USCEE REPORT #441

Slant Transform Image Coding

by

Wen-Hsiung Chen

May 1973

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

average coding of a monochrome image can be reduced from 8 bits/pixel to 1 bit/pixel or 1.5 bits/pixel for the threshold and zonal coding, respectively, without seriously degrading the image quality. Studies have also indicated that zonal coding has an extremely high noise immunity, and can be practically implemented.

Spatial redundancy of color images and the limitations of human color vision have also been exploited by slant transform coding to achieve a bandwidth reduction for natural color images. It has been found by computer simulation that the average coding of a color image can be reduced from 24 bits/pixel to about 2 bits/pixel while preserving good quality reconstruction.

SLANT TRANSFORM IMAGE CODING

TABLE OF CONTENTS

		Page
ABSTRACT		ii
LIST OF FIGURES		vi
LIST OF TABLES		x
LIST OF EXHIBITS		xi
ACKNOWLEDGEMENT		xii
CHAPTER		
1.	INTRODUCTION	1
	1.1 General Background	1
	1.2 Research Objectives	3
	1.3 Original Images	4
	1.4 Organization of Dissertation	4
2.	TWO DIMENSIONAL TRANSFORMS	9
	2.1 General Representation	9
	2.2 Fourier Transform	12
	2.3 Hadamard Transform	13
	2.4 Haar Transform	17
	2.5 Karhunen-Loeve Transform	18
3.	DEVELOPMENT OF THE SLANT TRANSFORM	27
	3.1 Definition of the Slant Transform	27
	3.2 Construction of the Slant Transform	28
	3.3 Fast Transform Algorithm	32
	3.4 Summary	39

Page

4. STATISTICAL ANALYSIS OF THE SLANT TRANSFORM. . . 46
   4.1 Moments 46
   4.2 Probability Density 49

5. SLANT TRANSFORM MONOCHROME IMAGE CODING 53
   5.1 Slant Transform Bandwidth Reduction 57
      5.1.1 Threshold Sampling 58
      5.1.2 Zonal Sampling 65
   5.2 Quantization Effects 79
      5.2.1 Calculation of M.S.E. - Approximate Method. . 79
      5.2.2 Calculation of M.S.E. - Exact Method 86
   5.3 Coding Effects 92
      5.3.1 Threshold Coding 92
      5.3.2 Zonal Coding and Bit Allocation 100

6. EFFECT OF CHANNEL ERROR FOR THE SLANT TRANSFORM IMAGE TRANSMISSION. 112
   6.1 Channel Error Effect - Without Bandwidth Reduction. 112
   6.2 Channel Error Effect - With Bandwidth Reduction. . . 114

7. SLANT TRANSFORM COLOR IMAGE CODING 120
   7.1 Color Image Coordinate Conversion 122
   7.2 Color Image Transform Coding 123
   7.3 Experimental Results 137

8. SUMMARY. 144

APPENDIX 148
REFERENCES 165

LIST OF FIGURES

Figure		Page
1-1	Original Monochrome Images	5
1-2	Original Color Images	6
2-1	Transform Image Coding System	10
2-2	Fourier Transform Waveforms	14
2-3	Hadamard Transform Waveforms	16
2-4	Haar Transform Waveforms	19
2-5	Karhunen-Loeve Transform Waveforms	21
2-6	Fourier Transform Domain Display	23
2-7	Hadamard Transform Domain Display	24
2-8	Haar Transform Domain Display	25
2-9	Karhunen-Loeve Transform Domain Display	26
3-1	Slant Transform Computational Flowchart of Order 4	33
3-2	Slant Transform Computational Flowchart of Order 8	36
3-3	Waveforms of Slant Matrix, S_{16}	43
3-4	Waveforms of Slant Matrix, $S_8 \times H_2$	44
3-5	Waveforms of Slant Matrix, $S_4 \times H_4$	45
5-1	Slant Transform Monochrome Image Coding System	54
5-2	Slant Transform Domain Display	56
5-3	Number of Slant Transform Samples Below Threshold Versus Threshold Level	60

Figure		Page
5-4	Number of Transform Samples of GIRL Below Threshold Versus Threshold Level	61
5-5	Slant Transform Threshold Sampling in 16 x 16 Pixel Blocks, Unquantized Transform	62
5-6	Hadamard and Haar Transform Threshold Sampling in 16 x 16 Pixel Blocks, Unquantized Transform	63
5-7	Fourier and Karhunen-Loeve Transform Threshold Sampling in 16 x 16 Pixel Blocks, Unquantized Transform	64
5-8	Full Size Slant Transform of Covariance Matrix-Markov Process	70
5-9	Transform Domain Variances	71
5-10	Mean Square Error Performance of Image Transforms as a Function of Block Size	73
5-11	Slant Transform Zonal Sampling in 16 x 16 Pixel Blcoks, Unquantized Transform	75
5-12	Hadamard and Haar Zonal Sampling in 16 x 16 Pixel Blocks, Unquantized Transform	76
5-13	Fourier and Karhunen-Loeve Transform Zonal Sampling in 16 x 16 Pixel Blocks, Unquantized Transform	77
5-14	Decision and Reconstruction Levels of dc Samples	82
5-15	Decision and Reconstruction Levels of ac Samples	83
5-16	Sample Reduction Versus Average Code Bits Per Pixel for Slant Transform Threshold Sampling of GIRL Image in 16 x 16 Pixel Blocks	96
5-17	Slant Transform Threshold Sampling in 16 x 16 Pixel Blocks, Quantized Transform	98

Figure		Page
5-18	Hadamard Transform Threshold Sampling in 16 x 16 Pixel Blocks, Quantized Transform.	99
5-19	Mean Square Error Versus Sample Reduction Factor for the Slant Transform Threshold Sampling of GIRL in 16 x 16 Pixel Blocks.	102
5-20	Typical Bit Assignments for the Slant Transform Zonal Coding in 16 x 16 Pixel Block	105
5-21	Mean Square Error Performance of Image Transforms as a Function of Block Size.	107
5-22	Slant Transform Zonal Sampling in 16 x 16 Pixel Blocks, Quantized Transform.	109
5-23	Hadamard Transform Zonal Sampling in 16 x 16 Pixel Blocks, Quantized Transform.	110
5-24	Comparison of the Optimal Threshold and Zonal Coding as a Function of Average Code Bits/Pixel for the GIRL Image.	111
6-1	Model of a Binary Symmetric Channel.	113
6-2	Spatial Domain Coding Effects of Channel Error, 8 Bits/Pixel.	115
6-3	Slant Transform Coding Effects of Channel Error, 8 Bits/Pixel.	116
6-4	Slant Transform Threshold Coding Effects of Channel Errors, Average Coding of 1.5 Bits/Pixel.	117
6-5	Slant Transform Zonal Coding Effects of Channel Errors, Average Coding of 1.5 Bits/Pixel.	118
7-1	Slant Transform Color Image Coding System.	121
7-2	Typical Bit Assignments for the Slant Transform Zonal Coding of Color Images.	126

Figure		Page
7-3	Mean Square Error for Various Color Plane Bit Assignments	136
7-4	R, G, B Tristimulus Color Planes of the Original Image	137
7-5	Y, I, Q Tristimulus Color Planes of the Original Image	138
7-6	Slant Transform Domain of Y, I, Q Images	139
7-7	Slant Transform Zonal Coding, 2 Bits/Pixel, $\hat{Y}, \hat{I}, \hat{Q}$	140
7-8	Slant Transform Zonal Coding, 2 Bits/Pixel, $\hat{R}, \hat{G}, \hat{B}$	141
7-9	Slant Transform Zonal Coding with Channel Errors, 2 Bits/Pixel	142
7-10	Slant Transform Zonal Coding with Channel Errors, 3 Bits/Pixel	143
A	Slant Transform Zonal Coding of Photographic Density	150
C-1	Rate Versus Distortion for Various Transforms for a First-Order Markov Process, N=16	153
C-2	Rate Versus Distortion for Various Transforms for a First-Order Markov Process, N=256	154

LIST OF TABLES

Table		Page
5-1	Mean Square Error Between the Original GIRL and its Threshold Sampling Reconstructions	66
5-2	Mean Square Error Between the Original GIRL and its Maximum Variance Zonal Sampling Reconstructions	78
5-3	Relationship Between Sample Reduction and Average Bits Per Pixel of Slant Transform Threshold Coded GIRL Image	95
5-4	Mean Square Error Versus Sample Reduction Factor of Slant Transform Threshold Coded GIRL Image	101
7-1	Energy Distribution of Color Planes	130
7-2	Mean Square Error Between the YIQ GIRL and Slant Transform Coded $\hat{Y}\hat{I}\hat{Q}$ GIRL	134

LIST OF EXHIBITS

Exhibit		Page
3-1	The Decomposition of a Generalized N x N Slant Matrix	37
3-2	The Decomposition of a Generalized N x N Inverse Slant Matrix	40

ACKNOWLEDGEMENT

The author wishes to express his most sincere gratitude to the chairman of his committee, Professor W. K. Pratt, for the guidance, advice and encouragement throughout the course of this research. The efforts of his committee, Professors L. R. Welch and T. E. Harris, are also greatly appreciated.

This research was supported by the Advanced Research Project Agency of the Department of Defense and was monitored by the Air Force Eastern Test Range under Contract No. F08606-72-C-0008.

1. INTRODUCTION

1.1 General Background

During the past twenty years the applications of electronic imagery have grown enormously. This growth has placed severe demands on the capabilities of communication systems since conventional television transmission requires exceptional wide bandwidths. One means of bandwidth reduction that has shown particular promise is the transform image coding process.

In 1968 the concept of coding and transmitting the two dimensional Fourier transform of an image, computed by a fast computational algorithm rather than the image itself, was introduced [1, 2]. This was followed shortly thereafter by the discovery that the Hadamard transform could be utilized in place of the Fourier transform with a considerable decrease in computational requirements [3]. Investigations then began into the application of the Karhunen-Loeve [4] and Haar [5] transforms for image coding. The Karhunen-Loeve transform provides minimum mean square error coding performance but, unfortunately, does not possess a fast computational algorithm. On the other hand, the Haar transform has the attribute of an extremely efficient computational algorithm, but results in a relatively large coding error. None of the transforms mentioned above,

however, has been expressly tailored to the characteristic of an image.

A major attribute of an image transform is that the transform compact the image energy to a few of the transform domain samples. A high degree of energy compaction will result if the basis vectors of the transform matrix "resemble" typical horizontal or vertical lines of an image. If the lines of a typical monochrome image are examined, it will be found that a large number of the lines are of constant grey level over a considerable length. The Fourier, Hadamard, and Haar transforms possess a constant valued basis vector that provides an efficient representation for constant grey level image lines, while the Karhunen-Loeve transform has a nearly constant basis vector suitable for this representation. Another type of typical image line is one which increases or decreases in brightness over the length in a linear fashion. None of the transforms previously mentioned possess a basis vector that efficiently represents such image lines.

Shibata and Enomoto have introduced orthogonal transforms containing a "slant" basis vector for data of vector lengths of four and eight [6]. The slant vector is a discrete sawtooth waveform decreasing in uniform steps over its length, which is suitable for efficiently representing gradual brightness changes in an image line. Their work gives no indication of a construction for larger size data vectors, nor exhibits the use of a fast computational algorithm. In order to achieve

a high degree of image coding compression with transform coding techniques, it is necessary to perform the two dimensional transform over block sizes of 16 x 16 picture elements or greater. For large block sizes, computation is usually not feasible unless a fast algorithm is employed.

1.2 Research Objectives

With this background research has been undertaken to develop a slant transform of variable block size possessing a fast computational algorithm. The specific objectives of this research project are the analysis and evaluation of the slant transform for image coding. The approach taken toward the fulfillment of these objectives is:

(1) Development of an image coding slant transform matrix possessing: (a) an orthogonal set of basis vectors; (b) a constant basis vector; (c) a slant basis vector; (d) sequency property; (e) variable size transformation; (f) high energy compaction; and (g) fast computational algorithm.

(2) Analysis of this slant transform image coding system supported by statistical picture measurements to determine its bandwidth reduction capability.

(3) Comparison of the slant transform with the Hadamard, Fourier, Haar, and Karhunen-Loeve transforms for image coding.

(4) Studying the channel error effects on the slant transform coding system.

(5) Application of the slant transform coding technique to color images.

1.3 Original Images

Figure 1-1 shows photographs of the three original monochrome images that have been used as test images for the evaluation of the slant transform monochrome image coding system. Each image contains 256 by 256 pixels with each pixel value uniformly quantized to 255 levels. The images were read from magnetic tape, displayed on a flying spot scanner cathode ray tube display, and photographed with polaroid type 52 film. Figure 1-2 shows photographs of the original color images that have been used as test images for the slant transform color image coding system. Both of these images contain 256 by 256 pixels, with each red, green and blue tristimulus values (NTSC receiver phosphor primary system) uniformly quantized to 255 levels. The images were read from magnetic tape. displayed on an Aerojet Model SG-D2219 display unit, and photographed on high speed Ektachrome film.

1.4 Organization of Dissertation

Chapter 1 is an introduction containing a discussion of general background and objectives of the research project. The chapter closes with a summary of the organization of the dissertation.

Chapter 2 presents a general representation of a two dimensional transform image coding system and definitions of the Fourier,

(a)

(b)

(c)

Figure 1-1. Original Monochrome Image:
(a) GIRL; (b) COUPLE; (c) MOON SCENE.

(a)

(b)

Figure 1-2. Original Color Images: (a) GIRL; (b) COUPLE.

Hadamard, Haar, and Karhunen-Loeve transforms. The energy compaction property of each transform is illustrated by transform domain pictures.

Chapter 3 presents a construction of the slant transform matrix of order 2^n where n is a positive integer. A fast computational algorithm based upon the matrix decomposition is also presented. The chapter closes by introducing a class of slant transform matrices which are constructed by a direct product of the various orders of slant and Hadamard matrices.

Chapter 4 contains a statistical analysis of the slant transform domain samples. A derivation of statistical mean and variance as well as an assignment of probability models for the transform domain samples are introduced. The method introduced is generally applicable to all transforms.

Chapter 5 is devoted to a presentation of the slant transform for monochrome image coding. Two means of achieving a bandwidth reduction for the transform domain samples are introduced. A derivation of the quantization and coding scheme, as well as a derivation of the mean square error between the original image and its reconstruction, are also introduced. Again the methods derived are generally applicable for all transforms. The evaluation of the slant transform is obtained by performing the transform coding to three test images as well as to an image statistically described by a Markov

process.

Chapter 6 presents the effects of transmitting the slant transform samples through a binary symmetrical channel. The superiority of transmitting a zonal coded slant transform sample is demonstrated.

Chapter 7 is mainly an extension of the work that has been presented in chapter 5. All the quantization and coding techniques introduced in the monochromatic case are used to code transform samples of each color plane. The color coordinate conversion between the NTSC receiver phospher primary system and NTSC transmission primary system is also included.

Chapter 8 summarizes the results of the dissertation.

2. TWO DIMENSIONAL TRANSFORMS*

Figure 2.1 shows a block diagram of a generalized transform image coding system. An original digital image, denoted by f(j, k), is defined here as an array of samples of a continuous two dimensional intensity pattern of light. The samples of this image undergo a two dimensional transformation over the entire image or some subsections of the image called blocks. The resultant transform samples, denoted by F(u, v), are then operated on by a sample selector, S(u, v), that decides which samples are to be transmitted on the basis of magnitude or geometrical location in the plane. A bandwidth reduction can be achieved by this selector simply by not transmitting all of the transform domain samples. Those samples that are to be transmitted are then quantized and coded. At the receiver the samples are decoded and inversely transformed to form the reconstructed image $\hat{f}(j, k)$. The following sections contain a general representation of a two dimensional transform image coding system and definitions of the Fourier, Hadamard, Haar and Karhunen-Loeve transforms. The definition of the slant transform is deferred to the next chapter.

2.1 General Representation

Mathematically, a two dimensional transform maps a two

---

*Part of the theory in this chapter is abstracted from the second chapter of USCEE Report 387, "Transform Image Coding".

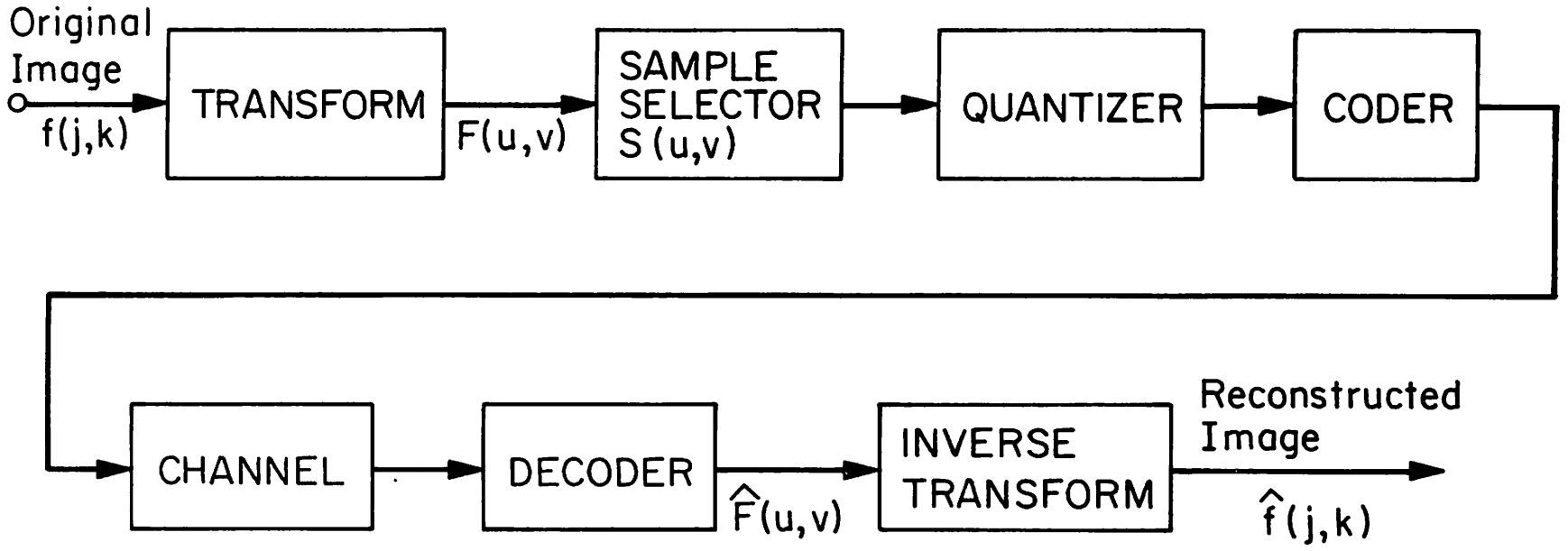


Figure 2-1. Transform Image Coding System.

dimensional image array of dimension** N×N into a two dimensional array of the same dimension by

$$F(u, v) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j, k)\, a(j, k, u, v) \qquad u, v = 0, 1, \ldots N-1 \qquad (2\text{-}1)$$

where a(j, k, u, v) is the forward transform kernel. A reverse transform is defined by

$$\hat{f}(j, k) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} f(u, v)\, b(j, k, u, v) \qquad j, k = 0, 1, \ldots N-1 \qquad (2\text{-}2)$$

where b(j, k, u, v) is the reverse transform kernel. When the function $\hat{f}(j, k)$ is equivalent to the original image f(j, k), the reverse transform is called an inverse transform.

