 - 2.60
2	0.62 - 0.66
3	0.66 - 0.72
4	0.80 - 1.00
5	1.00 - 1.40

From the above sequence channels 1 and 6 are contained in the same data block, then 2 and 3, and finally 4 and 5. The first column of the resulting 128x128 source covariance matrix for the 8x8x2 data block is presented in Figure 3.23.

Data rates achieved with the three dimensional data blocks and the K-L, Fourier and Hadamard encoders are presented in Figure 3.24. The K-L encoder realizes the lowest data rate. The Fourier and Hadamard encoders are quite similar, although the Hadamard is slightly better over the distortion range considered. This may be due to the relatively large differences in mean values between the channels as indicated in section 3.2. All three encoders achieve rates lower than the PCM encoder for distortion near 1% and greater.

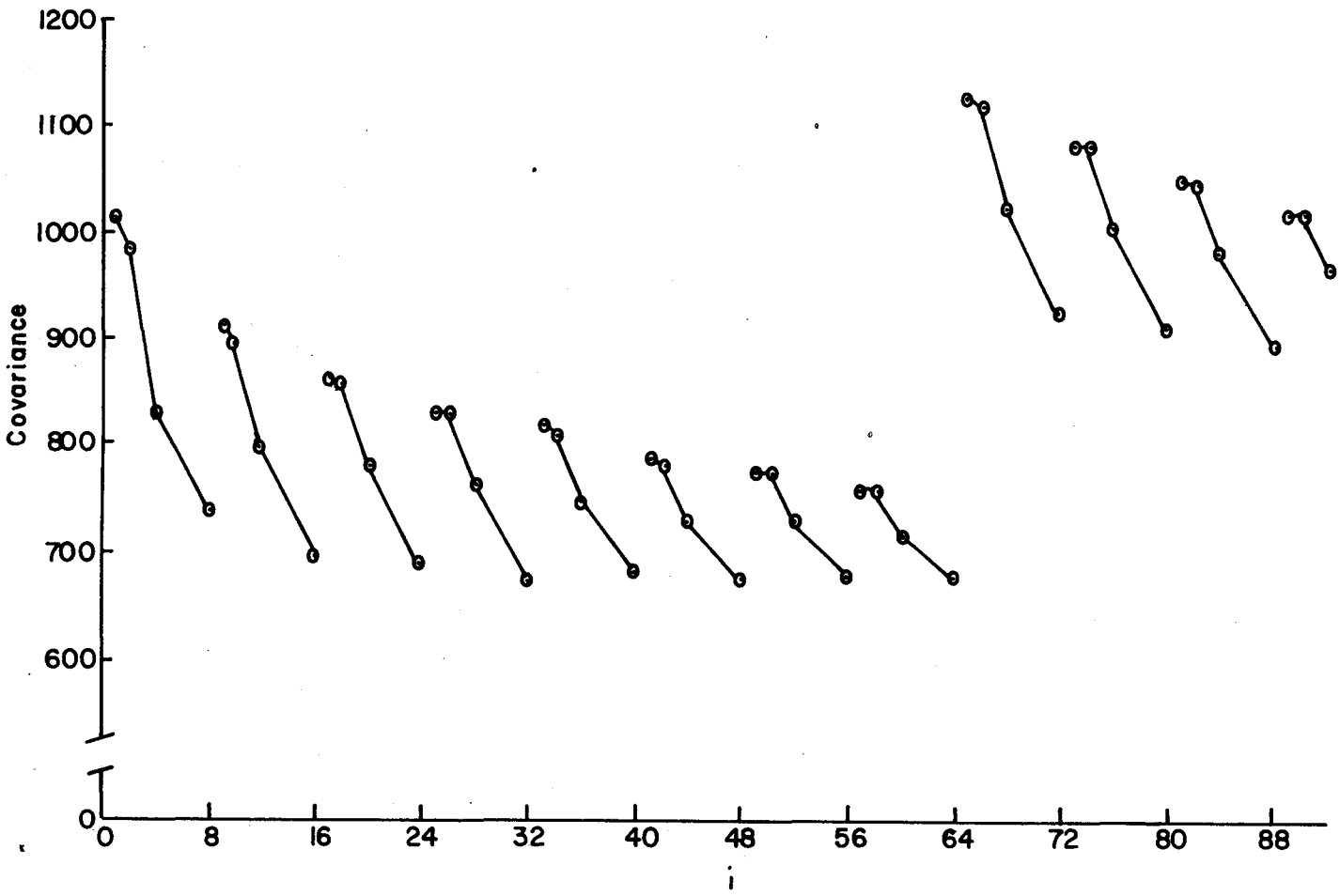


Figure 3.23. First Row of the 8x8x2 Data Block Covariance Matrix.

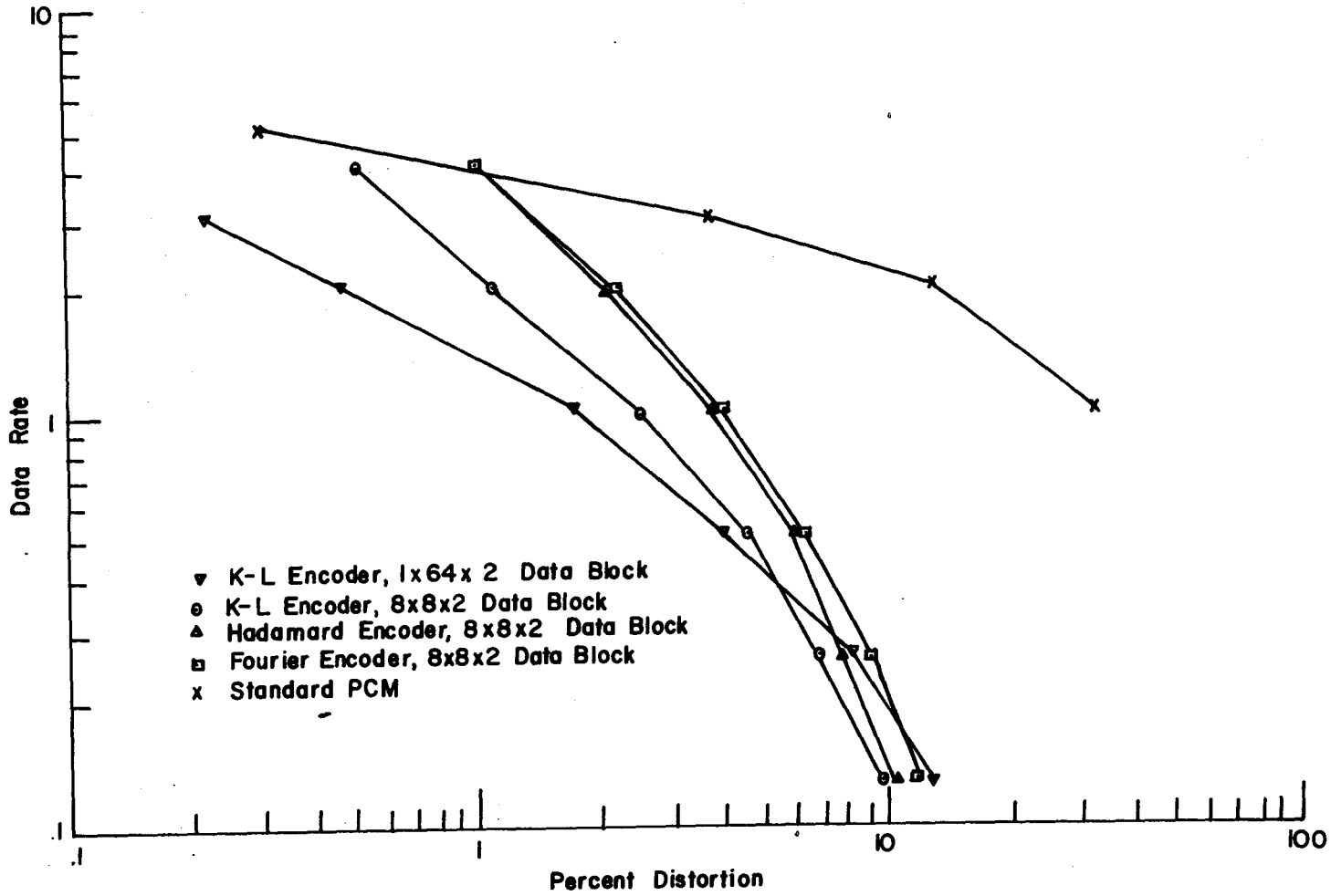


Figure 3.24. Data Rate Versus Percent Distortion Using the 8x8x2 and 1x64x2 Data Block.

Also included in Figure 3.24 is the rate versus distortion curve for the K-L encoder using a 1x64x2 data block. This data block is selected based on the performance of the 1x64x1 block over the 8x8x1 block as presented in section 3.5. As in the two dimensional case the 1x64x2 block achieves lower data rates than does the 8x8x2 block (d<5%).

3.7. Conclusions and Comparison of Results

The results of Chapter III indicate that correlations in all three dimensions of the multispectral source are sufficiently high to warrent the use of a transform coding scheme as outlined in Figure 1.4. Correlations are lowest in the spectral dimension and highest in the two spatial dimensions.

Data rates achieved with the several different data blocks considered (plus an additional 6x6x1 block) and the K-L encoder are compared in Figure 3.25. The 1x64x2 block is best for rates above 0.4 bits/vector element, while the 8x8x2 block is best for rates less than 0.4. The least efficient data block is the 1x1x6 block - the spectral channel vector.

The sensitivity of the K-L encoder to varying scene statistics is examined in Figure 3.26. The two middle curves represent data rates achieved with the 1x64x1 data block and the K-L encoder (1) optimized over the region of Figure 3.1 (these rates are identical to those for the K-L encoder in

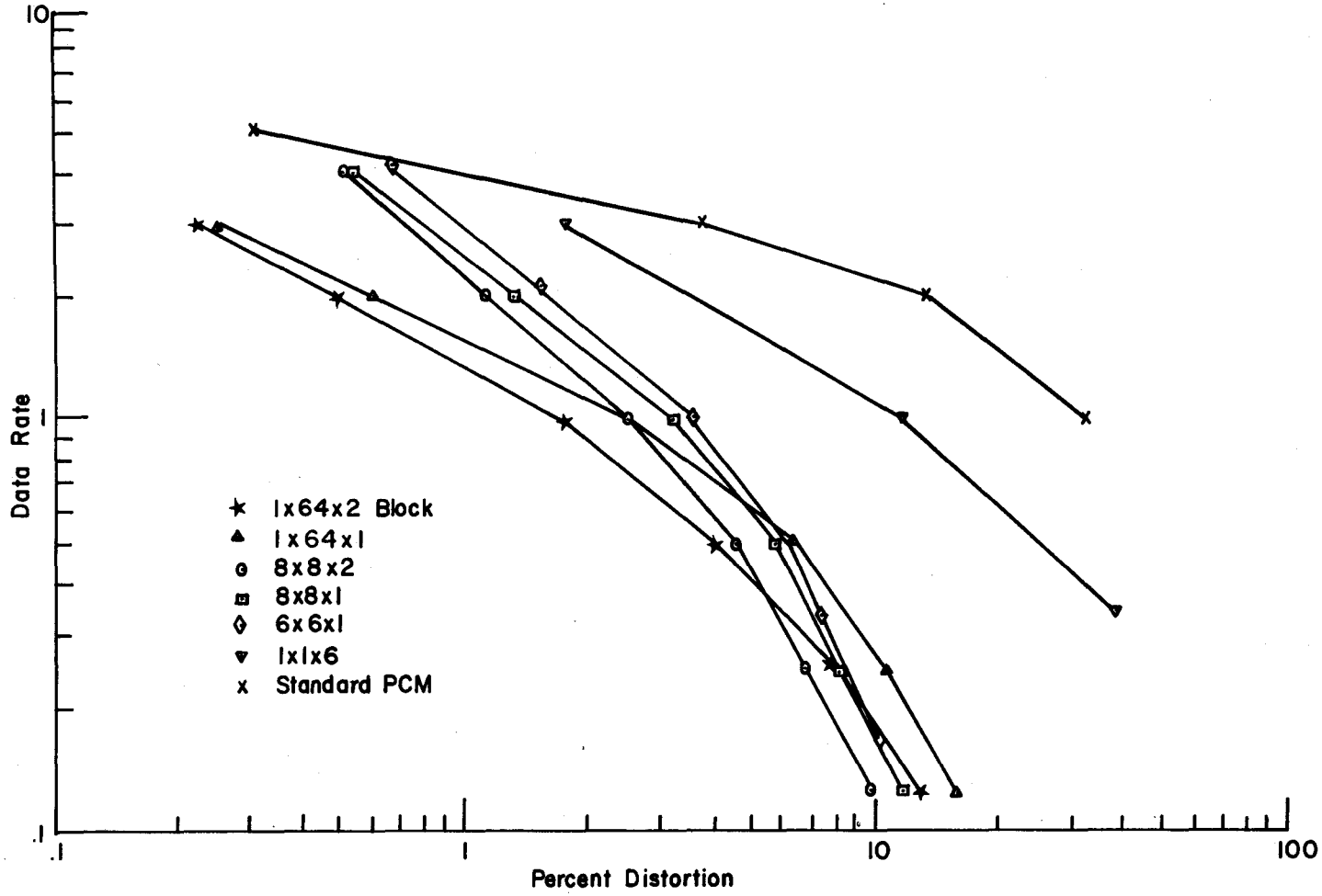


Figure 3.25. Data Rate Versus Percent Distortion Using the K-L Encoder and Various Data Blocks.