A forward (or reverse) transform kernel is said to be separable if it can be written as

$$a(j, k, u, v) = a_j(j, u)\, a_k(k, v) \qquad (2\text{-}3)$$

A separable two dimensional transform can be computed in two steps: a one dimensional transform along each row of the image f(j , k);

$$F(u, k) = \sum_{j=0}^{N-1} f(j, k)\, a_j(j, u) \qquad (2\text{-}4)$$

and then a one dimensional transform along each column of F(u, k).

$$F(u, v) = \sum_{k=0}^{N-1} F(u, k)\, a_k(k, v) \qquad (2\text{-}5)$$

** For simplicity all arrays are assumed to be square.

It is often useful to express two dimensional transforms in matrix form if the transform kernel is separable. Let [f] be an image matrix representation of the array f(j, k) and [F] be a transformed image matrix representation of F(u, v), then a two dimensional transform can be written as

$$[F] = [a_j][f][a_k] \tag{2-6}$$

where $[a_j]$ and $[a_k]$ are one dimensional transform matrices along rows and columns of an image. If $[a_j]$ and $[a_k]$ have inverses then a two dimensional inverse transform can be written as

$$[f] = [a_j]^{-1}[F][a_k]^{-1} \tag{2-7}$$

2.2 Fourier Transform

The two dimensional Fourier transform [1, 2, 7] of an image f(j, k) can be expressed in series form as

$$F(u, v) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j, k) \exp \left[- \frac{2\pi i}{N} (uj + vk) \right] \tag{2-8}$$

The inverse Fourier transform can be written as

$$f(j, k) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \exp \left[\frac{2\pi i}{N} (uj + vk) \right] \tag{2-9}$$

The two dimensional transform can be computed as two sequential one dimensional transforms since the transform kernel is separable.

The two dimensional Fourier transform can be put into matrix form by defining the symmetric unitary matrix

$$[\mathcal{J}] = \frac{1}{\sqrt{N}} \exp\left(-\frac{2\pi i}{N} ju\right) \tag{2-10}$$

Thus, the forward transform can be written as

$$[F] = [\mathcal{J}][f][\mathcal{J}]^T \tag{2-11}$$

and the inverse transform can be written as

$$[f] = [\mathcal{J}]^{*T}[F][\mathcal{J}]^* \tag{2-12}$$

Figure 2-2 illustrates a sketch of the Fourier transform waveforms with N=16. Figure 2-6 shows the full size two dimensional Fourier transform displays of the three original monochrome images shown in Figure 1-1. In each display the original image has been multiplied by the checkerboard function $(-1)^{j+k}$ in order to shift the zero frequency of the transform to the center of the display such that a photograph similar to the Fourier diffraction pattern of a coherent optical system can be achieved [7]. In addition, the logarithm of the absolute magnitude of each transform sample is displayed rather than the absolute magnitude itself in order to reduce the dynamic range of the display. From these displays it can be seen that most of the larger transform samples or energy are concentrating around the zero frequency region. The symmetrical property [7] of the Fourier transform domain can also be seen from the displays.

2.3 Hadamard Transform

The Hadamard transform [3] is based on the Hadamard matrix which is a square array of plus or minus ones. The lowest order

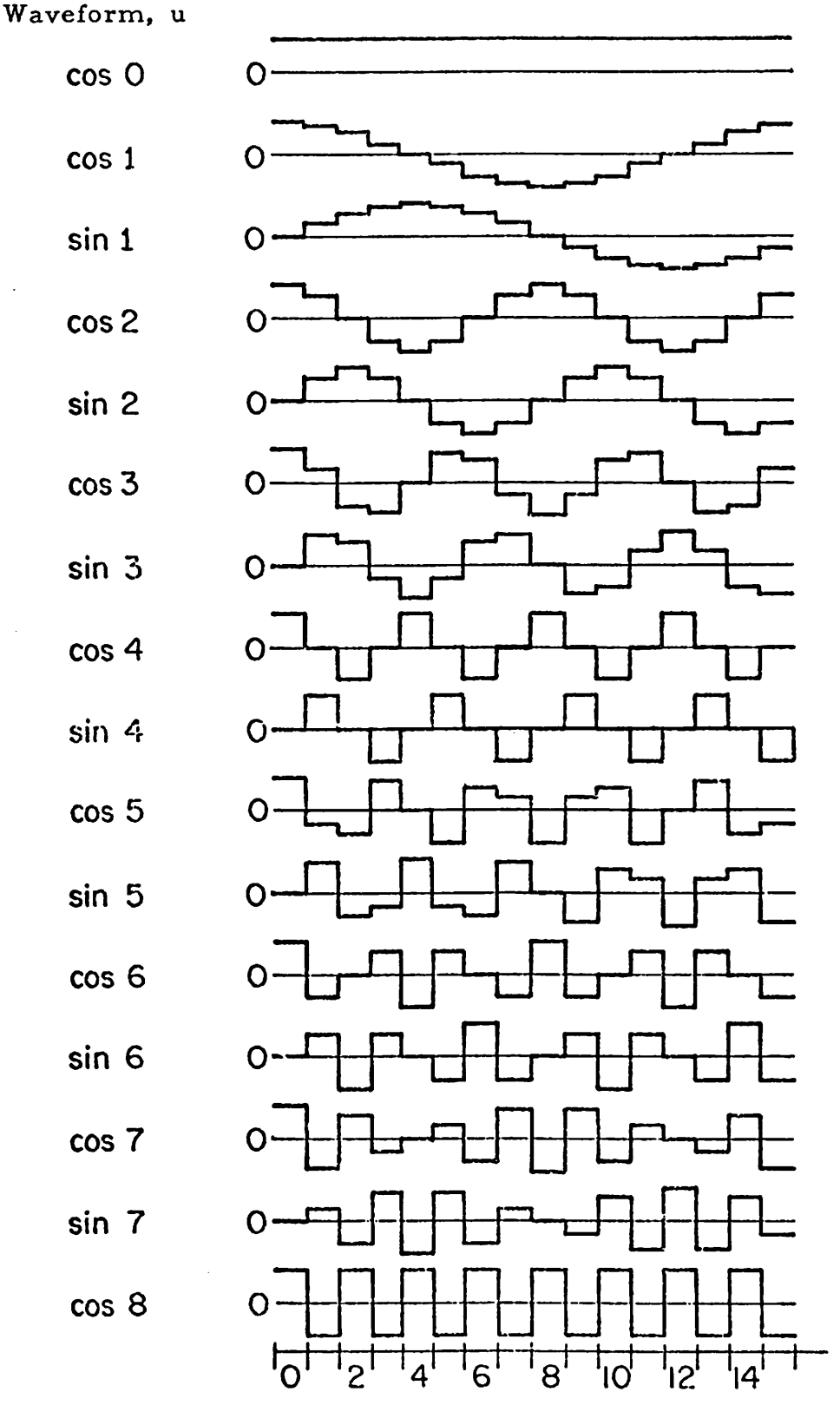


Figure 2-2. Fourier Transform Waveforms, N=16.

Hadamard matrix can be written as

$$[\mathcal{H}_2] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2-13}$$

and the construction of a Hadamard matrix of order N can be written by the following recursive relation:

$$[\mathcal{H}_N] = \frac{1}{\sqrt{2}} \begin{bmatrix} H_{\frac{N}{2}} & H_{\frac{N}{2}} \\ H_{\frac{N}{2}} & -H_{\frac{N}{2}} \end{bmatrix} \begin{bmatrix} \\ P \\ \\ \end{bmatrix} \tag{2-14}$$

where $N=2^n$ and n is an integer. The matrix [P] is a permutation matrix which permutes the rows of H_N such that the number of sign changes of each row increases with the row index. This is the sequency ordered Hadamard matrix.

Equation (2-6) and (2-7) can be used to express the two dimensional Hadamard transform matrix form by noting that the Hadamard matrix is real, symmetrical, and orthonormal. Therefore, the forward transform can be written as

$$[F] = [\mathcal{H}][f][\mathcal{H}] \tag{2-15}$$

and the inverse transform becomes

$$[f] = [\mathcal{H}][F][\mathcal{H}] \tag{2-16}$$

Figure 2-3 contains a sketch of the Hadamard transform waveforms of order 16. A sequency property and a constant basis vector can easily be seen in the waveforms. Figure 2-7 shows the two

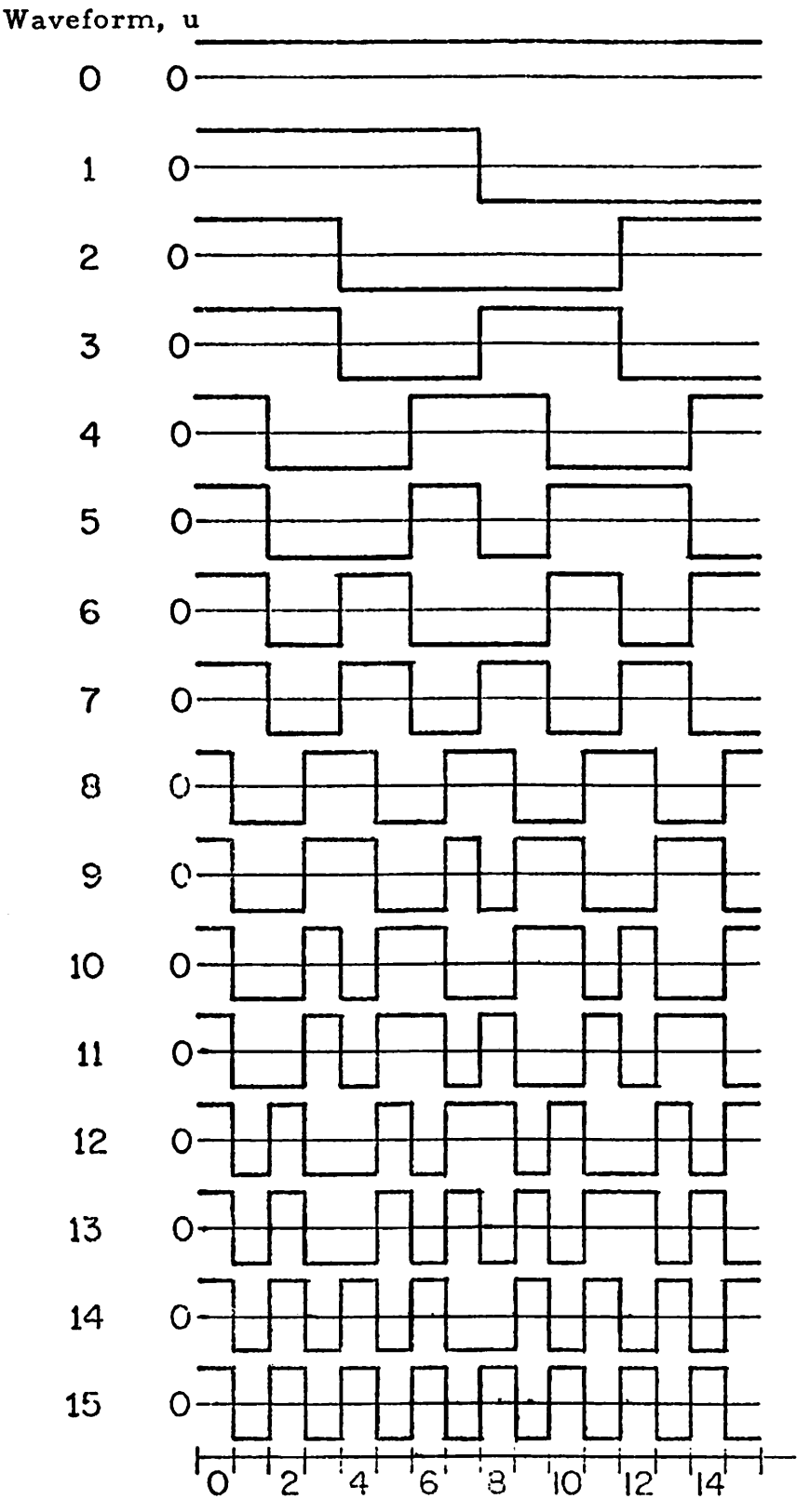


Figure 2-3. Hadamard Transform Waveforms, N=16.

dimensional Hadamard transform displays of the three original monochrome images in Figure 1-1. As in the Fourier transform display, the logarithm of the absolute value of each transform sample is displayed in order to reduce the dynamic range of transform samples. In these displays a certain degree of energy compaction can be seen on the upper-left corner of each display which is the low frequency region of transform samples.

2.4 Haar Transform

The Haar transform [5] is based on the Haar matrix which is a square array of elements plus one, minus one, and zero. A normalized Haar matrix of order 16 x 16 can be written in the following form

$$[R_{16}] = \frac{1}{\sqrt{16}} \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
\sqrt{2} & \sqrt{2} & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & -\sqrt{2} & -\sqrt{2} \\
2 & 2 & -2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 2 & -2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & -2 & -2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & -2 & -2 \\
2\sqrt{2} & -2\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2\sqrt{2} & -2\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2\sqrt{2} & -2\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{2} & -2\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{2} & -2\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{2} & -2\sqrt{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{2} & -2\sqrt{2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{2} & -2\sqrt{2}
\end{bmatrix} \tag{2-17}$$

Based on this pattern the Haar Matrices of order $N(=2^n; n=1, 2, 3, \ldots)$ can easily be constructed. It can be seen that the Haar matrices are real and orthonormal.

The two dimensional forward Haar transform can be written in matrix form as

$$[F] = [\mathcal{R}][f][\mathcal{R}]^T \tag{2-18}$$

and the inverse transform can be written as

$$[f] = [\mathcal{R}]^T[f][\mathcal{R}] \tag{2-19}$$

A sketch of the Haar transform waveform of order 16 is shown in Figure 2-4. Figure 2-8 shows the full size two dimensional Haar transform displays of the three original monochrome images. Again, the logarithm of the absolute value of each transform sample is displayed. The energy compaction property of Haar transform can also be seen in the upper-left corner but it is not nearly as good as the Hadamard or Fourier transform.

2.5 Karhunen-Loeve Transform

The Karhunen-Loeve transform is a special case of an eigenvector matrix transformation [9-13]. Let f_j be a column vector representing the rows of an image [f], then the covariance matrix of this vector can be written as

$$[C_j] = E\left\{[f_j(i) - \overline{f_j(i)}][f_j(ii) - \overline{f_j(ii)}]^T\right\} \tag{2-20}$$

where $i, ii = 0, 1, \ldots, N-1$ for the matrix of $N \times N$. The eigenvectors of

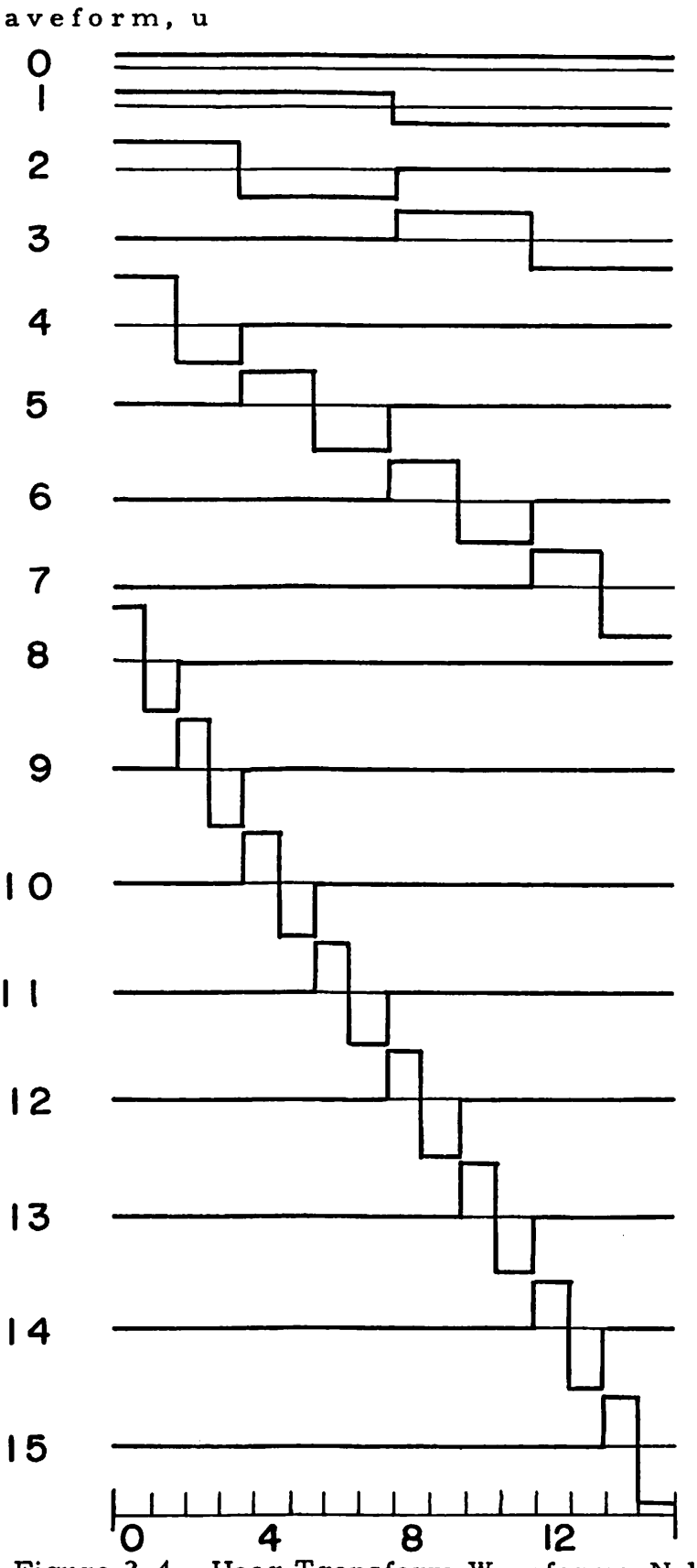


Figure 2-4. Haar Transform Waveforms, N=16.