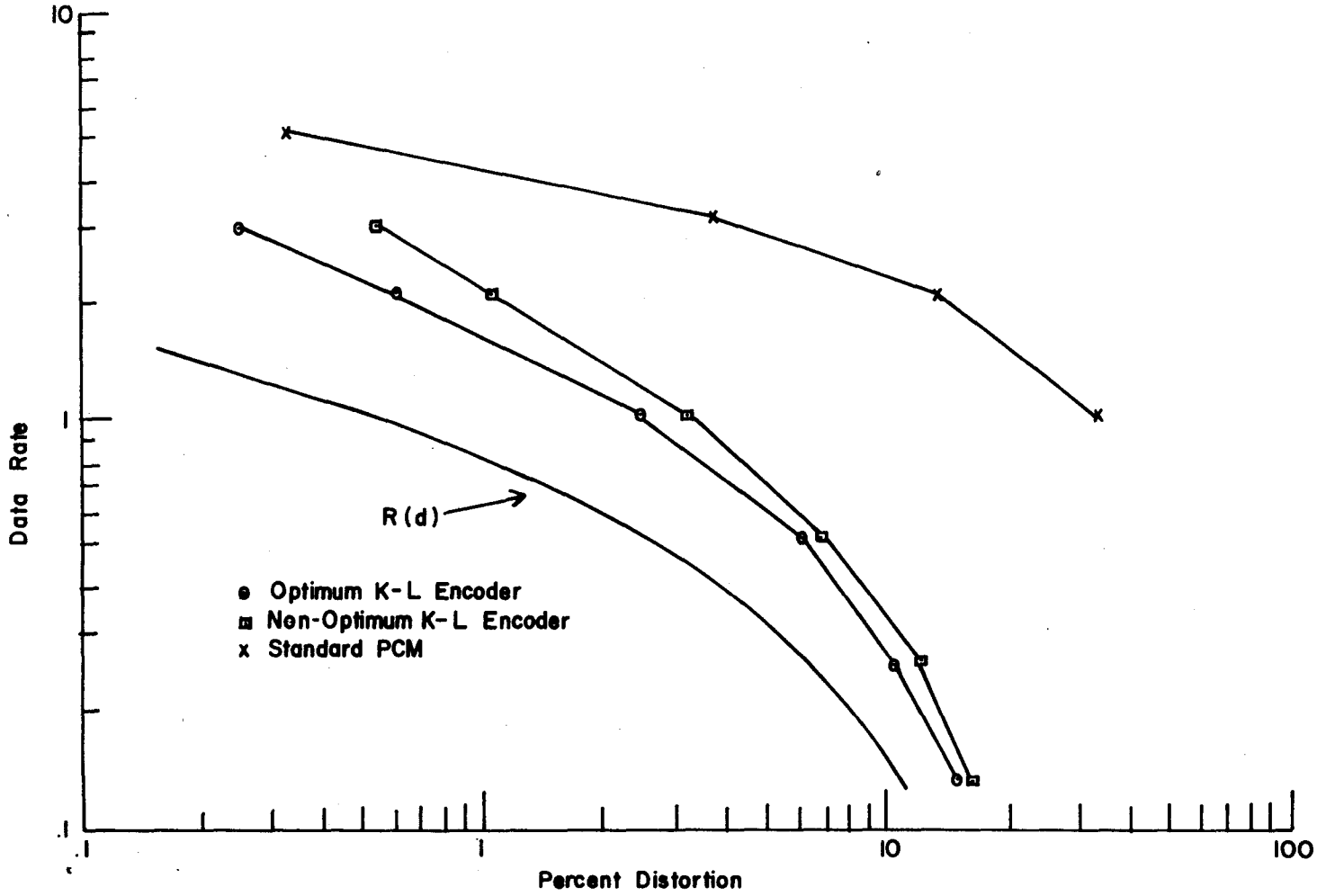


Figure 3.26. Data Rate Versus Percent Distortion Using the Optimum and Non-Optimum K-L Encoder with the 1x64x1 Data Block.

Figure 3.14), and (2) optimized over a much larger area (1 mile by 25 miles) of which the area in Figure 3.1 is a subset. The test area for both encoders is that of Figure 3.1. Although the non-optimum K-L encoder requires higher data rates than the optimum encoder, both are substantially below the PCM encoder.

CHAPTER IV

EXPERIMENTAL RESULTS PART II - SATELLITE DATA

4.1. Introduction and Description of the Source

In this chapter multispectral imagery gathered on the NASA Apollo 9 S065 experiment [49] over Imperial Valley, California (frame No. 3698), is subjected to the data compression system of Figure 1.4. The data is obtained from images taken by four 70mm Hasselblad cameras mounted in the Apollo 9 spacecraft window. Each camera has the particular film-filter combination as listed below

Table 4.1 The Four Film-Filter Combinations

Camera	Film	Filter	Wavelength Band (μm)
1	SO-180 Ektachrome Infrared	Photon 15	0.51-0.89
2	3400 Panatonic-X	Photon 58B	0.47-0.61
3	SO-246 B/W Infrared	Photon 89B	0.68-0.89
4	3400 Panatonic-X	Photon 25A	0.59-0.715

Images from the above camera-filter combinations were sampled with a scanning microdensitometer having a 25-micron aperture and sampling interval of 25 microns [50]. This sampling rate and the spacecraft altitude of approximately 115 miles gives each sample a theoretical ground resolution

of 200 ft. at nadir and a sample spacing of 200 ft. (actual ground resolution is 350 ft. at nadir). The data used in this study are the quantized (256 levels) samples from the last three wavelength bands listed above. To be useful in multispectral analysis the three images must be properly overlaid, i.e., each image resolution point must be in geometrical coincidence in all three channels. This problem is investigated in [51] and the results applied to the above three images. The resulting data is then a set of three geometrically coincident digital images, each representing the scene reflectance in the 0.47-0.61, 0.59-0.715, and the 0.68-0.89 μm bands. The three images are shown in Figure 4.1.

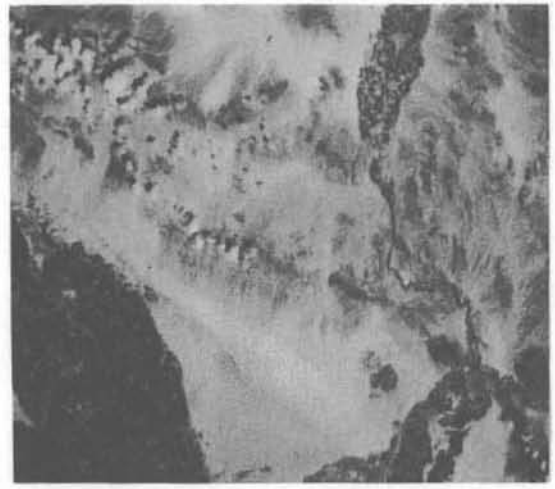
4.2. The Three Test Regions

Three quite different types of terrain are represented in the images of Figure 4.1 - vegetated areas, mountainous areas, and desert. It is felt that these three categories approximate the various types of data that an earth resources satellite might encounter while orbiting the earth. The areas chosen to represent the three categories are outlined in the channel 1 image shown in Figure 4.2. Each region is designated in the following manner

Table 4.2. The Three Test Regions

Region A - Vegetation
Region B - Mountain
Region C - Desert

Channel 1



Channel 2

Channel 3

Figure 4.1. The Three Original Spectral Images.

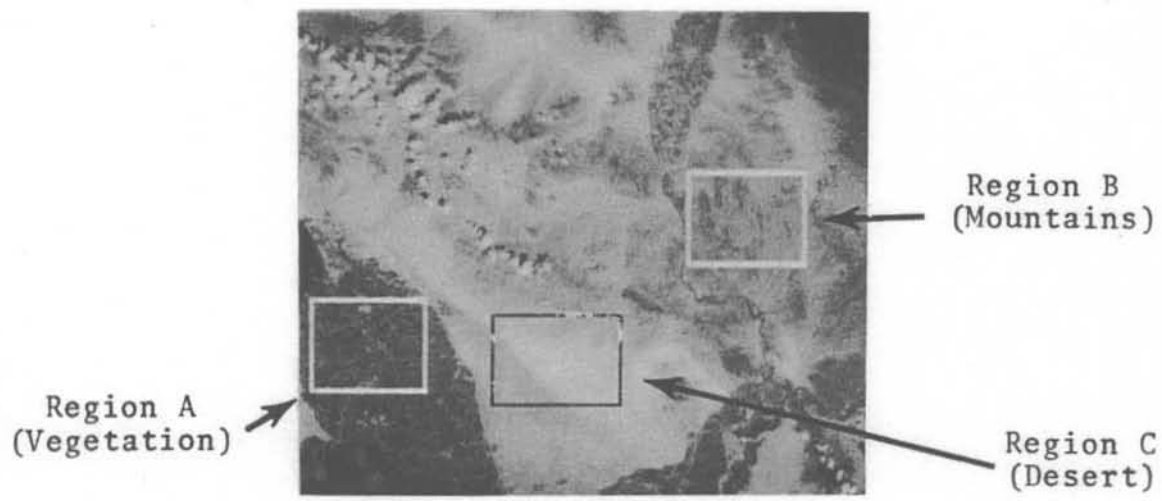


Figure 4.2. The Three Test Regions.

and each contains the same number of data points (384x300 per channel). The three regions are shown at maximum resolution in Figure 4.3.

4.3. Statistical Characteristics of the Three Regions

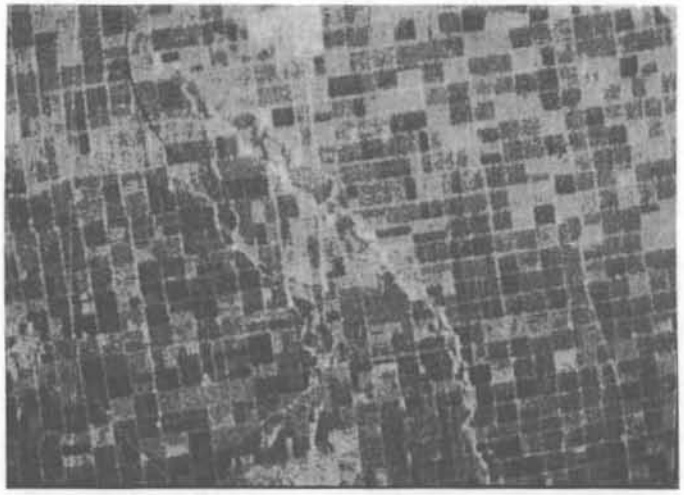
The mean and variance of Region A, B, and C (averaged over all three channels) are listed below

Table 4.3. Region Statistics

	Mean	Variance	Standard Deviation
A	85	839	29.0
B	94	375	19.4
C	134	240	15.5

The vegetated region A has the largest variance followed by the mountainous region B and the desert region C. The desert area has by far the highest average response while the mountainous and vegetated region have similar mean values. Histograms of the data from each region is presented in Figure 4.4.

The relative variances and spatial correlations existing in Regions A, B, and C are evident in Figure 4.5 where the first row of the source covariance matrix resulting from 8x8x1 data blocks is presented. Spatial correlations are evidently highest in Region C and lowest in the vegetated region A. This is in agreement with a subjective evaluation of the three images in Figures 4.2 and 4.3.



Region A

Region B

Region C

Figure 4.3. The Three Data Regions at Maximum Resolution.

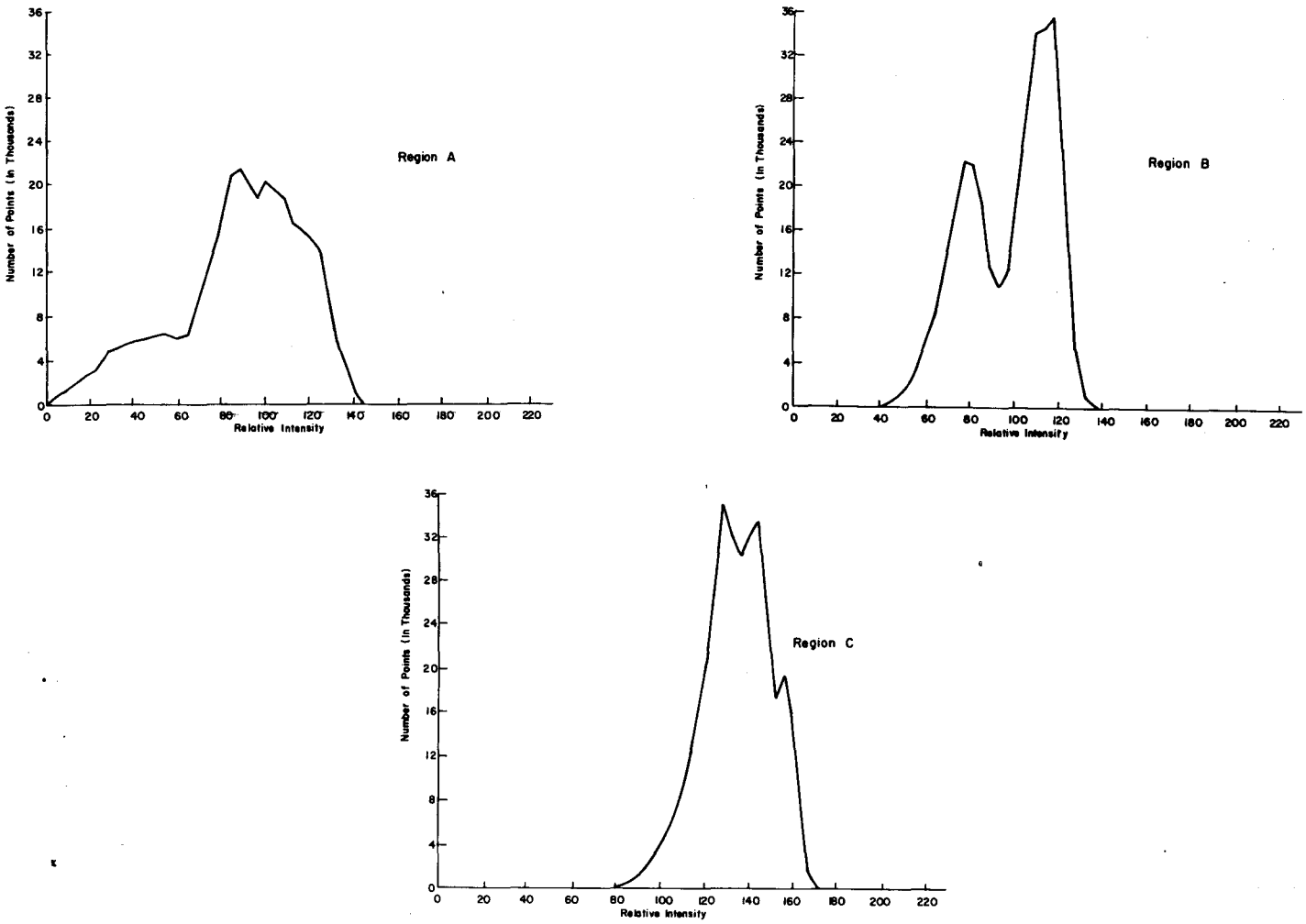


Figure 4.4. Histograms of the Three Test Regions.

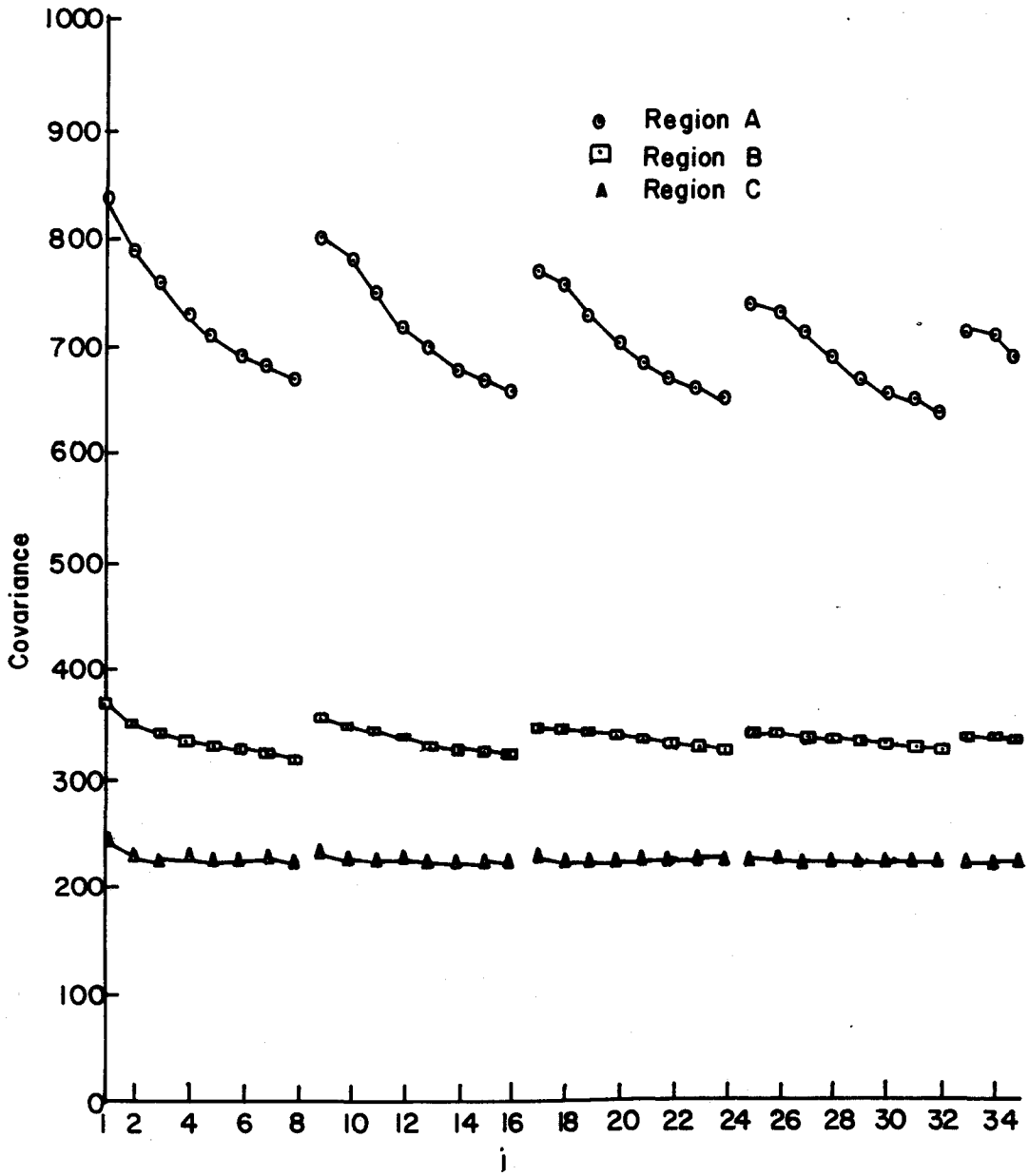


Figure 4.5. First Column of the 8x8x1 Data Block Covariance Matrix for Regions A, B, and C.

4.4. One and Two Dimensional Encoding

The data rates achieved using one and two dimensional spatial data blocks (see Figure 3.12) and the K-L, Fourier, and Hadamard encoders over Regions A, B, and C are presented in this section. The number of transform coefficients y_i retained and the resulting bit distribution over the coefficient are determined using the results of section 2.9. The two data blocks considered are the 1x64x1 and the 8x8x1 blocks.

4.4.1 Region A

The data rates achieved with the K-L, Fourier and Hadamard encoders over region A are presented in Figure 4.6. (The relative size of the 8x8x1 data blocks to the areas encoded is the same as that for the aircraft scanner data shown in Figure 3.16. However, one data block of the satellite imagery certainly includes more actual ground area than does the same block of aircraft scanner data.)

The K-L encoder achieves lower data rates than either the Fourier or the Hadamard encoder, while all three are substantially better than single sample PCM. As with the aircraft scanner data the rate distortion function for the gaussian source is included as a measure of the relative effectiveness of the various encoders. It is, as mentioned in Chapter III, an upper bound on the actual rate distortion function for the multispectral source. The Fourier and Hadamard encoders realize similar rates, although the Hadamard