$[C_j]$ are column vectors $[\mathcal{K}_{j_i}]$, $i = 0, 1, \ldots, N-1$, satisfying

$$[C_j][\mathcal{K}_{j_i}] = [\lambda_{j_i}][\mathcal{K}_{j_i}] \tag{2-21}$$

where $[\lambda_{j_i}]$ are the eigenvalues of $[C_j]$. The Karhunen-Loeve matrix constructed by the eigenvector columns which can be written as

$$[\mathcal{K}_j] = [\mathcal{K}_{j_0}\ \mathcal{K}_{j_1} \cdots \mathcal{K}_{j_{N-1}}] \tag{2-22}$$

if the eigenvalues are located along the diagonal of a matrix $[\lambda_j]$, then the following relation holds:

$$[C_j][\mathcal{K}_j] = [\mathcal{K}_j][\lambda_j] \tag{2-23}$$

Similarly, the Karhunen-Loeve matrix of a column vector representing those columns of the image can also be constructed by the following relation:

$$[C_k][\mathcal{K}_k] = [\mathcal{K}_k][\lambda_k] \tag{2-24}$$

A two dimensional separable Karhunen-Loeve transform can then be defined as

$$[F] = [\mathcal{K}_j][f][\mathcal{K}_k]^T \tag{2-25}$$

Figure 2-5 contains a sketch of the Karhunen-Loeve waveforms of order 16. These waveforms were obtained by assuming that the covariance matrix along an original image line is the first order Markov process with correlation coefficient $\rho = 0.95$. Figure 2-9 shows the two dimensional Karhunen-Loeve transform displays of the

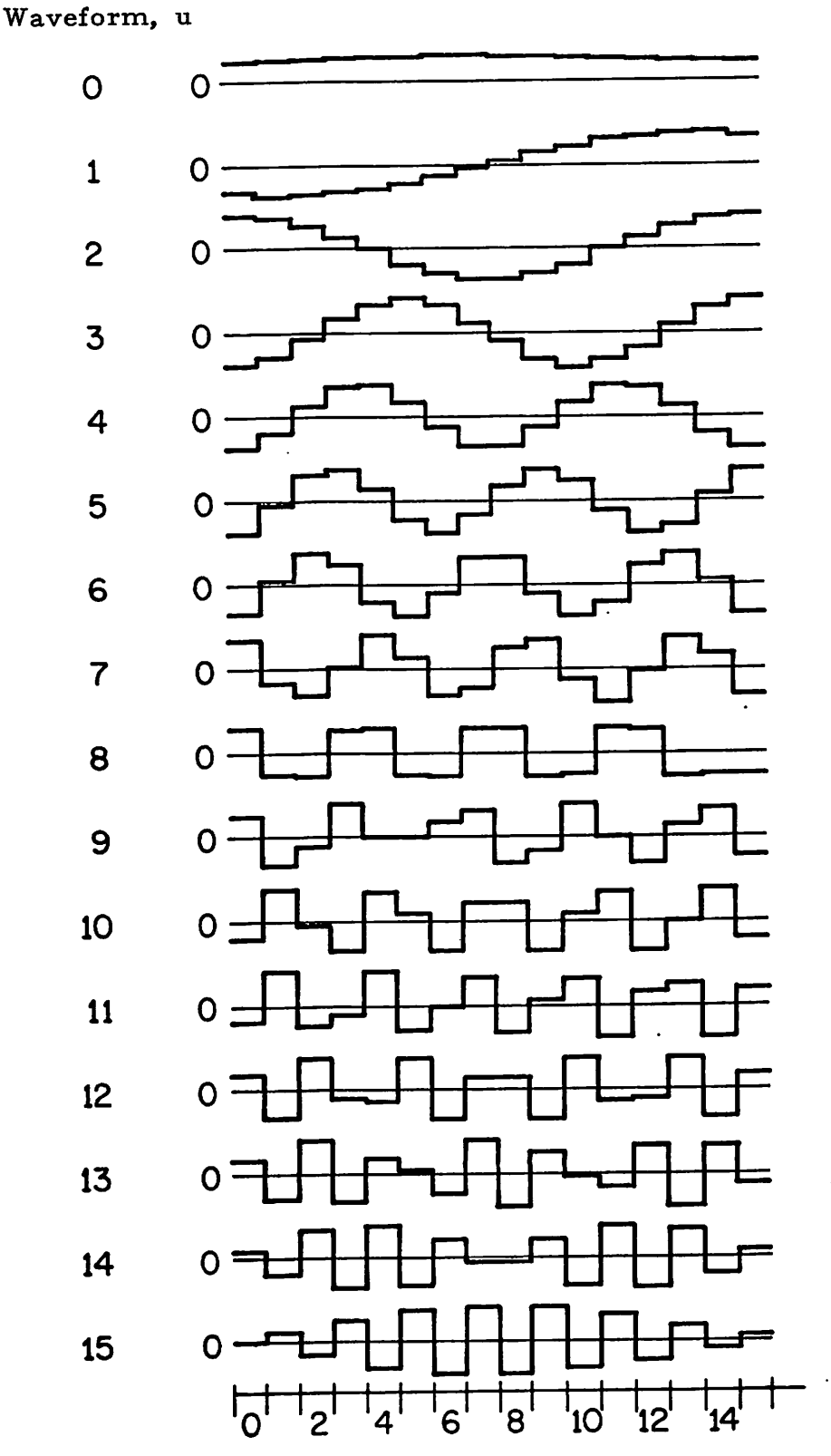
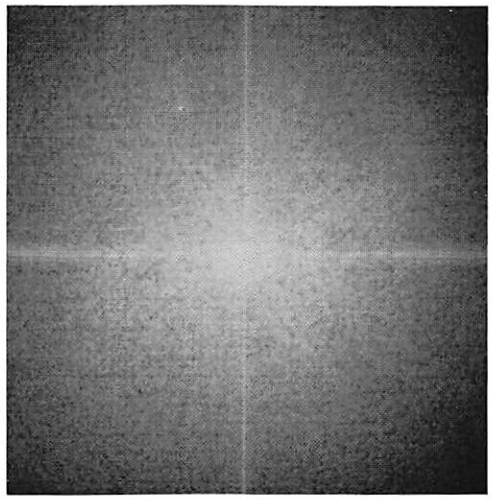
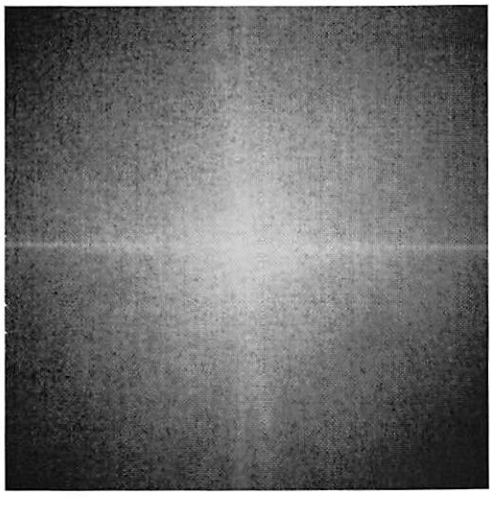


Figure 2-5. Karhunen-Loeve Transform Waveforms, N=16.

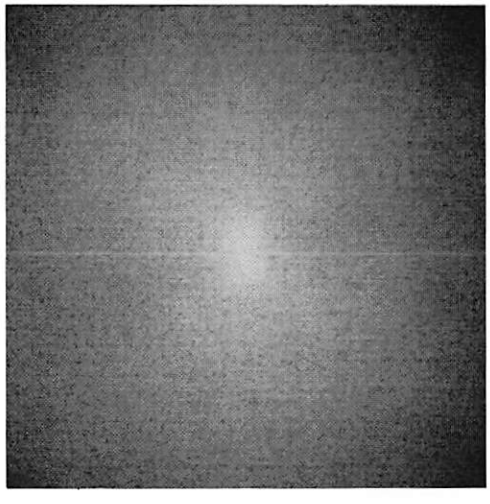
GIRL image. The transform was performed by brute force matrix multiplication since there is no known fast computational algorithm for the Karhunen-Loeve transformation.



(a)

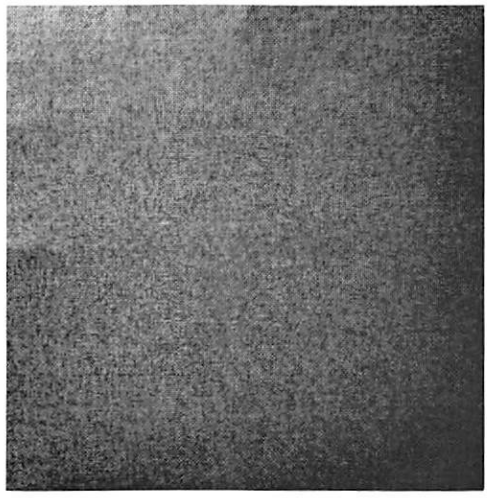


(b)

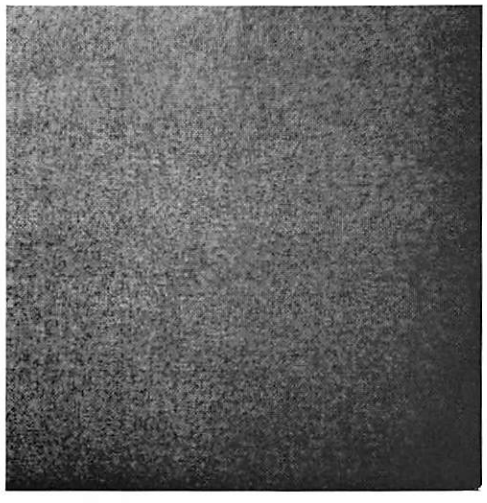


(c)

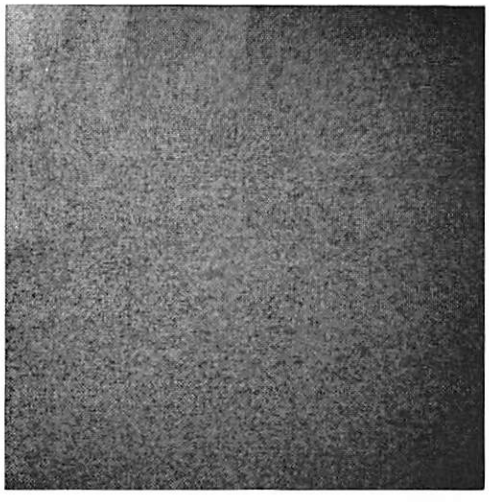
Figure 2-6. Fourier Transform Domain Display:
(a) GIRL; (b) COUPLE; (c) MOON SCENE.



(a)

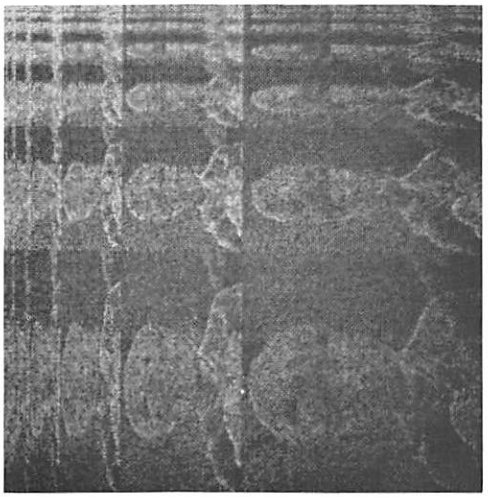


(b)

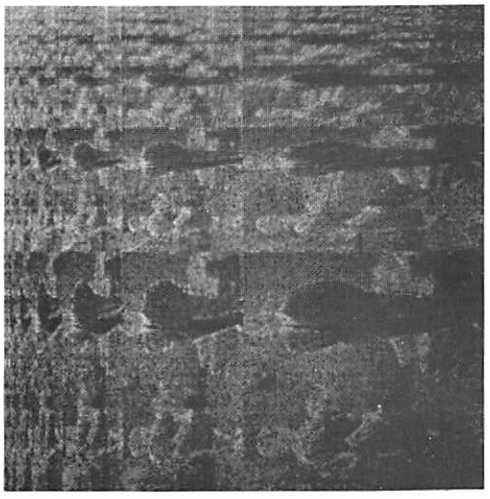


(c)

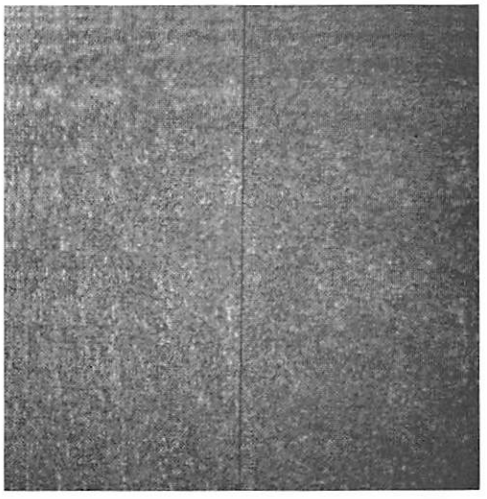
Figure 2-7. Hadamard Transform Domain Display: (a) GIRL; (b) COUPLE; (c) MOON SCENE.



(a)



(b)



(c)

Figure 2-8. Haar Transform Domain Display:
(a) GIRL; (b) COUPLE; (c) MOON SCENE.

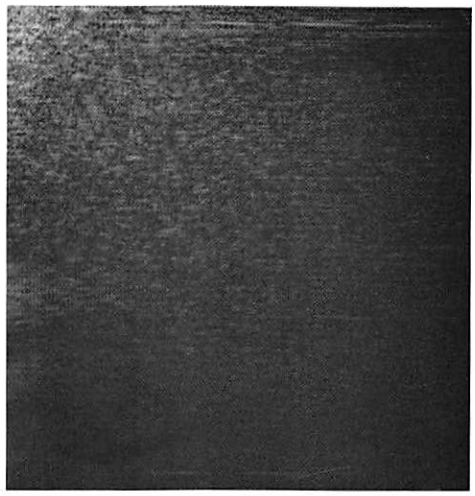


Figure 2-9. Karhunen-Loeve Transform Domain Display of GIRL Image.

3. DEVELOPMENT OF THE SLANT TRANSFORM

This chapter presents a detailed construction of the slant transform matrix of order 2^n (n is a positive integer). A fast computational algorithm based on the matrix decomposition is also presented. A computational flowchart of order four and eight which indicate steps of additions and multiplications needed for the fast transformation is also introduced. The chapter is summarized by introducing a class of slant matrices which are constructed by a direct product of the various orders of slant and Hadamard matrices.

3.1 Definition of the Slant Transform

The slant transform is based on the post multiplication of image lines by a slant matrix which is defined as an orthonormal matrix consisting of a constant basis vector, a slant basis vector, and possessing the sequency property. Let $[f_i]$ be a column vector of an image line composed of N pixels and $[\mathcal{S}]$ be the slant matrix of size N by N. Then the slant transform of this image line is

$$[F_i] = [\mathcal{S}][f_i] \tag{3-1}$$

Since the slant matrix is real and orthonormal, a two dimensional slant transform can be written as

$$[F] = [\mathcal{S}][f][\mathcal{S}]^T \tag{3-2}$$

and the inverse transform is

$$[f] = [\mathscr{s}]^T [f] [\mathscr{s}] \tag{3-3}$$

where [f] denotes image pixels in an N×N matrix, $[\mathscr{s}]$ denotes the slant transform kernel in an N×N matrix, and [F] denotes the two dimensional slant transform of [f].

3.2 The Construction of the Slant Transform Matrix

The slant transform matrix of order two consisting of a constant and slant basis vector is simply

$$S_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{3-4}$$

The slant transform matrix of order four can be written as

$$S_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ a_4+b_4 & a_4-b_4 & -a_4+b_4 & -a_4-b_4 \\ 1 & -1 & -1 & 1 \\ a_4-b_4 & -a_4-b_4 & a_4+b_4 & -a_4+b_4 \end{bmatrix} \tag{3-5}$$

where a_4 and b_4 are real constants to be determined by the conditions that S_4 must be orthogonal and that the step size of the slant basis vector must be the same throughout its length. The step sizes between adjacent elements of the slant vector are $2b_4$, $2a_4-2b_4$, and $2b_4$. By setting these step sizes equal, there results

$$a_4 = 2b_4 \tag{3-6}$$

Hence, the slant matrix of order 4 can be rewritten as

$$S_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3b_4 & b_4 & -b_4 & -3b_4 \\ 1 & -1 & -1 & 1 \\ b_4 & -3b_4 & 3b_4 & -b_4 \end{bmatrix} \tag{3-7}$$

By the orthonormality condition $[S_4][S_4]^T = [I]$, it is found that

$$b_4 = \frac{1}{\sqrt{5}} \tag{3-8}$$

Substituting this into equation (3-6) one obtains

$$a_4 = \frac{2}{\sqrt{5}} \tag{3-9}$$

Thus, the slant matrix of order 4 becomes

$$S_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{-1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} \\ 1 & -1 & -1 & 1 \\ \frac{1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} & \frac{3}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{bmatrix} \tag{3-10}$$

It is easily seen that S_4 is orthonormal. Furthermore, S_4 possesses the sequency property; each row has an increasing number of sign reversals from 0 to 3.