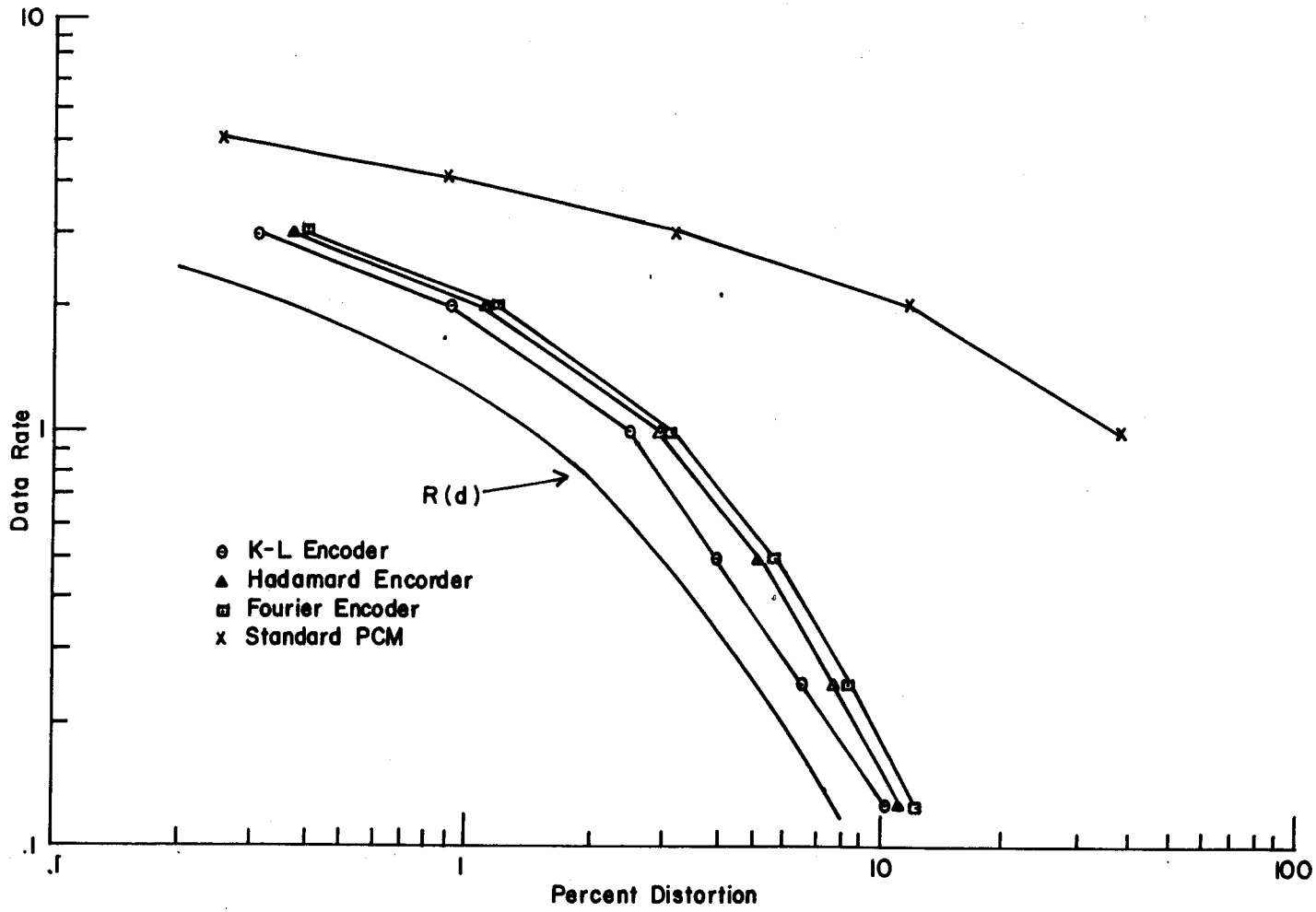


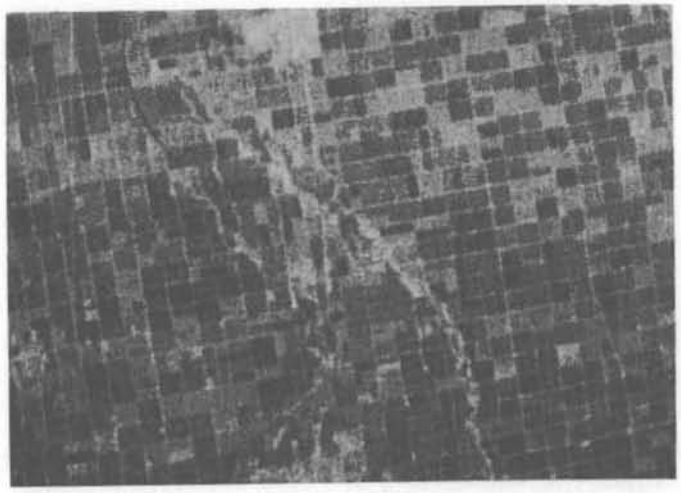
Figure 4.6. Data Rate Versus Percent Distortion Using the 8x8x1 Data Block Over Region A.

is slightly better. This may be due to the uniform nature of the fields within region A. There is little variation within a given field and a step-like variation between fields. This type of structure is quite similar to the Hadamard basis vectors as opposed to the more smooth variations present in the Fourier basis vectors.

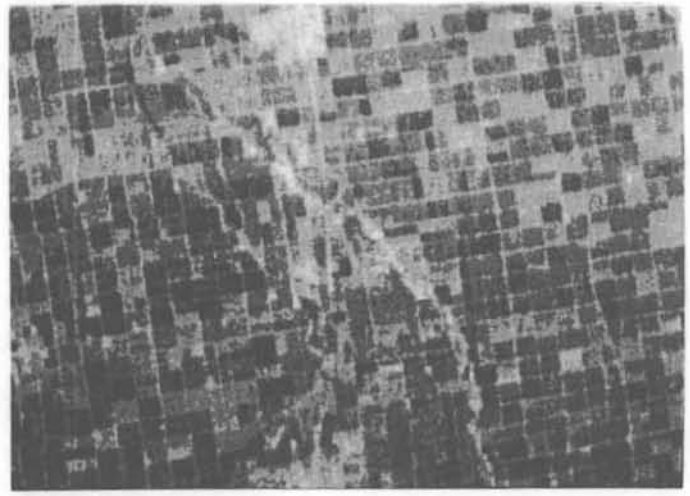
In the 1% distortion region the three encoders are capable of an apporximate 2:1 data rate reduction over standard PCM. Reconstructed channel 1 images at rates of 2, 1, and 0.5 bits/vector element using the K-L encoder are presented in Figure 4.7. Little or no distortion is evident at the 1 and 2 bit rates, while the 0.5 bit image contains detectable distortions.

The error image for the K-L encoder at 1 bit/vector element is also presented in Figure 4.8. Black represents no error and white represents a squared error of 255 or more. The area of large error in the upper-left corner of the image is due to the reconstruction error encountered over what is apparently a scratch in the original photograph. The results of the Fourier and Hadamard encoders using 8x8x1 data blocks at 1 bit/vector element are presented in Figure 4.9. Both are subjectively reasonable reconstructions of the original channel 2 image.

The effects of block structure on data rates is shown in Figure 4.10 where the rates achieved in region A using the K-L encoder are presented for the one dimensional (1x64x1)



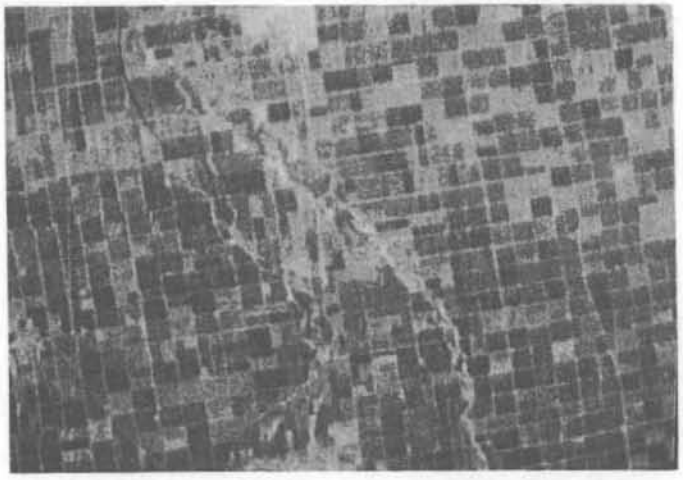
R = 2.0



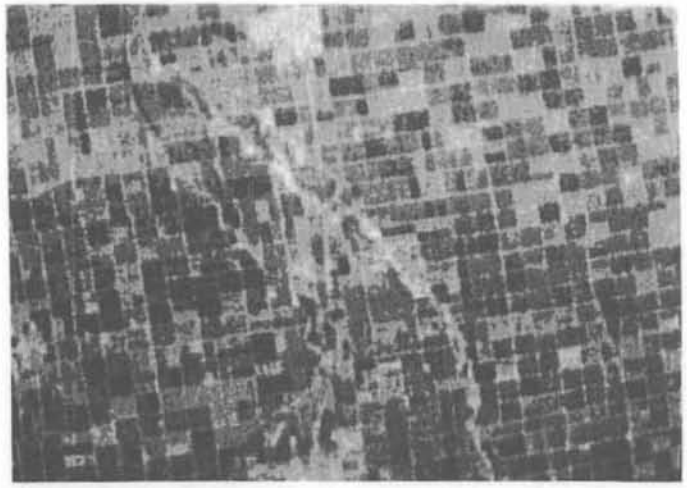
R = 1.0

R = 0.5

Figure 4.7. Reconstructed Channel 1 Image Using the K-L Encoder at Three Different Data Rates (8x8x1 Data Blocks).



Original Image



R = 1.0

Error Image

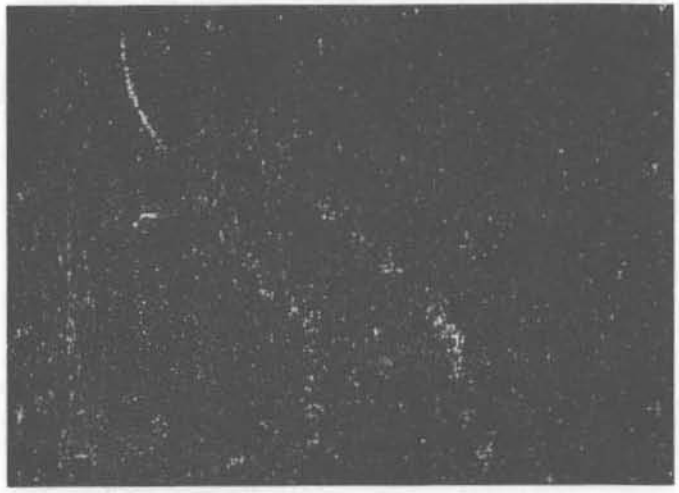


Figure 4.8. Error Image Between the Original and Reconstructed Channel 1 Image Using the K-L Encoder (8x8x1 Data Blocks).

Fourier Encoder

Hadamard Encoder

Figure 4.9. Reconstructed Channel 1 Image Using 8x8x1 Data Blocks with the Fourier and Hadamard Encoders at R = 1.0.

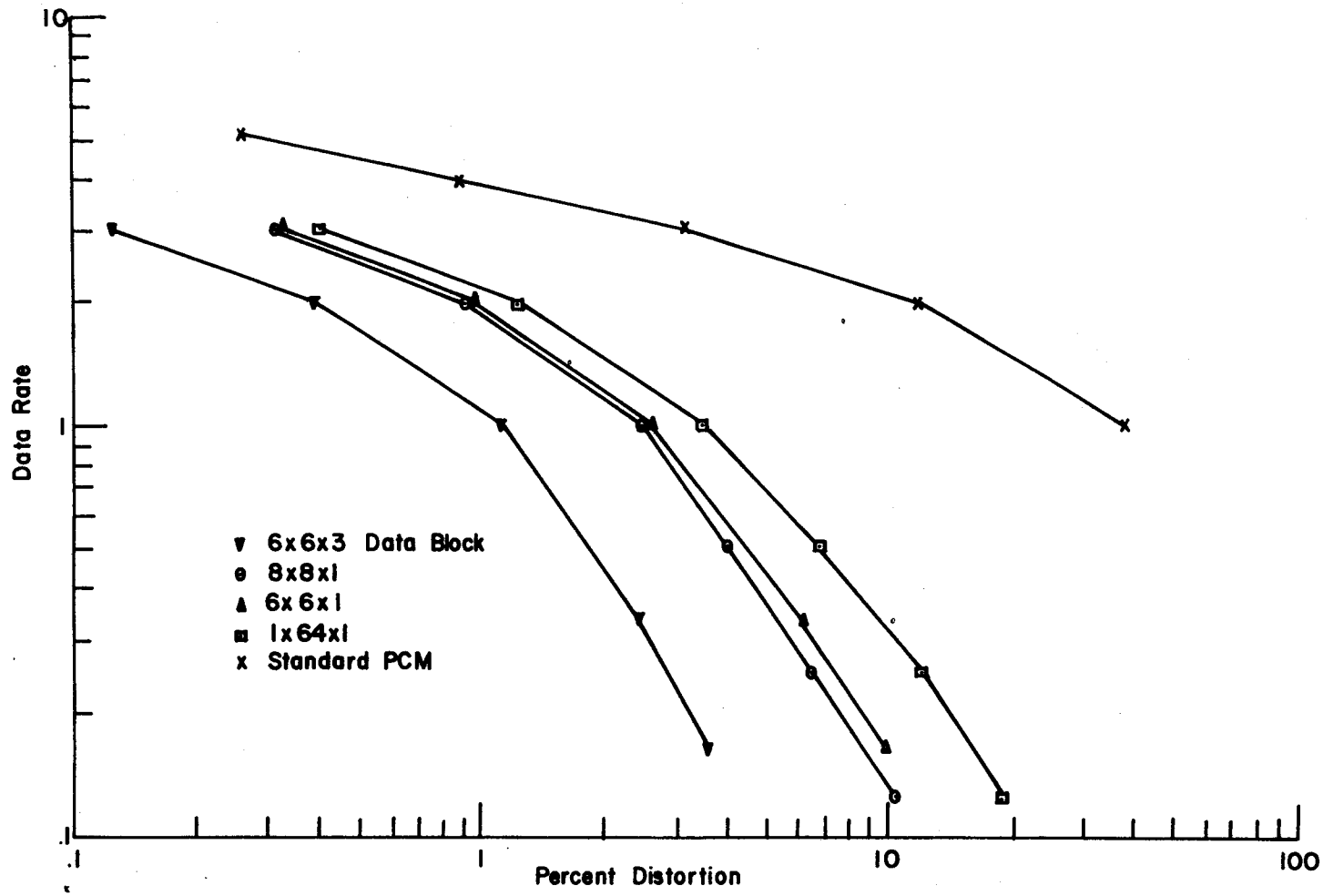


Figure 4.10. Data Rate Versus Percent Distortion Using the K-L Encoder and Various Data Blocks Over Region A.

and two dimensional (8x8x1) data blocks (the three dimensional block is also included and is discussed in section 4.6). The two dimensional block is uniformly better than the 1x64x1 block over the distortion range considered. This is in contrast to the results obtained with the aircraft scanner data where, as described in section 3.7, the one dimensional 1x64x1 block is better (d>2.5%) than the 8x8x1 block.

4.4.2 Region B

Data rates achieved using the K-L, Fourier and Hadamard encoders with 8x8x1 data blocks over the mountainous region B are presented in Figure 4.11. The K-L encoder realizes the lowest data rate while the Fourier and Hadamard are practically identical in their performance. This is evidently due to the low variance and high correlations in region B. The first few transform coefficients contain essentially all the average source energy E_s for both encoders. All three encoders achieve data rates substantially below the PCM encoder. In the 1% distortion range the K-L, Fourier and Hadamard encoders require approximately 2 bits/vector element while standard PCM requires 4 bits, giving the transform encoders a 2:1 rate reduction. A reconstructed channel 1 image using the K-L encoder and 8x8x1 data blocks at 1 bit/vector element is shown in Figure 4.12. Very little distortion is evident.

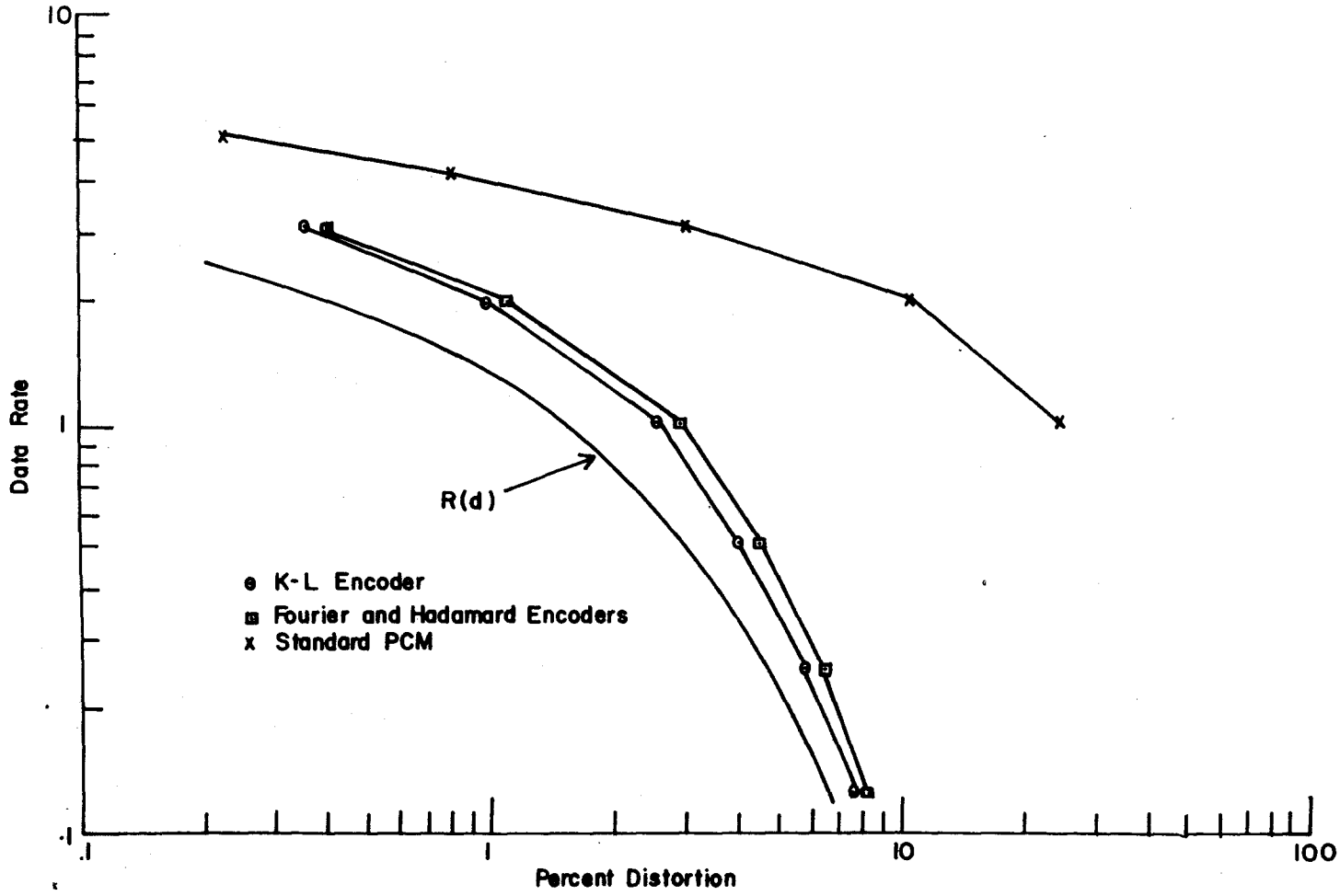


Figure 4.11. Data Rate Versus Percent Distortion Using the 8x8x1 Data Blocks Over Region B.

Region B (R = 1.0)

Region C (R = 1.0)

Figure 4.12. Reconstructed Channel 1 Image Using the 8x8x1 Data Block and the K-L Encoder Over Regions B and C.

4.4.3 Region C

Data rates achieved with the 8x8x1 data blocks and the K-L, Fourier and Hadamard encoders over the desert region C are presented in Figure 4.13. All three encoders are approximately indentical in their performance. The data in region C is so highly correlated (see Figures 4.2 and 4.5) that the first transform coefficient of each encoder contains essentially all the source energy E_s.

Table 4.4. Percent of Total Variance Contained in Each Transform Coefficient

	Coefficient					
	1	2	3	. . .	63	64
K-L	92.1	0.54	0.43	. . .	.07	.07
Fourier	92.1	0.41	0.36	. . .	.07	.07
Hadamard	92.1	0.46	0.39	. . .	.07	.07

The first Fourier and Hadamard basis vectors are identical, and both closely resemble the first K-L basis vector (for this particular data set). Thus the encoders use, in a sense, the same basis vectors, and the resulting data rates could be expected to be similar.

The transform encoders do realize rates substantially below the PCM encoder. As in regions A and B, the improvement is by a factor of approximately 2 in the 1% distortion region. A reconstructed channel 1 image using the K-L

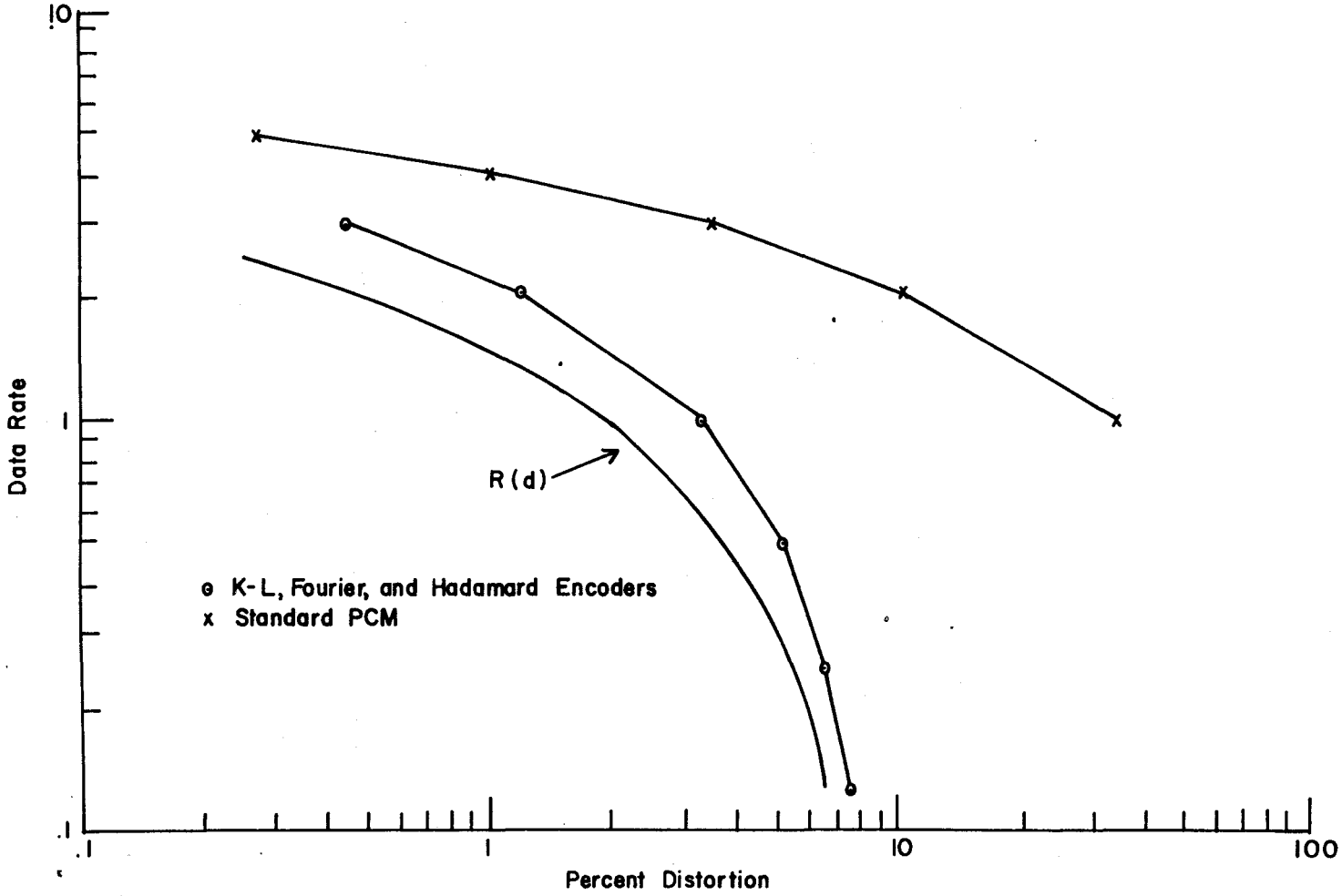


Figure 4.13. Data Rate Versus Percent Distortion Using the 8x8x1 Data Block Over Region C.

encoder and 8x8x1 data blocks at 1 bit/vector element is presented in Figure 4.12.

4.5. Comparison of Data Rates Over the Three Regions

Data rates achieved over regions A, B and C may be meaningfully compared on a total error basis, but not on a (non-adjusted) percent error basis. For example a 1% distortion in region A represent more total error than a 1% distortion in region C. This is because percent error as defined in 2.2 is a function of the data variance in the region considered. Total error is not so defined, but is difficult to interpret. In order to compare the data rates in the three regions on a percent error basis the "adjusted" percent error is defined. It is the total error encountered in a given region, divided by the average source energy E_s from the vegetated region A. These results are presented in Figure 4.14.

The encoder used in each region is the "optimum" encoder for that region (i.e., the transformation matrix $\underset{\sim}{T}$, the number of transform coefficients retained and the resulting m_i, are all based on the region covariance matrix). The most difficult region to encode is the vegetated region A, while the desert region is the least difficult. This is also true for the PCM encoder as evidenced by the top three curves.