An extension of the slant matrix to its next size of order 8 is given by

$$S_8 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ a_8 & b_8 & 0 & 0 & -a_8 & b_8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ -b_8 & a_8 & 0 & 0 & b_8 & a_8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} S_4 & 0 \\ 0 & S_4 \end{bmatrix} \qquad (3\text{-}11)$$

where a_8 and b_8 are constants to be determined to satisfy the slant, orthonormality, and sequency properties. It can be seen that in S_8, the slant vector is obtained by a simple scaling operation on S_4. The remaining terms in equation (3-11) are introduced to obtain the sequency and orthogonality properties.

Equation (3-11) can be generalized to give the slant matrix of order $N (N = 2^n,\ n = 3, 4, 5, \ldots)$ in terms of the slant matrix of order $\frac{N}{2}$ by the following recursive relation:

$$S_N = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & & 0 & 1 & 0 & & 0 \\ a_N & b_N & & & -a_N & b_N & & \\ & & I & & & & I & \\ 0 & & & I & 0 & & & I \\ 0 & 1 & & 0 & 0 & -1 & & 0 \\ -b_N & a_N & & & b_N & a_N & & \\ & & I & & & & I & \\ 0 & & & I & 0 & & & I \end{bmatrix} \begin{bmatrix} S_{\frac{N}{2}} & 0 \\ 0 & S_{\frac{N}{2}} \end{bmatrix} \qquad (3\text{-}12)$$

where I represents a 2x2 identity matrix.

A 16x16 slant matrix computed by this recursive relation is as follows:

$$S_{16} = \frac{1}{\sqrt{16}} \begin{bmatrix}
 & (1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1) \\
\frac{15}{\sqrt{85}} & (1 & \frac{13}{15} & \frac{11}{15} & \frac{9}{15} & \frac{7}{15} & \frac{5}{15} & \frac{3}{15} & \frac{1}{15} & \frac{-1}{15} & \frac{-3}{15} & \frac{-5}{15} & \frac{-7}{15} & \frac{-9}{15} & \frac{-11}{15} & \frac{-13}{15} & -1) \\
\frac{7}{\sqrt{21}} & (1 & \frac{5}{7} & \frac{3}{7} & \frac{1}{7} & \frac{-1}{7} & \frac{-3}{7} & \frac{-5}{7} & -1 & -1 & \frac{-5}{7} & \frac{-3}{7} & \frac{-1}{7} & \frac{1}{7} & \frac{3}{7} & \frac{5}{7} & 1) \\
\frac{1}{\sqrt{21\times 85}} & (\frac{35}{77} & \frac{19}{77} & \frac{3}{77} & \frac{-13}{77} & \frac{-28}{77} & \frac{-45}{77} & \frac{-61}{77} & -1 & 1 & \frac{61}{77} & \frac{45}{77} & \frac{28}{77} & \frac{13}{77} & \frac{-3}{77} & \frac{-19}{77} & \frac{-35}{77}) \\
\frac{3}{\sqrt{5}} & (1 & \frac{1}{3} & \frac{-1}{3} & -1 & -1 & \frac{-1}{3} & \frac{1}{3} & 1 & 1 & \frac{1}{3} & \frac{-1}{3} & -1 & -1 & \frac{-1}{3} & \frac{1}{3} & 1) \\
\frac{3}{\sqrt{5}} & (1 & \frac{1}{3} & \frac{-1}{3} & -1 & -1 & \frac{-1}{3} & \frac{1}{3} & 1 & -1 & \frac{-1}{3} & \frac{1}{3} & 1 & 1 & \frac{1}{3} & \frac{-1}{3} & -1) \\
\frac{17}{\sqrt{105}} & (\frac{7}{17} & \frac{-1}{17} & \frac{-9}{17} & -1 & 1 & \frac{9}{17} & \frac{1}{17} & \frac{-7}{17} & \frac{-7}{17} & \frac{1}{17} & \frac{9}{17} & 1 & -1 & \frac{-8}{17} & \frac{-1}{17} & \frac{7}{17}) \\
\frac{17}{\sqrt{125}} & (\frac{7}{17} & \frac{-1}{17} & \frac{-9}{17} & -1 & 1 & \frac{9}{17} & \frac{1}{17} & \frac{-7}{17} & \frac{7}{17} & \frac{-1}{17} & \frac{-9}{17} & -1 & 1 & \frac{9}{17} & \frac{1}{17} & \frac{-9}{17}) \\
 & (1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1) \\
 & (1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1) \\
 & (1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1) \\
 & (1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1) \\
\frac{3}{\sqrt{5}} & (\frac{1}{3} & -1 & 1 & \frac{-1}{3} & \frac{-1}{3} & 1 & -1 & \frac{1}{3} & \frac{1}{3} & -1 & 1 & \frac{-1}{3} & \frac{-1}{3} & 1 & -1 & \frac{1}{3}) \\
\frac{3}{\sqrt{5}} & (\frac{1}{3} & -1 & 1 & \frac{-1}{3} & \frac{-1}{3} & 1 & -1 & \frac{1}{3} & \frac{-1}{3} & 1 & -1 & \frac{1}{3} & \frac{1}{3} & -1 & 1 & \frac{-1}{3}) \\
\frac{3}{\sqrt{5}} & (\frac{1}{3} & -1 & 1 & \frac{-1}{3} & \frac{1}{3} & -1 & 1 & \frac{-1}{3} & \frac{-1}{3} & 1 & -1 & \frac{1}{3} & \frac{-1}{3} & 1 & -1 & \frac{1}{3}) \\
\frac{3}{\sqrt{5}} & (\frac{1}{3} & -1 & 1 & \frac{-1}{3} & \frac{1}{3} & -1 & 1 & \frac{-1}{3} & \frac{1}{3} & -1 & 1 & \frac{-1}{3} & \frac{1}{3} & -1 & 1 & \frac{-1}{3})
\end{bmatrix} \tag{3-13}$$

3.3 Slant Transform Fast Computational Algorithm

The fast computational algorithm of the slant transform is based on the property of matrix decomposition. The 4x4 slant matrix can be reformed as

$$S_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & \frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 1 & -1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{5}} & -\frac{3}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \tag{3-14}$$

If S_4 is post multiplied by a column data vector, the first computational pass requires 4 additions, the second pass requires 4 multiplications (the elements $\frac{3}{\sqrt{5}}$ and $\frac{1}{\sqrt{5}}$) and 4 additions. The total computational requirements, without counting the normalization factor $\frac{1}{\sqrt{4}}$, are 8 adds and 4 multiplies*. Figure 3-1 contains a flowchart of the computational operations for S_4. The 8x8 slant matrix can be written as

$$S_8 = \frac{1}{\sqrt{2}} \left[\begin{array}{cccc:cccc} 1 & 0 & 0 & 0 & & & & \\ 0 & b_8 & a_8 & 0 & & & & \\ 0 & 0 & 0 & 1 & & 0 & & \\ 0 & a_8 & -b_8 & 0 & & & & \\ \hdashline & & & & 1 & 0 & 0 & 0 \\ & 0 & & & 0 & 1 & 0 & 0 \\ & & & & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{cccc:cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ \hdashline 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{cc:cc} & & & \\ & S_4 & & 0 \\ & & & \\ \hdashline & & & \\ & 0 & & S_4 \\ & & & \end{array}\right] \tag{3-15}$$

* The normalization factor will not be encountered here.

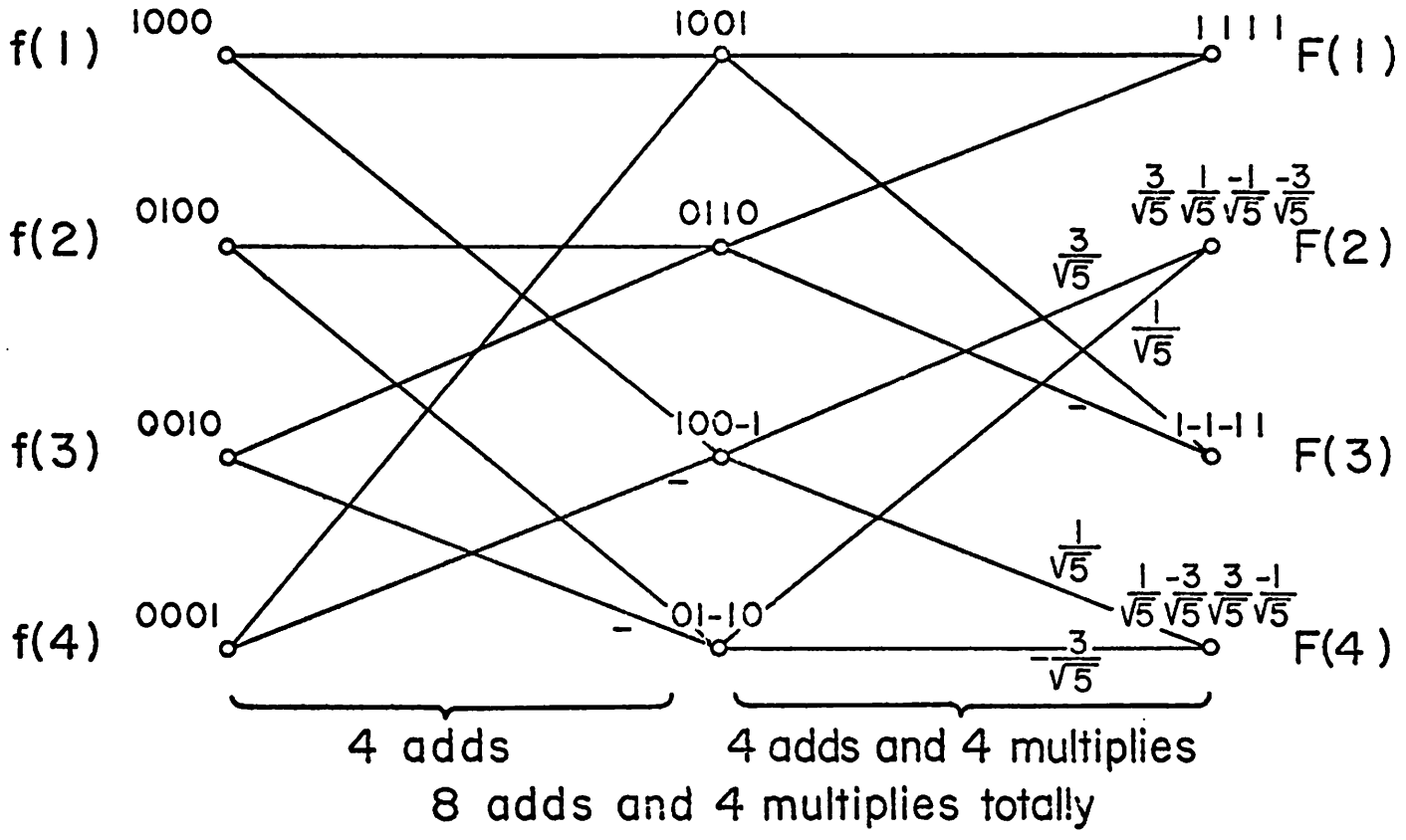


Figure 3-1. Slant Transform Computational Flowchart of Order 4.

If S_8 is post multiplied by a column data vector, the first and second computational passes require two times the total computations of S_4, and the third and fourth passes require eight additions and four multiplications (the element a_8 and b_8). The total computational requirements are 24 adds and 12 multiples. A flowchart of these computational operations is shown in Figure 3-2.

The decomposition of a generalized N×N slant matrix is shown in exhibit 3-1. If S_N is post multiplied by a column data vector, the total additions and multiplications can be computed by the recursive relation:

$$K_N = 2K_{\frac{N}{2}} + N \tag{3-16}$$

and

$$K'_N = 2K'_{\frac{N}{2}} + 4 \tag{3-17}$$

where K_N and K'_N are the number of additions and multiplications respectively. These terms can be rewritten as*

$$K_N = 2\left(2K_{\frac{N}{4}} + \frac{N}{2}\right) + N$$

$$= 4K_{\frac{N}{4}} + 2N$$

$$= 8K_{\frac{N}{8}} + 3N$$

$$= \cdots$$

* All logarithms are base two.

$$= \frac{N}{4} K_4 + (\log \frac{N}{4}) N$$

$$= \frac{N}{4} \cdot 8 + N \log \frac{N}{4}$$

$$= N(2 + \log \frac{N}{4})$$

$$= N \log N \qquad (3\text{-}18)$$

and

$$K'_N = 2 (2K'_{\frac{N}{4}} + 4) + 4$$

$$= 4 K'_{\frac{N}{4}} + 8 + 4$$

$$= 8 K'_{\frac{N}{8}} + 16 + 8 + 4$$

$$= \cdots$$

$$= \frac{N}{4} K'_4 + \frac{N}{2} + \frac{N}{4} + \cdots + 8 + 4$$

$$= N + \frac{N}{2} + \frac{N}{4} + \cdots + 8 + 4$$

$$= 2N-4 \qquad (3\text{-}19)$$

Therefore, the total operations of an N×N slant transform is N log N + 2N-4. For purpose of comparison an N×N Hadamard transform requires N log N operations. It is seen that the total number of operations of the slant transform is only slightly more then those of the Hadamard transform.

To determine the coefficients (a_N, b_N), one can proceed as

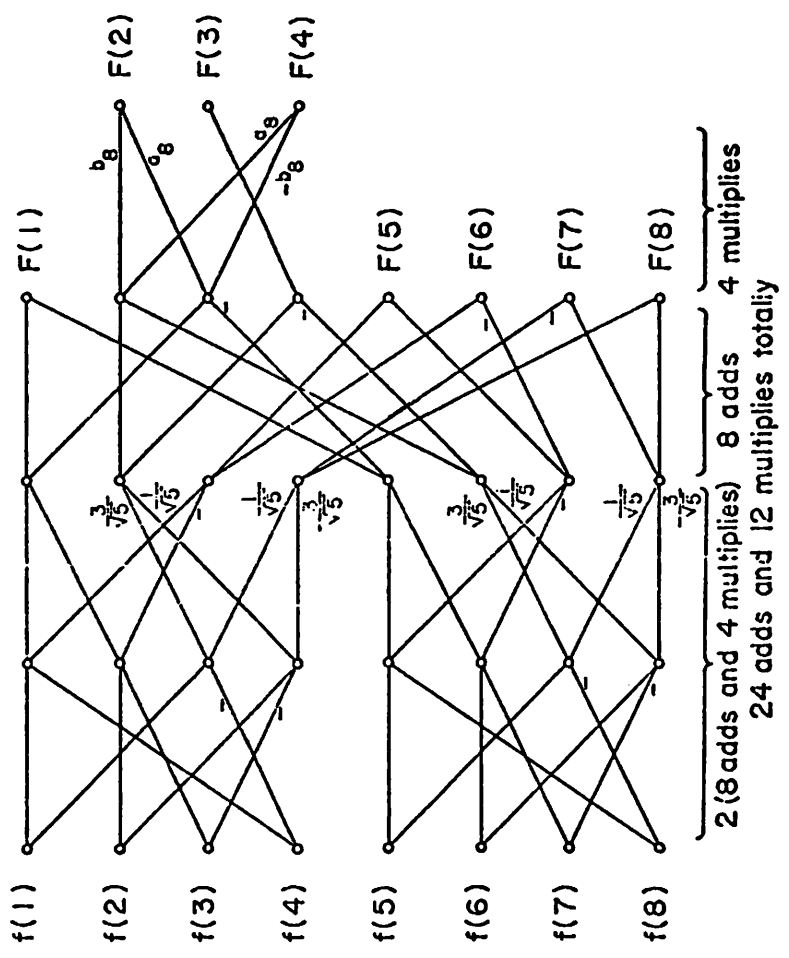


Figure 3-2. Slant Transform Computational Flowchart of Order 8.

$$S_N = \frac{1}{\sqrt{2}} \begin{bmatrix} \begin{array}{cccc|c} 1 & 0 & 0 & 0 & \\ 0 & b_N & a_N & 0 & 0 \\ 0 & 0 & 0 & 1 & \\ 0 & a_N & -b_N & 0 & \end{array} \\ \hline \begin{array}{c|c} & \\ 0 & I_{N-4} \\ & \end{array} \end{bmatrix} \begin{bmatrix} \begin{array}{cc|ccc|cc|ccc} 1 & 0 & & & & 1 & 0 & & & \\ 0 & 1 & & & & 0 & 1 & & & \\ 1 & 0 & & 0 & & -1 & 0 & & 0 & \\ 0 & 1 & & & & 0 & -1 & & & \end{array} \\ \hline \begin{array}{c|ccccc|c|ccccc} & 1 & 0 & 0 & \cdots & 0 & 0 & & 1 & 0 & 0 & \cdots & 0 & 0 \\ & 1 & 0 & 0 & \cdots & 0 & 0 & & -1 & 0 & 0 & \cdots & 0 & 0 \\ & 0 & 1 & 0 & \cdots & 0 & 0 & & 0 & -1 & 0 & \cdots & 0 & 0 \\ & 0 & 1 & 0 & \cdots & 0 & 0 & & 0 & 1 & 0 & \cdots & 0 & 0 \\ & 0 & 0 & 1 & \cdots & 0 & 0 & & 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 & -1 & \cdots & 0 & 0 \\ & \cdot & \cdot & \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ & \cdot & \cdot & \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ & \cdot & \cdot & \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ & 0 & 0 & 0 & \cdots & 0 & 1 & & 0 & 0 & 0 & \cdots & 0 & -1 \\ & 0 & 0 & 0 & \cdots & 0 & 1 & & 0 & 0 & 0 & \cdots & 0 & 1 \end{array} \end{bmatrix} \begin{bmatrix} \begin{array}{c|c} S_{\frac{N}{2}} & 0 \end{array} \\ \hline \begin{array}{c|c} 0 & S_{\frac{N}{2}} \end{array} \end{bmatrix}$$