In a practical, non-adaptive situation (see [53] for a discussion of an adaptive two dimensional encoder; also [55])

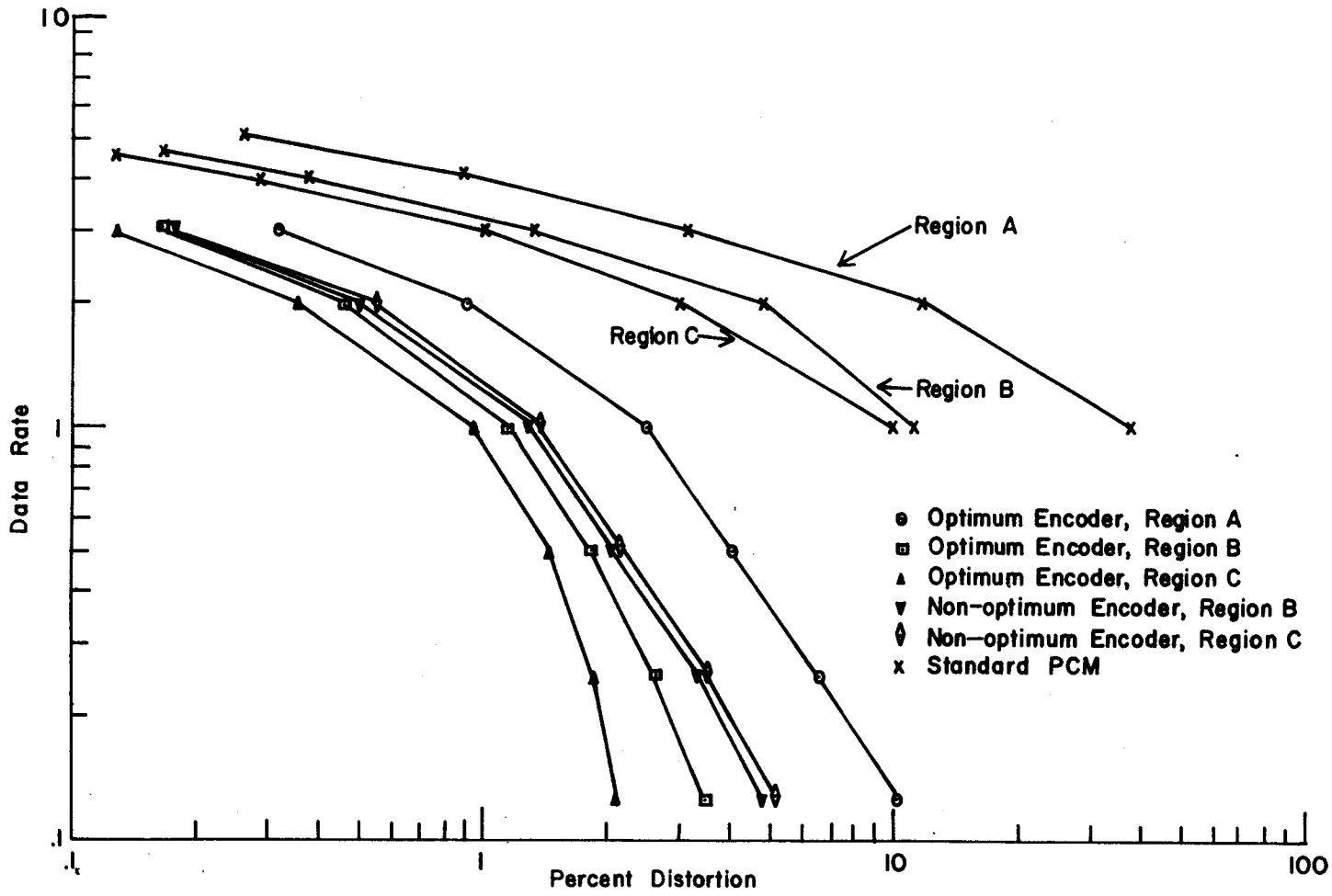


Figure 4.14. Data Rate Versus Percent Distortion Using the Optimum and Non-Optimum K-L Encoder (8x8x1 Data Block).

one encoder would be used for all three regions. The sensitivity of the data rates to encoders optimized over one region and applied to another is demonstrated in Figure 4.14. The K-L encoder is optimized over the more difficult region A and used to encode regions B and C.

The data rates achieved over region B and C are approximately equal although the region C curve is slightly higher. Both are above the data rates obtainable using the optimum encoder for each region, but both are still below the "optimized" PCM encoders. These results indicate that it might be reasonable to design an encoder to efficiently handle a particular type of terrain designated most important, and still perform well over statistically different areas. As an example, the reconstructed channel 1 images from regions B and C using the region A optimized K-L encoder at 1.0 bits/vector element are presented in Figure 4.15.

4.6. Three Dimensional Encoding

The three dimensional data block includes correlations existing in all three dimensions of the multispectral source. The structure of the three dimensional block is described in section 3.6 and Figure 3.22.

Data rates achieved using the K-L encoder over region A with 6x6x3 blocks (6 lines x 6 columns x 3 channels) are presented in Figure 4.10. Also included are the rates obtained with the 6x6x1 block. It is evident that the three

Region B

Region C

Figure 4.15. Reconstructed Channel 1 Image Using the Non-Optimum K-L Encoder Over Regions B and C (8x8x1 Data Block).

dimensional data block is superior to the other one and two dimensional blocks for the distortion range considered. At 1% distortion an approximate 5:1 rate reduction over standard PCM is indicated.

4.7. Principal Component Images

Although the satellite data analyzed in this study has only three spectral channels of information, it is of interest to examine the three principal component images. The images are constructed in the manner described in section 3.2. The data block (1x1x3) is the three element spectral vector.

The source covariance matrix is

$$\underset{\sim}{C} = \begin{matrix} 585.2 & & \\ 594.5 & 1034.5 & \\ 339.7 & 348.7 & 740.0 \end{matrix} \qquad 4.1$$

yielding the following three eigenvalues and K-L transformation matrix $\underset{\sim}{T}$

$$\lambda_1 = 1689.2, \quad \lambda_2 = 512.5, \quad \lambda_3 = 158.0 \qquad 4.2$$

$$\underset{\sim}{T} = \begin{bmatrix} 0.527 & 0.719 & 0.453 \\ -0.082 & -0.487 & 0.870 \\ 0.846 & -0.495 & -0.197 \end{bmatrix} \qquad 4.3$$

The weights given to the original three channels are given by the columns of $\underset{\sim}{T}$. The resulting three principal component images are shown in Figure 4.16.

Component 1

Component 2

Component 3

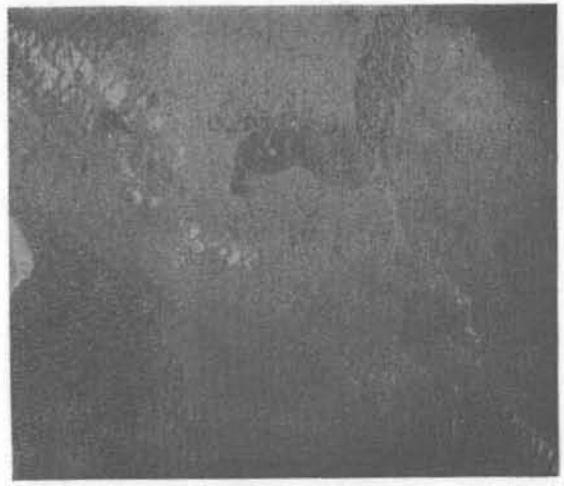


Figure 4.16. The Three Principal Component Images.

CHAPTER V

DISCUSSION OF THEORETICAL AND EXPERIMENTAL RESULTS

In this chapter the theoretical results of Chapter II and the experimental results of Chapters III and IV are discussed. The transform encoder is evaluated based on the two error criteria defined in section 2.1.

5.1. Theoretical Results

An approximate solution to the minimization of the total encoder system error over the number of retained transform coefficients and corresponding bit distribution for a fixed data rate and block size is presented and successfully applied to both the Markov source and to the multispectral data. The optimum number of retained coefficients is shown to be a function of the variances of the transform coefficients, as is the resulting bit distribution.

The K-dimensional linear transformation is shown to be representable by a single equivalent matrix multiplication of the re-ordered source output tensor. The resulting equivalent matrix for the Hadamard transformation is again a Hadamard matrix. This is not the case for the Fourier transformation, although the equivalent matrix is again orthogonal.

The selection of an optimum data block size under a maximum volume constraint is considered for the continuous source. The solution is in the form of a set of equations the K dimensions of the optimum block must satisfy, and is a function of the average error in a particular plane of the data.

5.2. Principal Component Imagery and Feature Selection

The transformation of multispectral data to principal components in the spectral dimension is found to be quite useful in providing an efficient means of data classification and storage. The first three principal component images (for the aircraft data) are shown to contain essentially all the information present in the original six spectral images. For the three channel satellite data only the first two principal component images are significant. The six Fourier component images from the aircraft data are found to exhibit substantially less information packing.

The problem of feature selection is simplified to a "first n" rule in the transform domain, and the results are experimentally verified over the aircraft data set. The results show that the first n transform features are at least as good as the "best" n spectral channels.

5.3. Encoder Performance Based on Mean Square Error

The results of Chapters III and IV indicate that correlations in all three dimensions of the multispectral source

are sufficiently high to warrant the use of a transform coding scheme. Encoder data rates are lowest for the K-L encoder, with the Fourier and Hadamard encoders typically performing at somewhat higher rates. All three encoders achieve data rates substantially below the standard (single sample) PCM encoder, and therefore represent significant reductions in bandwidth and/or storage requirements. Reductions (over standard PCM) range from approximately 3:1 at 1% distortion to more than 20:1 at 10% distortion using the K-L encoder over both the aircraft and satellite data sets.

The area to be encoded is found to be relatively insensitive to the encoder parameters. This result indicates that it might be reasonable to design an encoder to efficiently handle a particular type of terrain designated most important, and still perform well over statistically different areas.

The effects of data block structure are significant in that the encoder data rates are found to be a function of the dimensionality of the data block.

5.4. Encoder Performance Based On Classification Accuracy

The ability of the transform encoder to preserve class separability as determined by a maximum likelihood decision rule is quite good. Very little (≈6%) reduction in classification accuracy is evident for data rates as low as 0.25 bits/vector element using the K-L encoder, while the standard PCM encoder degrades rapidly (70% reduction at 1.0 bits/vector element).

LIST OF REFERENCES

LIST OF REFERENCES

[1] Special Issue on Remote Environmental Sensing, Proceedings of the IEEE, Vol. 57, No. 4, pp. 371-742, April 1969.

[2] Proceedings of the Seventh International Symposium on Remote Sensing of Environment, University of Michigan, May 1971.

[3] R.A. Holmes, R.B. MacDonald, "The Physical Basis of System Design for Remote Sensing in Agriculture," Proceedings of the IEEE, Vol. 5, No. 4, pp. 629-639, April 1969.

[4] Laboratory for Applications of Remote Sensing, (LARS) Purdue University, Research Bulletin No. 973, December 1970.

[5] D.A. Landgrebe, "Data Analysis and Remotely Sensed Data," AIAA Earth Resources Observations and Information Systems Meeting, AIAA Paper No. 70-292.

[6] Special Issue on Signal Processing for Digital Communications, IEEE Transactions on Communication Technology, Vol. Com-19, No. 6, pp. 872, 1155, February 1972.

[7] H.P. Kramer, M.V. Mathews, "A Linear Coding for Transmitting a Set of Correlated Signals," IRE Transactions on Information Theory, Vol. IT-2, pp. 41-46, September 1956.

[8] J.J.Y. Huang, P.M. Schultheiss, "Block Quantization of Correlated Gaussian Random Variables," IEEE Transactions on Communication Systems, pp. 289-296, September 1963.

[9] W.K. Pratt, J. Kane, H.C. Andrews, "Hadamard Transform Image Coding," Proceedings of the IEEE, Vol. 57, No. 1, pp. 58-67, January 1969.

[10] A. Habibi, P.A. Wintz, "Image Coding by Linear Transformation and Block Quantization," IEEE Transactions on Communication Technology, Vol. Com-19, No. 1, 1971.

[11] T.S. Huang, W.F. Schreiber, O.J. Tretiak, "Image Processing," Proceedings of the IEEE, pp. 1586-1607, November 1971.

[12] L.C. Wilkins, P.A. Wintz, "Bibliography on Data Compression, Picture Properties, and Picture Coding," IEEE Transactions on Information Theory, Vol. IT-17, pp. 180-197, March 1971.

[13] W.K. Pratt, "A Bibliography on Television Bandwidth Reduction Studies," IEEE Transactions on Information Theory, Vol. IT-13, No. 1, January 1967.

[14] A. Rosenfeld, "Bandwidth Reduction Bibliography," IEEE Transactions on Information Theory, Vol. IT-14, No. 4, pp. 601-602, July 1968.

[15] R.M. Haralick, I. Dinstein, "An Iterative Clustering Procedure," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-1, No. 3, pp. 275-289, July 1971.

[16] C.L. Crawford, R.H. Dye, D.S. Hanson, "Signature Data Processing Study," BSR 2949, The Bendix Corporation Aerospace Systems Division, August 1970.

[17] R. Kern, J. Silverman, "Multispectral Scanner Data Redundancy Study, "NASA Contract NA S5-21151, General Electric Company, Valley Forge Space Center, Philadelphia, Pennsylvania, 1971.

[18] H.F. Silverman, "On the Uses of Transforms for Satellite Image Processing," Proceedings of the Seventh International Symposium on Remote Sensing of Environment, Vol. II, pp. 1377-1386, May 1971.

[19] K.S. Fu, P.J. Min, "On Feature Selection in Multiclass Pattern Recognition," Technical Report TR-EE 68-17, Purdue University, July 1968.

[20] A.G. Wacker, "Minimum Distance Approach to Classification," Ph.D. Dissertation, Purdue University, pp. 110-253, January 1972.

[21] R.G. Gallager, Information Theory and Reliable Communications, Chapters 2 and 9, Wiley, 1968.

[22] T. Berger, Rate Distortion Theory: A Mathematical Basis for Data Compression, Prentice-Hall, 1971.

[23] B.J. Bunin, "Rate Distortion Functions for Correlated Gaussian Sources," Ph.D. Dissertation, Polytechnic Institute of Brooklyn, June 1970.

[24] R. Hogg, A. Craig, Introduction to Mathematical Statistics, Macmillan, Second Edition, 1966.

[25] C.J. Palermo, R.V. Palermo, H. Horwitz, "The Use of Data Omission for Redundancy Removal," Prepared for the Institute of Science and Technology, University of Michigan.

[26] P.A. Wintz, A.J. Kurtenbach, "Waveform Error Control in PCM Telemetry," IEEE Transactions on Information Theory, Vol. IT-14, No. 5, September 1968.

[27] H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," Journal of Educational Psychology, Vol. 24, pp. 417-441, 498-520, 1933.

[28] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, Inc., 1962.

[29] M. Loéve, Probability Theory, 3rd Edition, Van Nostrand, Princeton, 1963.

[30] K.S. Fu, Sequential Methods in Pattern Recognition and Machine Learning, Academic Press, 1968.

[31] H.J. Lendon, D. Slepion, "Some Computer Experiments in Picture Processing for Bandwidth Reduction," Bell Systems Technical Journal, Vol. 50, pp. 1525-1540, May-June 1971.

[32] J.D. Kennedy, S.J. Clark, W.A. Parkyn, Jr., "Digital Imagery Data Compression Techniques," Report No. MDC G 0402, Mc Donnel Douglas Corporation, Huntington Beach, California, January 1970.

[33] T.S. Huang, J.W. Woods, "Picture Bandwidth Compression by Linear Transformation and Block Quantization," Presented at the 1964 Symposium on Picture Bandwidth Compression, Massachusetts Institute of Technology, Cambridge, Massachusetts.

[34] J.L. Walsh, "A Colored Set of Orthogonal Functions," American Journal of Mathematics, Vol. 55, pp. 5-24, 1923.

[35] J. Hadamard, "Resolution d' une Question Relative aux Determinants," Bulletin of Science and Mathematics, series 2, Vol. 17, part 1, pp. 240-246, 1893.

[36] R.E.A.C. Paley, "On Orthogonal Matrices," Journal of Mathematical Physics, Vol. 12, pp. 311-320, 1933.

[37] J. Max, "Quantizing for Minimum Distortion," IRE Transactions on Information Theory, pp. 7-12, March 1960.

[38] S.P. Lloyd, "Least Square Quantizations in PCM," Bell Telephone Laboratory Memorandum, Murray Hill, N.S. (unpublished).

[39] J.F. Hayes, R. Bobilin, "Efficient Waveform Encoding," Technical Report TR-EE69-4, Purdue University, Lafayette, Indiana, February 1969.

[40] P.A. Wintz, A.J. Kurtenbach, "Analysis and Minimization of Message Error in PCM Telemetry Systems," Technical Report TR-EE67-19, Purdue University, Lafayette, Indiana, December 1967.

[41] R. Courant, D. Hilbert, Methods of Mathematical Physics, Vol. I, Interscience, New York, 1953.

[42] P.G. Hasell, Jr., L.M. Larsen, "Calibration of an Airborne Multispectral Optical Sensor," Technical Report ECOM-00013-137, Willow Run Laboratories, University of Michigan, Ann Arbor, Michigan, September 1968.

[43] Y.T. Chien, K.S. Fu, "Selection and Ordering of Feature Observations in a Pattern Recognition System," Information and Control, Vol. 12, May 1968.

[44] S. Watanabe, "The Loéve-Karhunen Expansion as a Means of Information Compression for Classification of Continuous Signals," AMRL-TR-65-114, Thomas J. Watson Research Center, IBM, October 1965.

[45] P.J. Ready, P.A. Wintz, "A Linear Transformation for Data Compression and Feature Selection in Multispectral Imagery," Purdue University, Lafayette, Indiana, LARS Information Note 072071, July 1971.

[46] K.S. Fu, D.A. Landgrebe, T.L. Phillips, "Information Processing of Remotely Sensed Agricultural Data," Proceedings of the IEEE, Vol. 57, No. 4, April 1969.

[47] P.H. Swain, T.V. Robertson, A.G. Wacker, "Comparison of the Divergence and B-Distance in Feature Selection," LARS Information Note 020871, Purdue University, Lafayette, Indiana, February 1971.

[48] R.M. Haralick, J.D. Young, D.R. Gael, K.S. Shanmugan, "A Comparative Study of Data Compression Techniques for Digital Image Transmission," Cadre Corporation, Lawrence, Kansas, January 1972.

[49] J.L. Kaltenbach, "Apollo 9 Multispectral Photographic Information," NASA Technical Memorandum X-1957 NASA, Washington, D.C., April 1970.

[50] P.E. Anuta, R.B. MacDonald, "Crop Surveys from Multiband Satellite Photography Using Digital Techniques," Remote Sensing of Environment, No. 2, pp. 53-67, 1971.

[51] P.E. Anuta, "Spatial Registration of Multispectral and Multitemporal Imagery Using Fast Fourier Transform Techniques," IEEE Transactions on Geoscience Electronics, Vol. GE-8, pp. 353-368, October 1970.

[52] M.R. Holter, W.L. Wolfe, "Optical-Mechanical Scanning Techniques," Proceedings of IRE, Vol. 47, No. 9, pp. 1546-1550, 1959.

[53] M. Tasto, P.A. Wintz, "Image Coding by Adaptive Block Quantization," IEEE Transactions on Communication Technology, Vol. COM-19, No. 6, pp. 957-972, December 1971.

[54] J.L. Brown, Jr. "Mean-Square Truncation Error in Series Expansion of Random Functions," Journal of the Society of Industrial Applied Mathematics, Vol. 8, No. 1, pp. 18-32, March 1960.

[55] C.A. Andrews, J.M. Davis, G.R. Schwarz, "Adaptive Data Compression, "Proceedings of the IEEE, Vol. 55, No. 3, March 1967.

[56] A.H. Koschman, J.G. Truxal, "Optimum Linear Filtering of Nonstationary Time Series," Proceedings of the National Electronics Conference, p. 126, 1954.

[57] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Part I, Wiley, 1968.

[58] T.J. Goblick, J.L. Holsinger, "Analog Source Digitization: A Comparison of Theory and Practice," IEEE Transactions on Information Theory (Correspondence), pp. 323-326, April 1967.

[59] P.A. Wintz, "Transform Picture Coding," Proceedings of the IEEE, Vol. 60, No. 7, July 1972.

[60] R.E. Totty, G.C. Clark, "Reconstruction Error in Waveform Transmission," IEEE Transactions on Information Theory, Vol. IT-13, April 1967.

[61] A.N. Kolmogorov, "On the Shannon Theory of Information Transmission in the Case of Continuous Signals," IEEE Transactions on Information Theory, pp. 102-108, December 1956.

APPENDICES

APPENDIX A

MEAN SQUARE ERROR AS A FUNCTION OF POSITION WITHIN THE DATA BLOCK

Consider the K-dimensional zero mean random process $x(\underline{t})$ having the following properties

1) The process is time aplitude continuous, with

$$x(\underline{t}) = x(t_1, t_2, \ldots, t_K) \qquad \text{A.1}$$

2) The process is correlation stationary [57]. The autocorrelation function is then $R(\underline{t}-\underline{\tau})$ and the process variance is

$$E\{x^2(\underline{t})\} = \sigma_x^2 \qquad \text{A.2}$$

Define the Karhunen-Loéve series expansion of $x(\underline{t})$ over the data block (interval) $[N_1, N_2, \ldots, N_n]$

$$x(\underline{t}) = \sum_{j=1}^{\infty} y_j \phi_j(\underline{t}) \qquad 0 \leq t_i \leq N_i \quad \forall i \qquad \text{A.3}$$

where the $\phi_j(\underline{t})$ are the orthonormal eigenfunctions of the integral equation

$$\lambda_j \phi_j(\underline{t}) = \int_{}^{N_1} \cdots \int_{}^{N_K} R(\underline{t}-\underline{\tau})\phi_j(\underline{\tau})\, d\underline{\tau} \qquad \text{A.4}$$

and the λ_i are the eigenvalues of A.4

The random coefficients y_j are determined by

$$y_j = \int_{}^{N_1} \cdots \int_{}^{N_K} x(\underline{t})\phi_j(\underline{t})\, d\underline{t} \qquad \text{A.5}$$

Define the mean square sampling error $d_s(\underline{t})$

$$d_s(\underline{t}) = E\{[x(\underline{t})-\hat{x}(\underline{t})]^2\} \qquad \text{A.6}$$

where

$$\hat{x}(\underline{t}) = \sum_{j=1}^{n} y_j \phi_j(\underline{t}) \qquad \text{A.7}$$

Expanding A.6 in terms of A.3 and A.7 gives

$$d_s(\underline{t}) = \sum_{i=n+1}^{\infty} \sum_{j=n+1}^{\infty} E\{y_i y_j\}\phi_i(\underline{t})\phi_j(\underline{t}) \qquad \text{A.8}$$

But

$$E\{y_i y_j\} = E\left\{\int_{}^{N_1}\cdots\int_{}^{N_K}\int_{}^{N_1}\cdots\int_{}^{N_K} x(\underline{t})x(\underline{\tau})\phi_i(\underline{t})\phi_j(\underline{\tau})\,d\underline{t}\,d\underline{\tau}\right\}$$

$$= \int_{}^{N_1}\cdots\int_{}^{N_K}\int_{}^{N_1}\cdots\int_{}^{N_K} R(\underline{t}-\underline{\tau})\phi_i(\underline{t})\phi_j(\underline{\tau})\,d\underline{t}\,d\underline{\tau} \qquad \text{A.9}$$

From A.4 equation A.9 becomes

$$E\{y_i y_j\} = \int_{}^{N_1}\cdots\int_{}^{N_K} \lambda_j\phi_j(\underline{t})\phi_i(\underline{t})\,d\underline{t} \qquad \text{A.10}$$

The $\phi_i(\underline{t})$ are orthonormal. Thus A.10 gives

$$E\{y_i y_j\} = \begin{cases}\lambda_j & i=j \\ 0 & i\neq j\end{cases} \qquad \text{A.11}$$

or

$$E\{y_j^2\} = \lambda_j \qquad \text{A.12}$$

and the eigenvalues of A.4 are the variances of the y_i. Substituting A.11 into the expression for $d_s(\underline{t})$ in A.8 gives

$$d_s(\underline{t}) = \sum_{j=n+1}^{\infty} \lambda_j \phi_j^2(\underline{t}) \qquad \text{A.13}$$

From Mercer's Theorem [57]

$$\sigma_x^2 = \sum_{j=1}^{\infty} \lambda_j \phi_j^2(\underline{t}) \qquad \text{A.14}$$

Thus A.13 becomes

$$d_s(\underline{t}) = \sigma_x^2 - \sum_{j=1}^{n} \lambda_j \phi_j^2(\underline{t}) \qquad \text{A.15}$$

Equations A.13 and A.15 are the desired result. They express the total sampling error as a function of position ($\underline{t}$) within the $N_1 x N_2 x \ldots x N_K$ data block.