Exhibit 3-1 The decomposition of a generalized N×N slant matrix.

follows: the first vector is a constant

$$S_N(1, i) = \frac{1}{\sqrt{N}} \tag{3-20}$$

The second vector (slant vector) is a linear function of the column index which is orthonormal to the first row. It must, therefore, be of the form

$$S_N(2, i) = X_N \cdot (N + 1 - 2i) \tag{3-21}$$

Now, by the recursion indicated in equation (3-12), one obtains

$$S_N(2, i) = \frac{1}{\sqrt{2}} a_N S_{\frac{N}{2}}(1, i) + \frac{1}{\sqrt{2}} b_N S_{\frac{N}{2}}(2, i) \tag{3-22}$$

or

$$X_N \cdot (N + 1 - 2i) = \frac{1}{\sqrt{N}} a_N + \frac{1}{\sqrt{2}} b_N X_{\frac{N}{2}} \cdot (\frac{N}{2} + 1 - 2i) \tag{3-23}$$

From this it is found that

$$X_N = \frac{1}{\sqrt{2}} b_N X_{\frac{N}{2}} \tag{3-24a}$$

$$a_N = \sqrt{2} \, (\frac{N}{2})^{\frac{3}{2}} X_N \tag{3-24b}$$

and by induction

$$a_N = 2 b_N a_{\frac{N}{2}} \tag{3-25}$$

Since $S_{\frac{N}{2}}(1, .)$ and $S_{\frac{N}{2}}(2, .)$ are orthonormal vectors in $\frac{N}{2}$ dimensions and $S_N(2, .)$ is a unit vector in N dimensions, the above recursion

implies

$$\| S_N(2,.) \| = a_N^2 + b_N^2 = 1 \tag{3-26}$$

These two relations can be used to obtain the coefficients, (a_N, b_N), recursively:

$$a_2 = 1$$

$$b_N = 1 \Big/ \sqrt{1 + 4\, a_{\frac{N}{2}}^2}$$

$$a_N = 2 b_N\, a_{\frac{N}{2}} \tag{3-27}$$

For the inverse transform the computational operations are similar to the forward transform described above. The decomposition of a generalized N×N inverse slant matrix is shown in exhibit 3-2. If S_N is post multiplied by a column vector it can easily be seen that the total number of operations is exactly the same as for the forward transform. Again the coefficients, (a_N, b_N), can be computed by equation (3-27).

3.4 Summary

The slant transform matrix thus far discussed contains a slant basis vector decreasing in uniform steps over its entire vector length. There are a class of slant matrices that can be constructed with the slant basis vector decreasing or increasing in steps a number of times over its vector length. Taking a 16×16 matrix as an example,

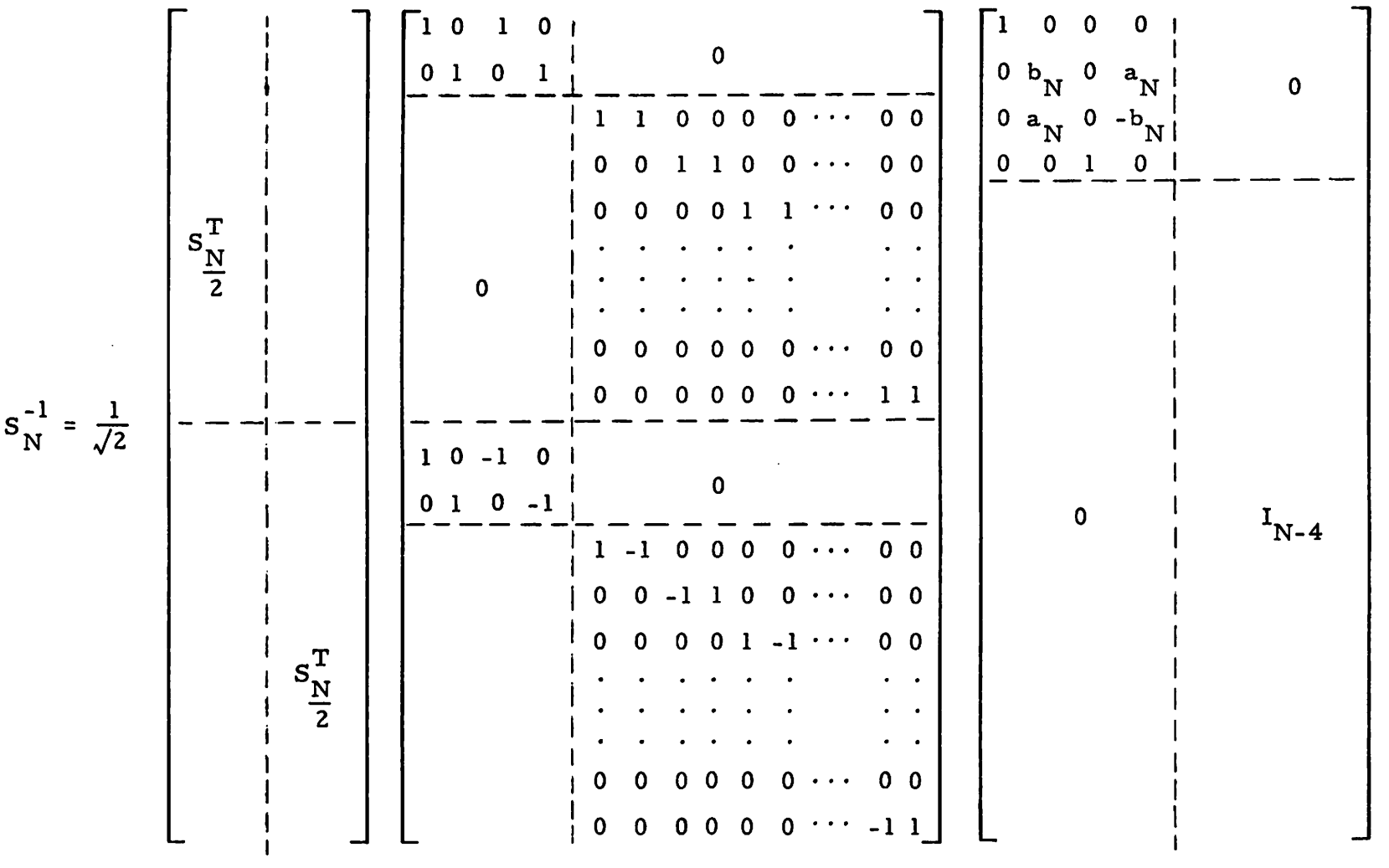


Exhibit 3-2 The decomposition of a generalized N×N inverse slant matrix.

there are two other type of matrices, in addition to the one shown on equation (3-13), which can be constructed by the direct product* of slant and Hadamard matrices, namely, $S_8 \times H_2$ and $S_4 \times H_4$. In $S_8 \times H_2$, the slant basis vector decreases in uniform steps to the middle point of its vector length and then increases in uniform steps to the rest of its vector length. In $S_4 \times H_4$, the slant basis vector decreases uniformly for the first quarter length, increases for the second quarter, decreases for the third quarter, and increases again for the final quarter. It can easily be shown that a matrix of order N×N has a class of $(\log_2 N-1)$ matrices each containing a "slant" basis vector. Figures 3-3, 3-4 and 3-5 show the class of slant matrices of order 16. It can be seen that all these matrices possess sequency and orthogonality properties. The performance capability of these matrices has not been investigated. From a purely mathematical point of view the

---

* A direct product of a matrix $A = a_{ij}$, $1 \leq i, j \leq m$, and a matrix $B = b_{ij}$, $1 \leq i, j \leq n$, is defined as

$$A \times B = \begin{bmatrix} Ab_{11} & Ab_{12} & \cdots & Ab_{1m} \\ Ab_{21} & Ab_{12} & \cdots & Ab_{2m} \\ \vdots & & & \\ Ab_{m1} & Ab_{m2} & \cdots & Ab_{mm} \end{bmatrix}$$

existence of a slant matrix of order other then 2^n, n=1, 2, 3, . . . , may also be an interesting topic to explore, though it will not be considered here for image coding.

The next two chapters contain a general analysis of the slant image transform.

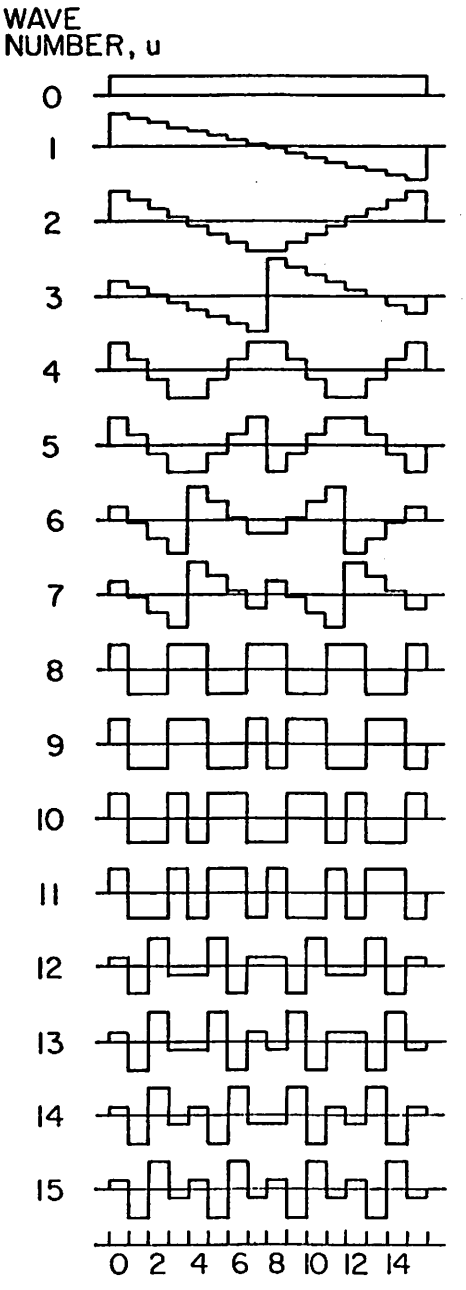


Figure 3-3. Waveforms of Slant Matrix S_{16}.

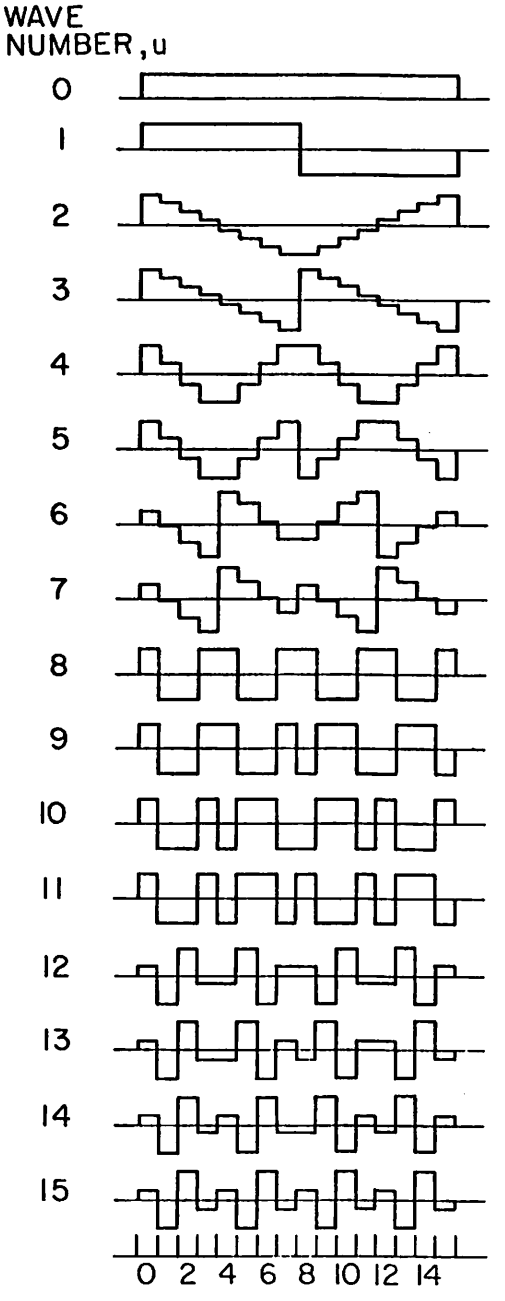


Figure 3-5. Waveforms of Slant Matrix $S_8 \times H_2$.

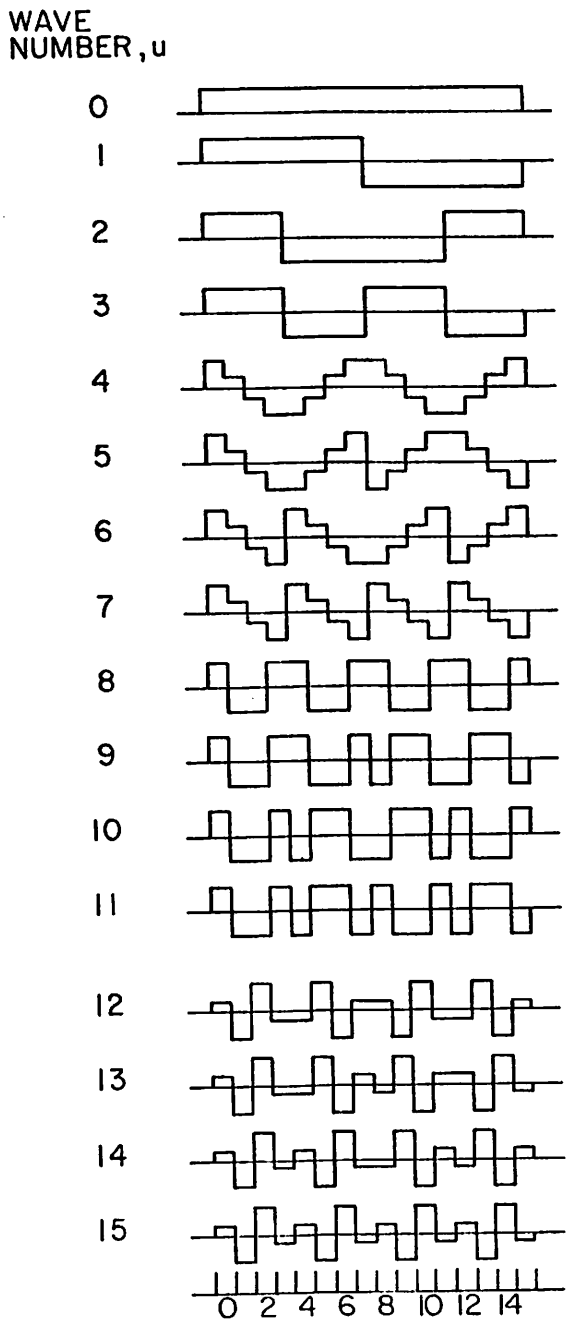


Figure 3-4. Waveforms of Slant Matrix $S_4 \times H_4$.

4. STATISTICAL ANALYSIS OF THE SLANT TRANSFORM

The development of efficient quantization and coding methods for slant transform samples requires an understanding of the statistical properties of the transform domain samples. This chapter presents a derivation of the statistical mean and variance of slant transform samples, and also the development of stochastic probability models for slant transform samples. The material introduced here is generally applicable not only for slant transform but for all transforms as well.

4.1 Statistical Mean and Variance

Suppose each sample of an original image, denoted by the function f(j, k) over spatial coordinates, is considered as a two dimensional stochastic process. The spatial mean

$$E\{f(j,k)\} \equiv \overline{f(j,k)} \tag{4-1}$$

and the covariance

$$E\{[f(j_1,k_1) - \overline{f(j_1,k_1)}][f(j_2,k_2) - \overline{f(j_2,k_2)}]\} \equiv C\{j_1,k_1,j_2,k_2\} \tag{4-2}$$

are assumed known or at least estimable. Then, for a generalized forward transform as shown in equation (2-1), the mean of the transform samples can be written as

$$E\{F(u,v)\} \equiv \overline{F(u,v)} = \sum_j \sum_k \overline{f(j,k)}\ a(j,k,u,v) \tag{4-3}$$

The covariance function of the transform samples is defined as

$$C\{u_1, u_2, v_1, v_2\} \equiv E\{[F(u_1, v_1) - \overline{F(u_1, v_1)}][F(u_2, v_2) - \overline{F(u_2, v_2)}]\} \tag{4-4}$$

Substitution of equation (2-1) and (4-3) into equation (4-4) gives

$$C\{u_1, u_2, v_1, v_2\} \equiv E\{\sum_{j_1}\sum_{k_1}[f(j_1, k_1) - \overline{f(j_1, k_1)}] \cdot a(j_1, k_1, u_1, v_1) \cdot \sum_{j_2}\sum_{k_2}[f(j_2, k_2) - \overline{f(j_2, k_2)}] \cdot a(j_2, k_2, u_2, v_2)\} \tag{4-5}$$

or

$$C\{u_1, u_2, v_1, v_2\} = \sum_{j_1}\sum_{j_2}\sum_{k_1}\sum_{k_2} E\{[f(j_1, k_1) - \overline{f(j_1, k_1)}][f(j_2, k_2) - \overline{f(j_2, k_2)}]\} \cdot a(j_1, k_1, u_1, v_1)\, a(j_2, k_2, u_2, v_2) \tag{4-6}$$

Upon substitution of equation (4-2), the result is

$$C(u_1, u_2, v_1, v_2) = \sum_{j_1}\sum_{j_2}\sum_{k_1}\sum_{k_2} C(j_1, j_2, k_1, k_2)\, a(j_1, k_1, u_1, v_1)\, a(j_2, k_2, u_2, v_2) \tag{4-7}$$

The generalized expression for the variance of transform domain samples is thus

$$\sigma^2(u, v) = C(u, u, v, v) = \sum_{j_1}\sum_{j_2}\sum_{k_1}\sum_{k_2} C(j_1, j_2, k_1, k_2)\, a(j_1, k_1, u, v)\, a(j_2, k_2, u, v) \tag{4-8}$$

If the covariance matrix of the original image is separable in j and k direction and if the transform kernel is separable, then the transform domain variance can be computed as

$$\sigma^2(u, v) = \sigma^2(u)\,\sigma^2(v) \tag{4-9}$$

Since the transform kernel of the slant transform in series form is mathematically difficult to describe, it is desirable to introduce an alternate matrix formulation.

A matrix formulation of a separable two dimensional slant transform can be written as

$$[F] = [S][f][S]^T \tag{4-10}$$

where [f] denotes an image pixel in N×N matrix, [S] denotes slant transform kernel in N×N matrix, and [F] denotes the two dimensional slant transform of [f]. Let $[f_j]$ and $[f_k]$ be column vectors representing the rows and columns of [f], respectively. Then the covariance matrices of rows and columns are

$$\left[C_{f_j}\right] = E\{[f_j - \bar{f}_j][f_j - \bar{f}_j]^{*T}\} \tag{4-11}$$

and

$$\left[C_{f_k}\right] = E\{[f_k - \bar{f}_k][f_k - \bar{f}_k]^{*T}\} \tag{4-12}$$

Now define

$$[F_j] = [S][f_j] \tag{4-13}$$

and

$$[F_k] = [S][f_k] \tag{4-14}$$

Then the covariance matrices of $[F_j]$ and $[F_k]$ are respectively

$$\left[C_{F_j}\right] = E\{[F_j - \overline{F}_j][F_j - \overline{F}_j]^{*T}\} \tag{4-15}$$

and

$$\left[C_{F_k}\right] = E\{[F_k - \overline{F}_k][F_k - \overline{F}_k]^{*T}\} \tag{4-16}$$

substituting equation (4-13) into equation (4-15) gives

$$\left[C_{F_j}\right] = E\{([S][f_j] - [S][\overline{f}_j])([S][f_j] - [S][\overline{f}_j])^{*T}\}$$

$$= E\{[S][f_j - \overline{f}_j][f_j - \overline{f}_j]^{*T}[S]^{*T}\}$$

or

$$\left[C_{F_j}\right] = [S]\left[C_{f_j}\right][S]^{*T} \tag{4-17}$$

Similarly, equation (4-16) can be reduced to

$$\left[C_{F_k}\right] = [S]\left[C_{f_k}\right][S]^{*T} \tag{4-18}$$

The variances of the slant transform samples are, therefore,

$$[V_F] = \left[V_{F_j}\right]\left[V_{F_k}\right]^{*T} \tag{4-19}$$

where $\left[V_{F_j}\right]^T = \left[C_{F_j}(0,0),\ C_{F_j}(1,1), \ldots, C_{F_j}(N-1, N-1)\right]$

$\left[V_{F_k}\right]^T = \left[C_{F_k}(0,0),\ C_{F_k}(1,1), \ldots, C_{F_k}(N-1, N-1)\right]$

4.2 Probability Densities

The probability density of slant transform samples is generally very difficult to obtain since the probability density of the original image is not usually well defined, and also, the slant transform

representation is mathematically complex. Nevertheless, since the transform operation forms a weighted sum over all of the pixels in the original image, one can evoke qualitative arguments based upon the central-limit theorem [15-17] to determine probability densities of transform domain samples.

A two dimensional slant transform in matrix form, shown in equation (4-10), can be rewritten as

$$F(0,0) = [S_0][f][S_0]^T \tag{4-20a}$$

and

$$F(u,v) = [S_i][f][S_j]^T;\ u,v = 0,1,2,\ldots,N-1;\ (u,v) \neq (0,0) \tag{4-20b}$$

where $[S_k]$ is kth vector of slant matrix $[S]$, i.e., $[S_k] = [S(k,0)\ S(k,1)\cdots S(k,N-1)]$. From equation (4-20a), since all elements of the vector $[S_0]$ are positive constants and all pixels of $[f]$ are non-negative, the value of $F(0,0)$ is always non-negative. The probability density of $F(0,0)$ will resemble the histogram of $f(j,k)$ and generally can be represented by a Rayleigh density, i.e.,

$$p_{F(0,0)}(x) = \frac{x}{\alpha^2} \exp(-x^2/2\alpha^2)\ U(x) \tag{4-21}$$

The probability densities of the samples of $F(u,v)$, for $(u,v) \neq (0,0)$, are generally indeterminable. By examining equation (4-20b), since half the elements of $[S_i]$ $(i \neq 0)$ are positive and half are negative, and also the magnitude of these elements are periodically

defined, therefore, the only information that one can obtain is that the probability densities of $F(u, v)$, $(u, v) \neq (0, 0)$, are expected to be symmetrical about the origin. Based upon this information two different types of probability models may be defined, namely, the Gaussian and two-sided exponential models:

For the Gaussian model, the probability density can be written as

$$p_{F(u,v)}(x) = \frac{1}{\sqrt{2\pi}\,\sigma(u, v)} \exp(-x^2/2\sigma^2(u, v))$$
$$u, v = 0, 1, \ldots, N-1;\ (u, v) \neq (0, 0) \tag{4-22}$$

and for the two-sided exponential model, the probability density can be written as

$$p_{F(u,v)}(x) = \frac{\alpha(u, v)}{2} \exp(-\alpha(u, v)x)$$
$$u, v=0, 1, \ldots, N-1;\ (u, v) \neq (0, 0) \tag{4-23}$$

Since the magnitude of $F(0, 0)$ is always non-negative and the magnitudes of the remaining transform samples fluctuate about the origin, they will be termed respectively the dc and ac transform domain samples.

Suppose that the ac samples, as defined in equation (4-20b), are independent and identically distributed, then from the central-limit theorem the probability density of $\sum_{u} \sum_{v} F(u, v)$ $_{(u,v)\neq(0,0)}$ will tend to be Gaussian with mean zero and variance $\sigma = \sum_{u} \sum_{v} \sigma(u, v)$ $_{(u,v)\neq(0,0)}$, i.e.,

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2}\right] \tag{4-24}$$

where X denotes the summation of the ac random variables, F(u, v), $(u, v) \neq (0, 0)$.

5. SLANT TRANSFORM MONOCHROME IMAGE CODING

Figure 5-1 contains a block diagram of the slant transform coding system for monochrome images. In operation a two dimensional slant transform is taken of the image pixels over the entire image or repeatedly over subsections of the image called blocks. The transform domain samples are then operated upon by a sample selector that decides which samples are to be transmitted. Those selected samples are then quantized, coded, and transmitted over a channel. At the receiver the received data is decoded and an inverse slant transformation is performed to reconstruct the original image.

The basic premise of an image transform coding system is that the two dimensional transform of an image has an energy distribution more amenable to coding than the spatial domain representation. As a result of the inherent pixel to pixel correlation of natural images, the energy in the transform domain tends to be clustered into a relatively few number of transform samples. The slant transform of an image has exactly the same property which can be exploited to achieve a bandwidth reduction.

Figure 5-2(a), (b), and (c) shows the full size two dimensional slant transform displays of the three original monochrome images shown in Figure 1-1. The logarithm of the absolute value of each

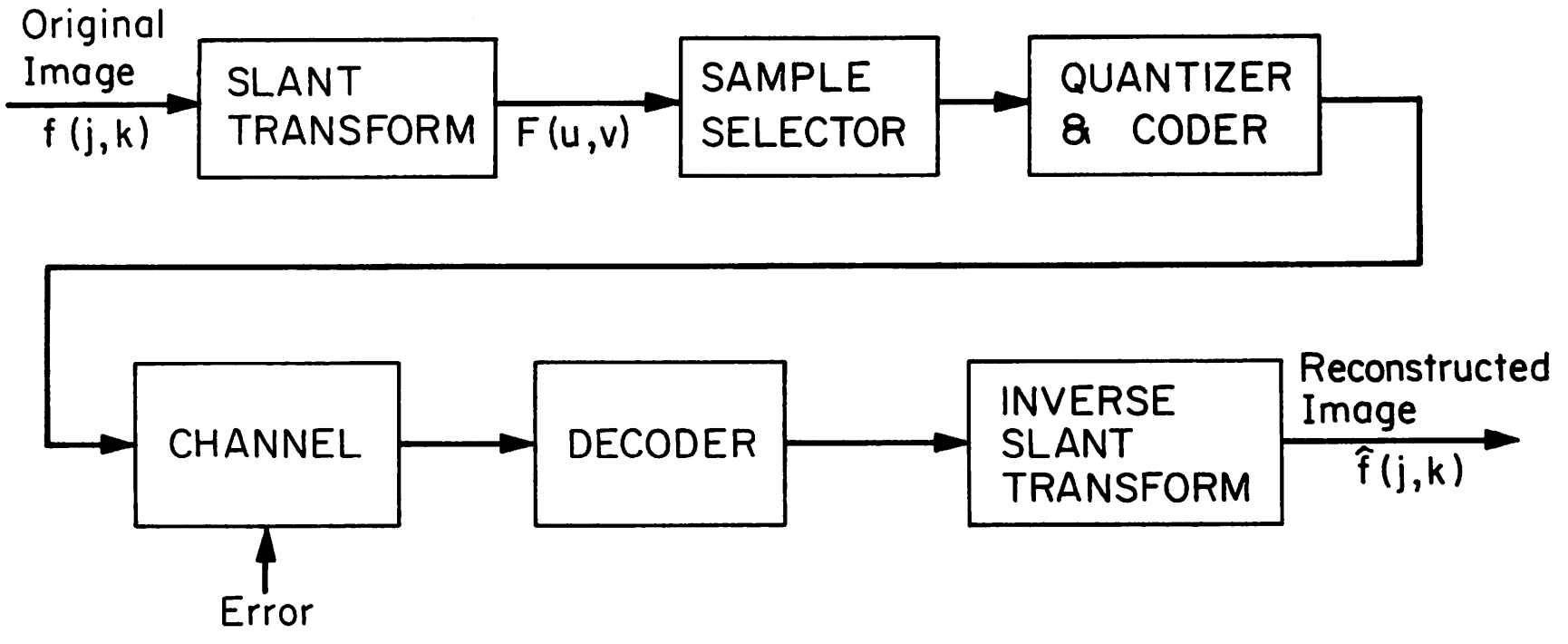


Figure 5-1. Slant Transform Monochrome Image Coding System.

transform sample is displayed rather than the absolute value itself in order to reduce the dynamic range of transform samples. Figure 5-2(d) and 5-2(e) show two different views of the threshold display of Figure 5-2(a) where all samples whose magnitude is below a specified threshold are set to zero and all samples whose magnitude is above the threshold remain unchanged. A typical energy distribution of the slant transform canbe seen from these pictures. A high degree of energy compaction is seen on the upper-left hand corner of each picture which is the low frequency zone of the transform samples.

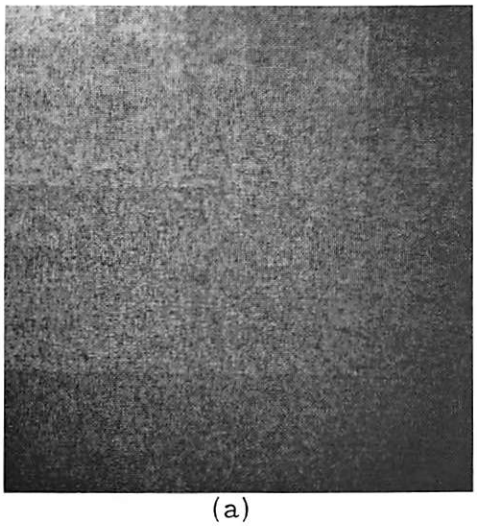
There are two bandwidth reduction factors that are often stated as image coding performance measures [23]:

$$\text{Sample reduction factor} = \frac{\text{number of original image samples}}{\text{number of samples selected to be coded and transmitted}} \tag{5-1}$$

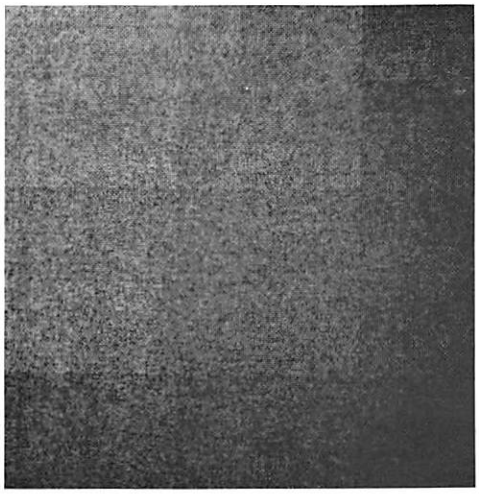
$$\text{Bit reduction factor} = \frac{\text{number of original image code bits}}{\text{number of selected sample code bits}} \tag{5-2}$$

The sample and bit reduction factors are identical if the same number of bits are assigned to both original and coded image samples.

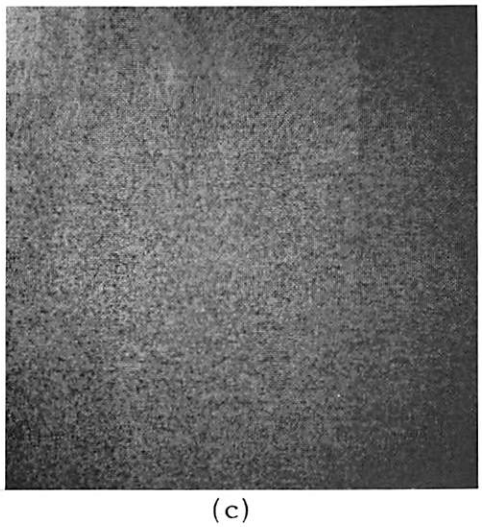
A bandwidth reduction can be achieved with the transform coding system in two basic ways: threshold sampling and zonal sampling [14]. In threshold sampling the image reconstruction is made with a subset of the samples which are larger than a specified threshold. In zonal sampling the reconstruction is made with a subset



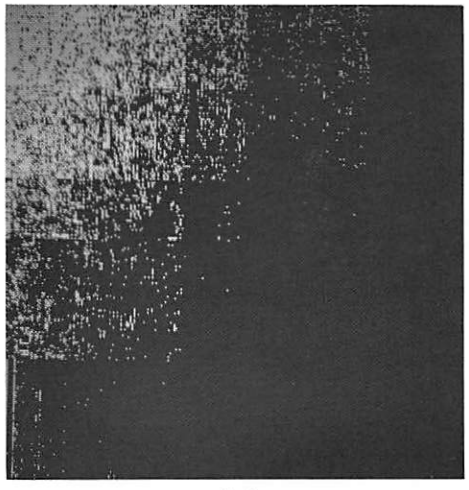
(a)



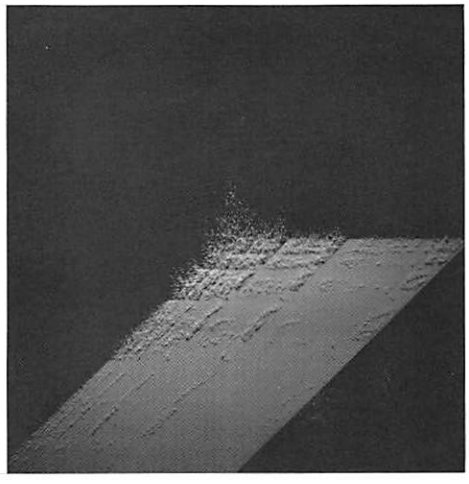
(b)



(c)



(d)



(e)

Figure 5-2. Slant Transform Domain Display: (a) GIRL; (b) COUPLE; (c) MOON SCENE; (d) Transform Threshold GIRL; (e) Perspective View of Transform Threshold GIRL.

of samples which lie in a certain geometrical zone – usually the lower frequency samples. This chapter will present the performance of the slant transform for threshold and zonal samplings in an error free channel. A derivation of the quantization and coding scheme as well as a derivation of the mean square error between the original image and its transform coded reconstruction will also be presented. For threshold sampling a runlength coding technique is used and for zonal sampling a bit allocation matrix based on the "maximum variance matrix" and rate distortion theory is introduced. Again the method introduced is generally applicable for all transforms.

5.1 Slant Transform Bandwidth Reduction

The slant transform sample reduction technique can be analyzed by defining a two dimensional sample selector as a function S(u, v) which takes on the value zero or one according to an apriori or adaptive rule. The reconstructed image, with those selected transform samples unquantized and uncoded, is then given by

$$f_s(j,k) = \sum_u \sum_v F(u,v)\, S(u,v)\, b(j,k,u,v) \tag{5-3}$$

where b(j, k, u, v) is the inverse transform Kernel.

The mean square error between this reconstructed image and original is, therefore

$$\epsilon = \frac{1}{N^2} \sum_j \sum_k E\{[f(j,k) - f_s(j,k)]^2\} \tag{5-4}$$

Substituting equations (2-2) and (5-3) into (5-4) gives

$$\epsilon = \frac{1}{N^2} \sum_j \sum_k E\{[\sum_u \sum_v F(u, v)\,(1-S(u, v))\, b(j, k, u, v)]^2\} \tag{5-5}$$

Expanding the series and changing the order of operations gives

$$\epsilon = \frac{1}{N^2} \sum_u \sum_{u'} \sum_v \sum_{v'} E\{F(u, v)\,[1-S(u, v)]\, F(u', v')\,[1-S(u', v')] \cdot \sum_j \sum_k b(j, k, u, v)\, b(j, k, u', v')\} \tag{5-6}$$

By the orthonormality of the slant transform Kernels the second set of summations is non-zero only if u=u' and v=v'. Thus

$$\epsilon = \frac{1}{N^2} \sum_u \sum_{u'} \sum_v \sum_{v'} E\{F(u, v)\,[1-S(u, v)]\, F(u', v')\,[1-S(u', v')] \cdot \delta(u-u', v-v')\} \tag{5-7}$$

or

$$\epsilon = \frac{1}{N^2} \sum_u \sum_v E\{[F(u, v)\,(1-S(u, v))]^2\} \tag{5-8}$$

Equation (5-8) shows that the mean square error between the original image and its sample reduced reconstruction may be computed from the transform domain samples.

5.1.1 Threshold Sampling

In threshold sampling the selection of transform domain samples is made after the transform has been taken on a particular image. A threshold is established apriori or adaptively, and only those samples whose magnitudes are greater than the threshold are selected to be quantized, coded, and transmitted over a channel.

Figure 5-3 contains a plot of the percentage of transform

domain samples lying below a magnitude threshold level for the slant transform of the three original monochrome images. The energy compaction for the GIRL image for various transforms is illustrated in Figure 5-4. In the figure it is seen that the energy compaction of the slant transform is exceeded only by the Karhunen-Loeve transform.

Figure 5-5 shows slant transform threshold sampling processed images of the original pictures. A two dimensional slant transform was taken of the image pixels repeatedly over subsections of an image in 16x16 pixel blocks. A threshold was assigned to make the sample reduction and then these reduced samples were inversely transformed immediately without any quantization and/or coding. The purpose of these pictures is mainly to illustrate the threshold sampling effect of the slant transform. It can be seen that the slant transform threshold sampling provides good quality reconstructions for sample reduction factors up to 12:1. For purposes of comparison Figures 5-6 and 5-7 contain threshold sampling processed GIRL pictures for the Hadamard, Haar, Fourier, and Karhunen-Loeve transforms. The quality rating of these transforms, from the standpoint of subjective quality, is the Karhunen-Loeve first, followed by the slant, Haar, Hadamard, and Fourier transforms. The orders of these ratings are exactly the same as the orders of those curves shown in Figure 5-4. It is rather interesting that the performance

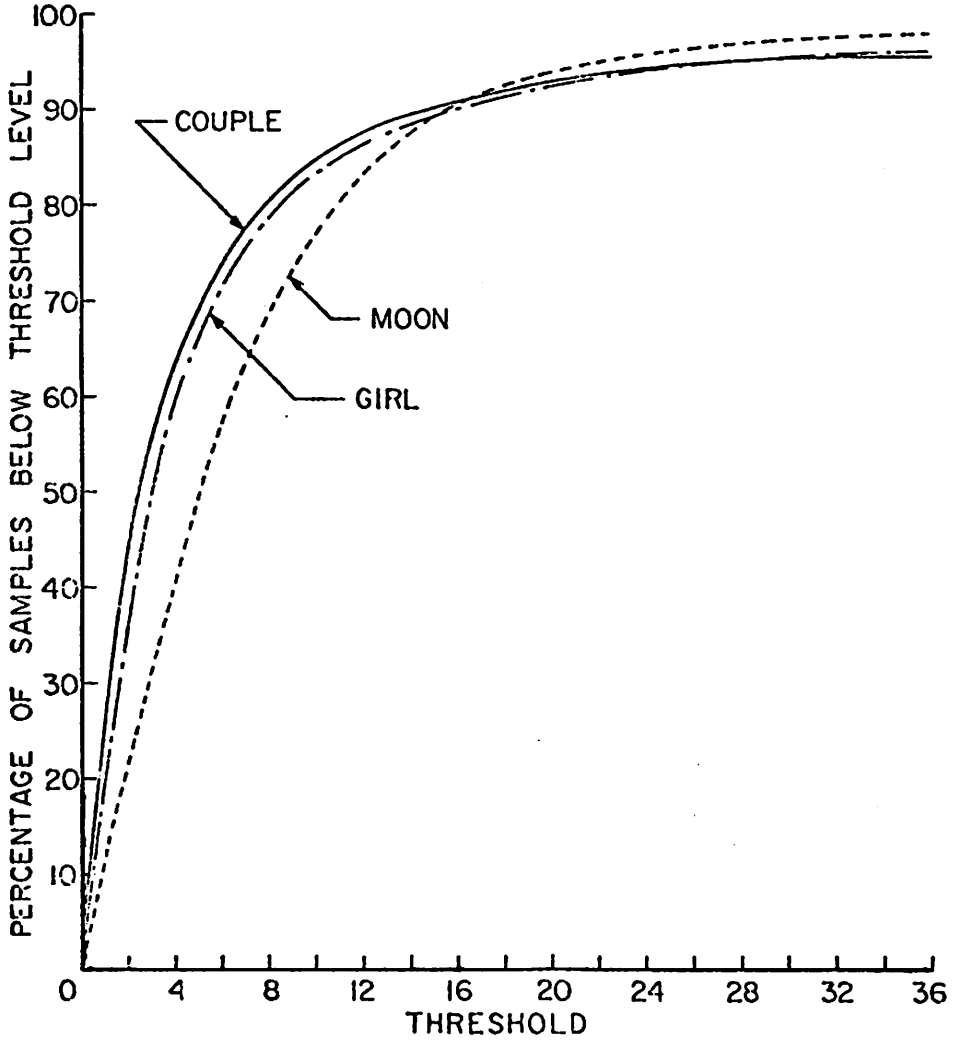


Figure 5-3. Number of Slant Transform Samples Below Threshold Versus Threshold Level (Transform is Performed in 16 x 16 Pixel Blocks).

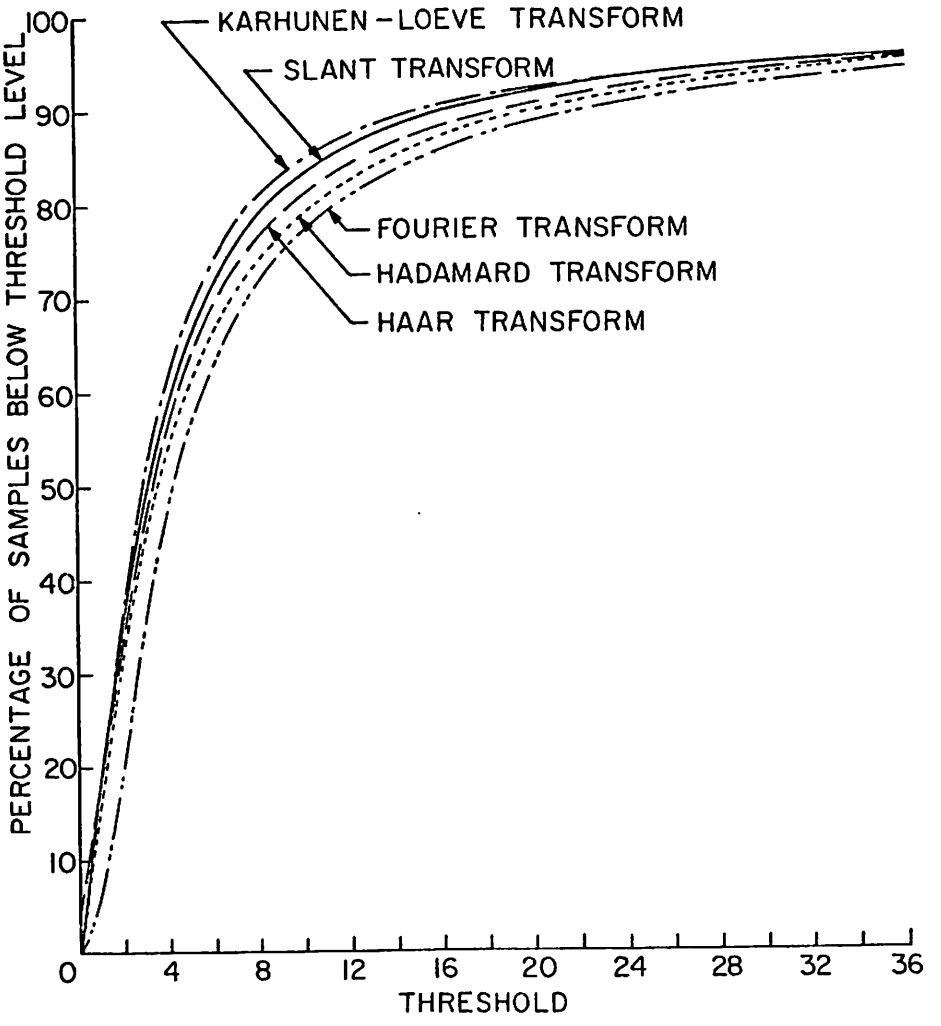


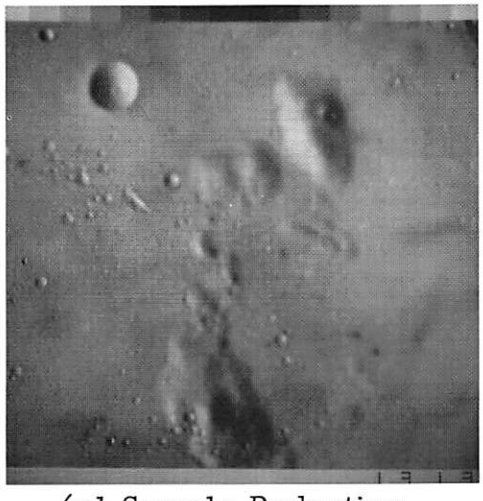
Figure 5-4. Number of Transform Samples of GIRL Below Threshold Versus Threshold Level (Transform is Performed in 16 x 16 Pixel Blocks).

6:1 Sample Reduction

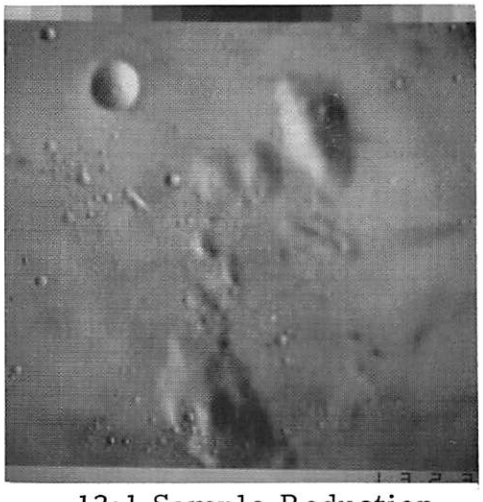
12:1 Sample Reduction

6:1 Sample Reduction

12:1 Sample Reduction



6:1 Sample Reduction



12:1 Sample Reduction

Figure 5-5. Slant Transform Threshold Sampling in 16 x 16 Pixel Blocks, Unquantized Transform.

Hadamard Transform
6:1 Sample Reduction

Hadamard Transform
12:1 Sample Reduction

Haar Transform
6:1 Sample Reduction

Haar Transform
12:1 Sample Reduction

Figure 5-6. Hadamard and Haar Transform Threshold Sampling in 16 x 16 Pixel Blocks, Unquantized Transform.

Fourier Transform
6:1 Sample Reduction

Fourier Transform
12:1 Sample Reduction

Karhunen-Loeve Transform
6:1 Sample Reduction

Karhunen-Loeve Transform
12:1 Sample Reduction

Figure 5-7. Fourier and Karhunen-Loeve Transform Threshold Sampling in 16 x 16 Pixel Blocks, Unquantized Transform.

of the Haar transform, with the fastest transform algorithm, is better than either the Hadamard or Fourier transform. Table 5-1 exhibits the mean square errors between the original GIRL and its threshold sampling reconstructions for the slant, Haar, Hadamard, Fourier, and Karhunen-Loeve transforms with various sample reductions. As expected, the Karhunen-Loeve transform has the best mean square error followed closely by the slant transform.

5.1.2 Zonal Sampling

In zonal sampling, rather than selecting those larger magnitude samples, a specific zone in the transform domain is established. Those samples lying inside the zone are selected and transmitted over a channel. Since the slant transform compacts energy over the upper-left hand corner of the transform domain, it is possible to design a special zone in this corner which covers most of the larger samples.

There are a number of zones that could be employed in the transform domain samples:

(1) rectangular zone

$$\begin{aligned} S(u, v) &= 1 \quad \text{for } u \leq u_c,\ v \leq v_c \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{5-9}$$

(2) elliptical zone

$$\begin{aligned} S(u, v) &= 1 \quad \text{for } \frac{u^2}{u_c^2} + \frac{v^2}{v_c^2} \leq 1 \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{5-10}$$

Transform / Sample Reduction	Slant	Hadamard	Haar	Fourier	Karhunen-Loeve
4:1	0.164%	0.223%	0.192%	0.310%	0.144%
6:1	0.297%	0.429%	0.354%	0.547%	0.261%
12:1	0.665%	1.020%	0.851%	1.250%	0.571%

Table 5-1 Mean Square Error Between the Original GIRL and Its Threshold Sampling Reconstructions (Transforms are performed in 16 x 16 pixel blocks).

$$\text{M. S. E.} = \sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\left[f(j,k)-\hat{f}(j,k)\right]^2 \Big/ \sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\left[f(j,k)\right]^2$$

(3) triangular zone

$$S(u,v) = 1 \quad \text{for } (u+v) \leq K \qquad (5\text{-}11)$$

$$= 0 \quad \text{otherwise}$$

where $S(u,v)$ denotes a sample selecting function, and u_c, v_c, and K are constants. It could easily be seen from figure 5-2(d) and 5-2(b) that for a full size slant transform none of the zones listed closely resembles the energy distribution of transform samples. Nevertheless, the degradation is not too serious if a rectangular zone is used for the slant transform of very high sample reductions (about 20:1), and a circular or a triangular zone for medium sample reductions (about 8:1).

There is a special zone termed "maximum variance zone" which is optimum for a mean square error criterion. Consider the pixels along a row of an image as denoted by a sequence of random variables $f_j(0), f_j(1), \ldots, f_j(N-1)$. Then the autocovariance function may be expressed as a covariance matrix of the form

$$\left[C_{f_j}\right] =$$

$$\begin{bmatrix} E\{[f_j(0)-\overline{f_j(0)}]^2\} & E\{[f_j(0)-\overline{f_j(0)}][f_j(1)-\overline{f_j(1)}]\} & \cdots \\ E\{[f_j(1)-\overline{f_j(1)}][f_j(0)-\overline{f_j(0)}]\} & E\{[f_j(1)-\overline{f_j(1)}]^2 & \cdots \\ \vdots & \vdots & \\ E\{[f_j(N-1)-\overline{f_j(N-1)}][f_j(0)-\overline{f_j(0)}]\} & E\{[f_j(N-1)-\overline{f_j(N-1)}][f_j(1)-\overline{f_j(1)}]\} & \cdots \end{bmatrix}$$

(5-12)

If the image is considered a sample of a Markov process with a correlation coefficient of ρ $(0 < \rho < 1)$ between each adjacent pixels and self correlation coefficient of unity, then equation (5-12) reduces to

$$\left[C_{f_j}\right] = \sigma_j^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & 1 & \rho & \cdots & \rho^{N-2} \\ \vdots & \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \cdots & 1 \end{bmatrix} \tag{5-13}$$

where σ_j^2 denotes the variance of pixels along the row. Similarly, considering the pixels along a column of the image as a Markov process, the covariance matrix can be written as

$$\left[C_{f_k}\right] = \sigma_k^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & 1 & \rho & \cdots & \rho^{N-2} \\ \vdots & \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \cdots & 1 \end{bmatrix} \tag{5-14}$$

where σ_k^2 is the variance of pixels along the column. With covariance matrices of rows and columns as defined in equation (5-13) and (5-14) the variance of the two dimensional slant transform samples may be written as

$$\left[V_F\right] = \left[V_{F_j}\right]\left[V_{F_k}\right]^T \tag{5-15}$$

where $\left[V_{F_j}\right]^T = \left[C_{F_j}(0,0)\ C_{F_j}(1,1) \ldots C_{F_j}(N-1,N-1)\right]$

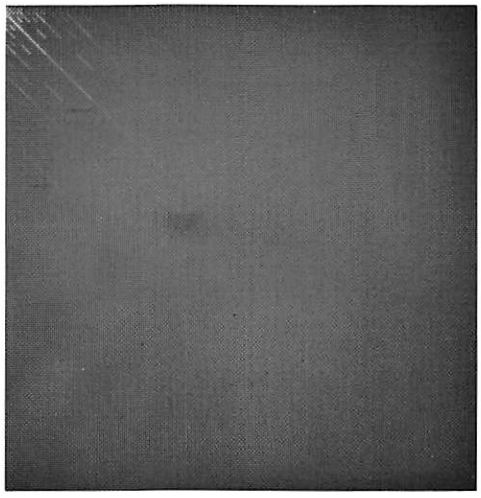
$\left[V_{F_k}\right]^T = \left[C_{F_k}(0,0)\ C_{F_k}(1,1) \ldots C_{F_k}(N-1,N-1)\right]$

and $\left[C_{F_j}\right] = \left[\mathscr{A}\right]\left[C_{f_j}\right]\left[\mathscr{A}\right]^T$

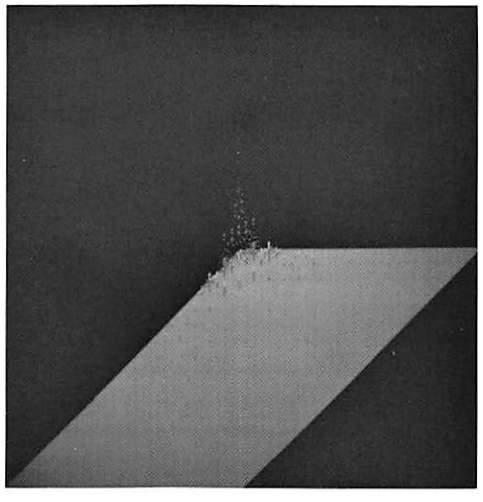
$\left[C_{F_k}\right] = \left[\mathscr{A}\right]\left[C_{f_k}\right]\left[\mathscr{A}\right]^T$

By setting some smaller elements of $\left[V_{F_j}\right]$ and $\left[V_{F_k}\right]$ zero, a special zone in $\left[V_F\right]$ which is constructed by non-zero elements will be formed. This special type of zone is called the "maximum variance zone."

Figure 5-8 contains a display of the covariance matrix, $\left[C_{F_j}\right]$ or $\left[C_{F_k}\right]$, of the slant transform with correlation coefficient 0.95 and matrix size of 256 x 256. It can be seen that $\left[V_{F_j}\right]$ or $\left[V_{F_k}\right]$ are the diagonal elements of $\left[C_{F_j}\right]$ or $\left[C_{F_k}\right]$ which appear in descending order. There are a few numbers of off diagonal elements in this picture whose values are much smaller than those in the diagonal. It should be noted that all these off-diagonal elements are zero for the Karhunen-Loeve transform. Figure 5-9 contains a plot of the variance function $\left[V_{F_x}\right]$ or $\left[V_{F_y}\right]$ of the slant transform as a



(a)



(b)

Figure 5-8. Full Size Slant Transform of Covariance Matrix - Markov Process (Correlation Coefficient = 0.95): (a) Transform; (b) Perspective View of Transform.

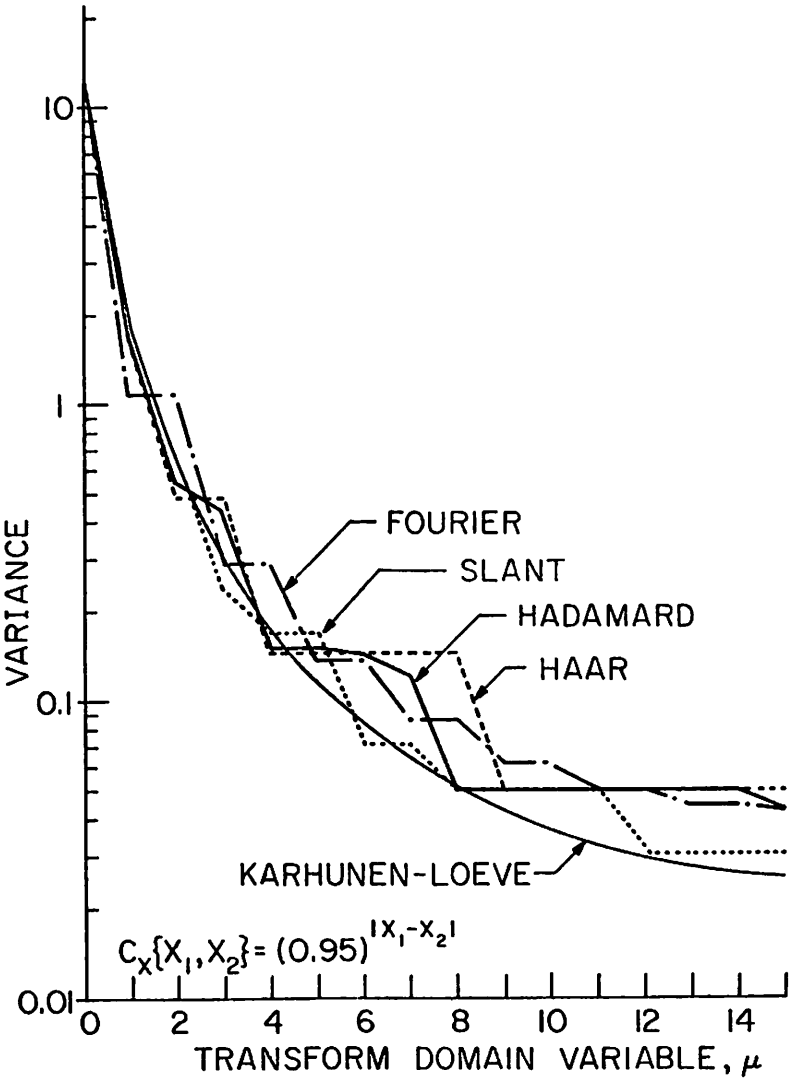


Figure 5-9. Transform Domain Variances; N=16, Element Correlation = 0.95.

function of frequency. The variance functions for the Hadamard, Fourier, Haar and Karhunen-Loeve transforms are included for comparison. It is seen that the variance function for the slant transform is reasonably close to the variance function of the Karhunen-Loeve transform which is known to provide the best energy compaction for the Markov source.

The statistical evaluation of the performance of various transforms in a form of the "maximum variance" zonal sampling can be specified in terms of the mean square error between the original image and its reconstruction for a statistical class of images. Figure 5-10 contains a plot of this mean square error as a function of block size for various transforms. This plot was obtained for an image statistically described by a Markov process. In the sampling process 25% of the transform samples with the largest variances, $V_F(u,v)$, were selected and the remainder were discarded according to the maximum variance zone. From the figure it is seen that the Karhunen-Loeve transform provides the best mean square error, while the slant transform results in only a slightly greater error. Also to be noted is that the rate of decrease in mean square error for larger block sizes becomes quite small after a block size of 32×32.

Figure 5-11 shows the slant transform "maximum variance" zonal sampling of the original images. A two dimensional slant

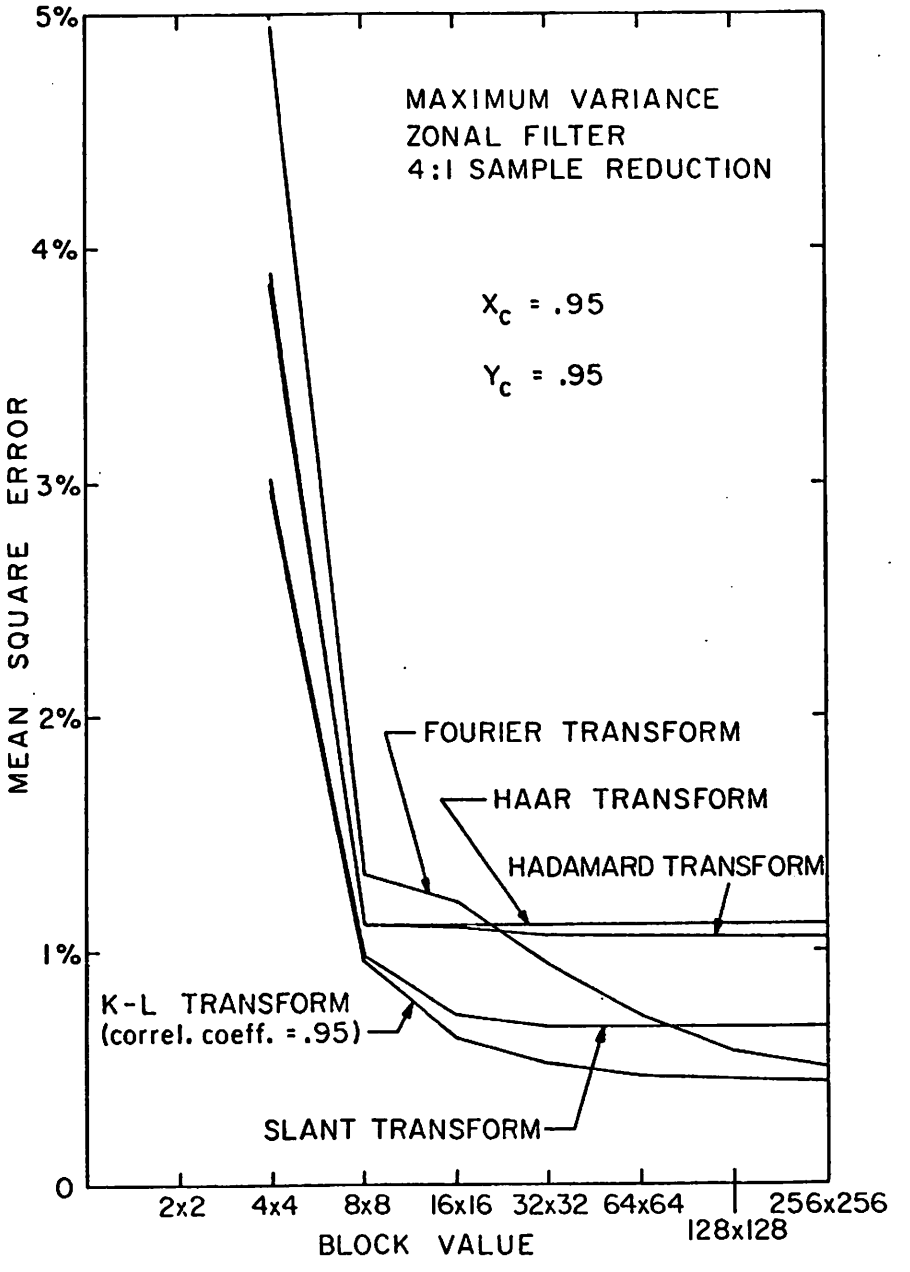


Figure 5-10. Mean Square Error Performance of Image Transforms as a Function of Block Size.

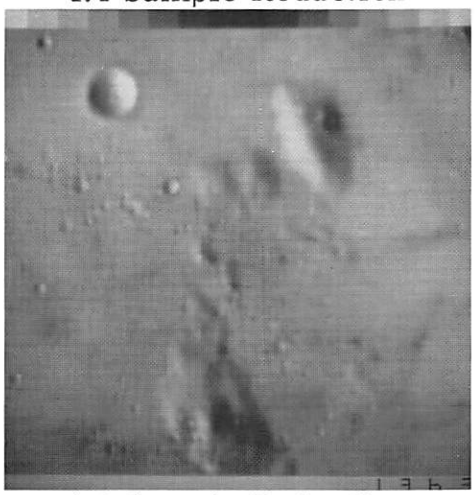
transform was taken of the image pixels repeatedly over subsections in 16 x 16 pixel blocks. A maximum variance zone was assigned to the transform samples by selecting only these samples inside the zone and an inverse transform was taken to form the reconstruction. Again, the reconstruction was formed without quantization or coding of transform domain samples. For purposes of comparison a series of experiments of the GIRL image have been performed for the Hadamard, Haar, Fourier and Karhunen-Loeve transforms which are shown in figures 5-12 and 5-13. It can be seen that the slant transform generally performs better than the rest of the transforms which possess a fast computational algorithm. Table 5-2 exhibits the mean square errors between the original GIRL image and its "maximum variance" zonal sampling reconstructions for various transforms with sample reductions of 2:1, 4:1, and 6:1. It can be noted from this table that the Haar transform generally does not perform very well under the maximum variance zonal sampling. This is due to the fact that most larger transform domain samples are not concentrated on the low frequency zone. It can also be noted that the Fourier transform with the sample reduction of 4:1 does rather well under the "maximum variance" zonal sampling. This indicates that the "maximum variance" model is particularly favorable to the Fourier transform zonal sampling with the sample reduction of 4:1. The slant transform generally performs well under

4:1 Sample Reduction

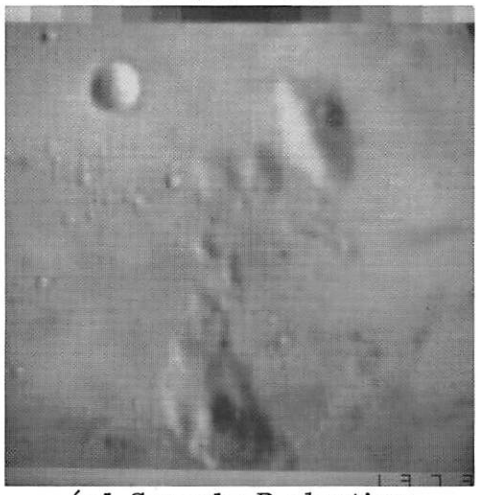
6:1 Sample Reduction

4:1 Sample Reduction

6:1 Sample Reduction



4:1 Sample Reduction



6:1 Sample Reduction

Figure 5-11. Slant Transform Zonal Sampling in 16 x 16 Pixel Blocks, Unquantized Transform.

Hadamard Transform
4:1 Sample Reduction

Hadamard Transform
6:1 Sample Reduction

Haar Transform
4:1 Sample Reduction

Haar Transform
6:1 Sample Reduction

Figure 5-12. Hadamard and Haar Transform Zonal Sampling in 16 x 16 Pixel Blocks, Unquantized Transform.

Fourier Transform
4:1 Sample Reduction

Fourier Transform
6:1 Sample Reduction

Karhunen-Loeve Transform
4:1 Sample Reduction

Karhunen-Loeve Transform
6:1 Sample Reduction

Figure 5-13. Fourier and Karhunen-Loeve Transform Zonal Sampling in 16 x 16 Pixel Blocks, Unquantized Transform.

Sample Reduction \ Transform	Slant	Hadamard	Haar	Fourier	Karhunen-Loeve
2:1	0.296%	0.396%	0.397%	0.299%	0.256%
4:1	0.960%	1.192%	1.100%	0.960%	0.774%
6:1	1.435%	1.860%	1.937%	1.730%	1.206%

Table 5-2 Mean Square Error Between the Original GIRL and Its Maximum Variance Zonal Sampling Reconstructions (Transforms are performed in 16 x 16 pixel blocks).

$$M.S.E = \sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\left[f(j,k)-\hat{f}(j,k)\right]^2 \Big/ \sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\left[f(j,k)\right]^2$$

the "maximum variance" model. But better results could no doubt be achieved if a better statistical model could be found.

5.2 Quantization Effect

There are two basic approaches to the quantization of transform samples: each sample could be quantized to the same number of levels; or the number of levels could be varied from sample to sample. Since the transform samples differ significantly in magnitude from sample to sample, the latter approach results in the most efficient coding. The following subsections present analytic methods of minimizing the mean square error between an original image and its reconstruction with the reconstruction taking into account the quantization effect in its transform samples. Two methods are considered: an approximate and an exact method. The analysis considers the quantization with various levels; quantization with the same number of levels is simply a special case of this analysis.

5.2.1 Calculation of Optimal Mean Square Error - Approximate Method

As denoted in equation (5-8) the mean square error between an original image and its reconstruction with a selected set of samples, but without quantization, may be computed by

$$\epsilon = \frac{1}{N^2} \sum_u \sum_v E\{[F(u,v)\,(1-S(u,v))]^2\} \qquad (5\text{-}16)$$

or

$$\epsilon = \frac{1}{N^2} \sum_{\substack{u \\ S(u,v)=0}} \sum_{v} E[F^2(u,v)] \tag{5-17}$$

Now with those selected samples from the output of the sample selector quantized, an additional quantization error will be introduced and, therefore, the total mean square error may be written as

$$\epsilon_T = \frac{1}{N^2} \left\{ \epsilon(0,0) + \sum_{\substack{u \\ (u,v)\neq(0,0) \\ S(u,v)=1}} \sum_{v} \epsilon(u,v) + \sum_{\substack{u \\ S(u,v)=0}} \sum_{v} E[F^2(u,v)] \right\} \tag{5-18}$$

where $\epsilon(0,0)$ and $\epsilon(u,v)$ denote the mean square quantization errors of dc and ac samples respectively. Now let the dc samples be quantized into L_1 levels by a set of decision levels, $Q_K(0,0)$, and a set of reconstruction levels, $F_K(0,0)$, as shown in figure 5-14. Then $\epsilon(0,0)$ in equation (5-18) may be written as

$$\epsilon(0,0) = \sum_{K=1}^{L_1} \int_{Q_{K-1}(0,0)}^{Q_K(0,0)} [F(0,0) - F_K(0,0)]^2 p_1\{F(0,0)\}\, dF(0,0) \tag{5-19}$$

where p_1 denotes the probability density of the dc samples. By the method introduced in appendix D, $\epsilon(0,0)$ can be minimized and written as

$$\epsilon(0,0) = \frac{1}{12 L_1^2} \left(\int_0^{A_1} p_1^{\frac{1}{3}} (F)\, d\, F \right)^3 \tag{5-20}$$

Let the ac samples be quantized into $L_2(u,v)$ levels by a set of decision levels, $Q_K(u,v)$, and a set of reconstruction levels, $F_K(u,v)$, as shown in figure 5-15. Then $\epsilon(u,v)$ in equation (5-18) may be written as

$$\epsilon(u,v) = \sum_{K=\frac{-L_2(u,v)}{2}}^{\frac{L_2(u,v)}{2}} \int_{Q_{K-1}(u,v)}^{Q_K(u,v)} [F(u,v)-F_K(u,v)]^2 p_2\{F(u,v)\}\, dF(u,v) \tag{5-21}$$

where p_2 denotes the probability density of ac samples. By the method introduced in appendix D, equation (5-21) can be minimized and written as

$$\epsilon(u,v) = \frac{2}{3 L_2^2(u,v)} \left(\int_0^{A_2(u,v)} p_2^{\frac{1}{3}} (F)\, dF \right)^3 \tag{5-22}$$

With substitutions of equations (5-20) and (5-22), equation (5-18) becomes

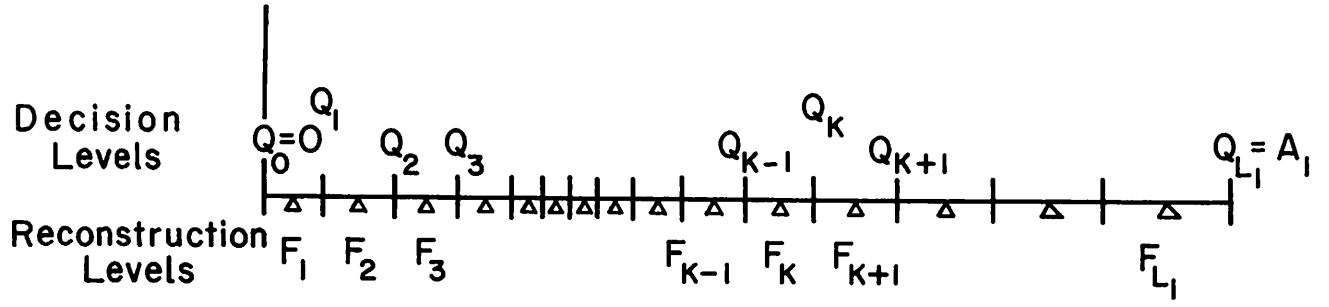


Figure 5-14. Decision and Reconstruction Levels of dc Samples.