The mean integral square sampling error is

$$d_s = \int_{}^{N_1} \cdots \int_{}^{N_K} d_s(\underline{t})\, d\underline{t} \qquad \text{A.16}$$

Substituting A.13 into A.16 gives

$$d_s = \sum_{j=n+1}^{\infty} \lambda_j \int_{}^{N_1} \cdots \int_{}^{N_K} \phi_j^2(\underline{t})\, d\underline{t} \qquad \text{A.17a}$$

$$= \sum_{j=n+1}^{\infty} \lambda_j \qquad \text{A.17b}$$

a familiar result [54,56,57].

APPENDIX B

THE OPTIMUM BLOCK SIZE

Consider the Karhunen-Loéve series expansion of the K-dimensional random process $x(\underline{t})$ defined in Appendix A. The following problem is examined. Choose the N_i, $i=1,2,\cdots,K$ that minimize the mean integral sampling error d_s as defined in A.17, subject to the constraint

$$\prod_{j=1}^{K} N_j = N \qquad \text{B.1}$$

Introducing the Lagrange multiplier Λ the problem becomes the solution to the following K+1 equations

$$\frac{\partial}{\partial N_i}\left\{d_s + \Lambda \prod_{i=1}^{K} N_i\right\} = 0 \qquad i = 1,2,\cdots,K \qquad \text{B.2}$$

$$\prod_{j=1}^{K} N_j = N \qquad \text{B.3}$$

Using the results of Appendix A, B.2 becomes

$$\frac{\partial}{\partial N_i}\left\{\sum_{j=n+1}^{\infty}\lambda_j + \Lambda\prod_{j=1}^{K}N_j\right\} = 0 \qquad \text{B.4}$$

or

$$\sum_{j=n+1}^{\infty}\frac{\partial\lambda_j}{\partial N_i} + \Lambda\frac{N}{N_i} = 0 \qquad \text{B.5}$$

The first term in B.5 is now examined. Multiplying both sides of A.4 by $\phi_j(\underline{t})$ and integrating gives

$$\lambda_j = \int_{N_1}\cdots\int_{N_K}\int_{N_1}\cdots\int_{N_K} R(\underline{t}-\underline{\tau})\phi_j(\underline{t})\phi_j(\underline{\tau})\,d\underline{\tau}\,d\underline{t} \qquad \text{B.6}$$

The partial derivative of λ_j is then

$$\frac{\partial\lambda_j}{\partial N_i} = \oint_i\oint_i\frac{\partial}{\partial N_i}\int_{N_i}\int_{N_i} R(\underline{t}-\underline{\tau})\phi_j(\underline{t})\phi_j(\tau)\,d\underline{\tau}\,d\underline{t} \qquad \text{B.7}$$

where $\oint$ is introduced as the multiple integral

$$\oint_i \Rightarrow \int_{N_1}\cdots\int_{N_{i-1}}\int_{N_{i+1}}\cdots\int_{N_K} \qquad \text{B.8}$$

i.e., integration over all N_j, $j \neq i$

Carrying out the differentiation in B.7 gives

$$\frac{\partial \lambda_j}{\partial N_i} = \oint_i \oint_i \Bigg[\int^{N_i} R(\underline{t}'-\underline{\tau}', N_i-\tau_i)\phi_j(\underline{\tau})\phi_j(\underline{t}', N_i)d\tau_i +$$

$$\int^{N_i} \Big(R(\underline{t}'-\underline{\tau}', t_i-N_i)\phi_j(\underline{\tau}', N_i)\phi_j(\underline{t}) +$$

$$\int^{N_i} R(\underline{t}-\underline{\tau})\dot{\phi}_j(\underline{t})\phi_j(\underline{\tau})d\tau_i +$$

$$\int^{N_i} R(\underline{t}-\underline{\tau})\phi_j(\underline{t})\dot{\phi}_j(\underline{\tau})d\tau_i \Big) dt_i \Bigg] d\underline{t}' d\underline{\tau}' \qquad \text{B.9}$$

where

$$\underline{t}' = (t_1 \cdots, t_{i-1}, t_{i+1}, \cdots, t_K) \qquad \text{B.10}$$

and

$$\dot{\phi}_j(\underline{t}) = \frac{\partial}{\partial N_i}\Big[\phi_j(\underline{t})\Big] \qquad \text{B.11}$$

Rewriting B.9

$$\frac{\partial \lambda_j}{\partial N_i} = \oint_i \oint_i \int^{N_i} R(\underline{t}'-\underline{\tau}', N_i-\tau_i)\phi_j(\underline{\tau})\phi_j(\underline{t}', N_i)\,d\underline{\tau}\,dt' +$$

$$\oint_i \oint_i \int^{N_i} R(\underline{t}'-\underline{\tau}', t_i-N_i)\phi_j(\underline{\tau}', N_i)\phi_j(\underline{t})\,d\underline{\tau}'\,d\underline{t} +$$

$$2\oint_i \oint_i \int^{N_i}\int^{N_i} R(\underline{t}-\underline{\tau})\dot{\phi}_j(\tau)\phi_j(\underline{t})\,d\underline{t}\,d\underline{\tau} \qquad \text{B.12}$$

After some manipulation and using A.4 the three integrals in B.12 become

$$\text{First integral} = \oint_i \lambda_j \phi_j^2(\underline{t}', N_i)\,d\underline{t}' \qquad \text{B.13}$$

$$\text{Second integral} = \oint_i \lambda_j \phi_j^2(\underline{\tau}', N_i)\,d\underline{\tau}' \qquad \text{B.14}$$

$$\text{Third integral} = 2\int^{N_1} \cdots \int^{N_K} \lambda_j \phi_j(\underline{t})\dot{\phi}_j(\underline{t})\,d\underline{t} \qquad \text{B.15}$$

Evaluation of B.15 is accomplished using the orthonormality of the $\phi_j(\underline{t})$

$$\frac{\partial}{\partial N_i}\left[\int_{}^{N_1}\cdots\int^{N_K}\phi_j^2(\underline{t})\,d\underline{t}\right] = 0 \qquad \text{B.16}$$

or

$$\oint_i \frac{\partial}{\partial N_i}\int^{N_i}\phi_j^2(\underline{t})\,d\underline{t} = 0 \qquad \text{B.17}$$

Carrying out the differentiation and rearranging terms gives

$$\int_{}^{N_1}\cdots\int^{N_K}\dot{\phi}_j(\underline{t})\,\phi_j(\underline{t})\,d\underline{t} = \oint_i \phi_j^2(\underline{t}',N_i)\,d\underline{t}' \qquad \text{B.18}$$

Substituting B.18, B.13 and B.14 into B.12 gives

$$\frac{\partial\lambda_j}{\partial N_i} = 2\oint_i \lambda_j\phi_j^2(\underline{t}',N_i)\,d\underline{t}' - \oint_i \lambda_j\phi_j^2(\underline{t}',N_i)\,d\underline{t} \qquad \text{B.19a}$$

$$= \oint_i \lambda_j\phi_j^2(\underline{t}',N_i)\,d\underline{t}' \qquad \text{B.19b}$$

Now B.19 may be used in B.5

$$\sum_{j=n+1}^{\infty} \int_{i} \lambda_j \phi_j^2(\underline{t}', N_i) d\underline{t}' + \frac{\Lambda N}{N_i} = 0 \qquad \text{B.20}$$

or

$$N_i \sum_{j=n+1}^{\infty} \int_{i} \lambda_j \phi_j^2(\underline{t}', N_i) d\underline{t}' = -\Lambda N \qquad \text{B.21}$$

Since ΛN is a constant B.21 implies

$$N_i \sum_{j=n+1}^{\infty} \int_{i} \lambda_j \phi_j^2(\underline{t}', N_i) d\underline{t}' =$$

$$N_l \sum_{j=n+1}^{\infty} \int_{l} \lambda_j \phi_j^2(\underline{t}', N_l) d\underline{t}' \quad \forall i, l \qquad \text{B.22}$$

The set of $N_i, i=1,2,\cdots,K$ that satisify B.22 are the desired result. Based on the results of Appendix A the following interpretation may be given to B.22. First, rewrite B.22 in the following form

$$N_i \int_{i} \sum_{j=n+1}^{\infty} \lambda_j \phi_j^2(\underline{t}', N_i) d\underline{t}' = N_l \int_{l} \sum_{j=n+1}^{\infty} \lambda_j \phi_j^2(\underline{t}', N_l) d\underline{t}' \qquad \text{B.23}$$

From A.13

$$d_s(\underline{t}) = \sum_{j=n+1}^{\infty} \lambda_j \phi_j^2(\underline{t}) \qquad \text{B.24}$$

Thus the integrals in B.23 represent the sampling error as a function of $\underline{t}'$ in the plane $t_i=N_i$. Integrating over $\underline{t}'$ gives the average sampling error in the plane $t_i=N_i$. Equation B.23 states that the optimum choice of N_i, $i=1,2,\cdots,K$ is such that N_i times the average error in the $t_i=N_i$ plane is equal to N_l times the average error in the $t_l=N_l$ plane.

If the process autocorrelation function is separable, i.e.

$$R(\underline{t}-\underline{\tau}) = \prod_{j=1}^{K} R_j(t_j-\tau_j) \qquad \text{B.25}$$

then B.23 may be simplified to give

$$N_i \sum_{j=n+1}^{\infty} \lambda_j \theta_{j,i}^2(N_i) = N_l \sum_{j=n+1}^{\infty} \lambda_j \theta_{j,l}^2(N_l) \qquad \text{B.26}$$

where the $\Theta_{j,i}(N_i)$ are the factors in the separable eigenfunctions

$$\phi_j(\underline{\tau}) = \prod_{i=1}^{K} \Theta_{j,i}(\tau_i) \qquad \text{B.27}$$

APPENDIX C

BIT DISTRIBUTION FOR THE K-L ENCODER

Tables C.1 and C.2 below represent integer bit assignments resulting from the application of equations 2.60 and 2.63 to the aircraft scanner data using the K-L encoder and 1x1x6 and 1x64x1 data blocks, respectively. The resulting percent distortions are presented in Figures 3.7 and 3.14.

Table C.1. Bit Distribution for the K-L Encoder and 1x1x6 Data Blocks

No. Bits	R=5.0	R=3.0	R=1.0	R=0.33
8	1	-	-	-
7	-	-	-	-
6	2	1	-	-
5	3	-	-	-
4	4-5	2	-	-
3	6	3	1	-
2	-	4-5	2	1
1	-	6	3	-

Table C.2. Bit Distribution for the K-L Encoder and 1x64x1 Data Blocks

No. Bits	R=2.0	R=1.0	R=0.50	R=0.25	R=0.125
9	1	-	-	-	-
8	-	-	-	-	-
7	2	1	-	-	-
6	3-5	-	1	-	-
5	6-11	2	-	1	-
4	12-21	3-4	2	-	4
3	22-26	5-10	3	2	-
2	27-29	11-20	4-8	3	2
1	30-32	21-26	9-17	4-9	3-4

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
School of Electrical Engineering
Purdue University

2a. REPORT SECURITY CLASSIFICATION
unclassified

2b. GROUP

3. REPORT TITLE
Multispectral Data Compression Through Transform Coding and Block Quanitization

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
scientific report

5. AUTHOR(S) *(First name, middle initial, last name)*
Patrick J. Ready
Paul A. Wintz

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
May 5, 1972		

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
NASA Grant No. NGL 15-005-112	TR-EE 72-2
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	LARS INFORMATION NOTE #050572

10. DISTRIBUTION STATEMENT
unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	National Aeronautics and Space Administration

13. ABSTRACT

Transform coding and block quantization techniques are applied to multispectral data for data compression purposes. Two types of multispectral data are considered, (1) aircraft scanner data, and (2) digitized satellite imagery. The multispectral source is defined and an appropriate mathematical model proposed.

The Karhunen-Loeve, Fourier, and Hadamard encoders are considered and are compared to the rate distortion function for the equivalent gaussian source and to the performance of the single sample PCM encoder.

Minimization of the total encoder system error over the number of retained transform coefficients and corresponding bit distribution for a fixed data rate and block size is considered and an appropriate solution proposed. Minimization of the sampling error over the data block size for the continuous source is also considered.

The Karhunen-Loeve transformation is applied to the spectral dimension of the multispectral source and the resulting principal components are evaluated as feature vectors for use in data classification.

Experimental results using the transform encoder and several different (i.e., one, two, and three dimensional) data blocks are presented for both the satellite and aircraft data sets. Performances of the encoders over the three test regions within the satellite data are evaluated and compared.

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT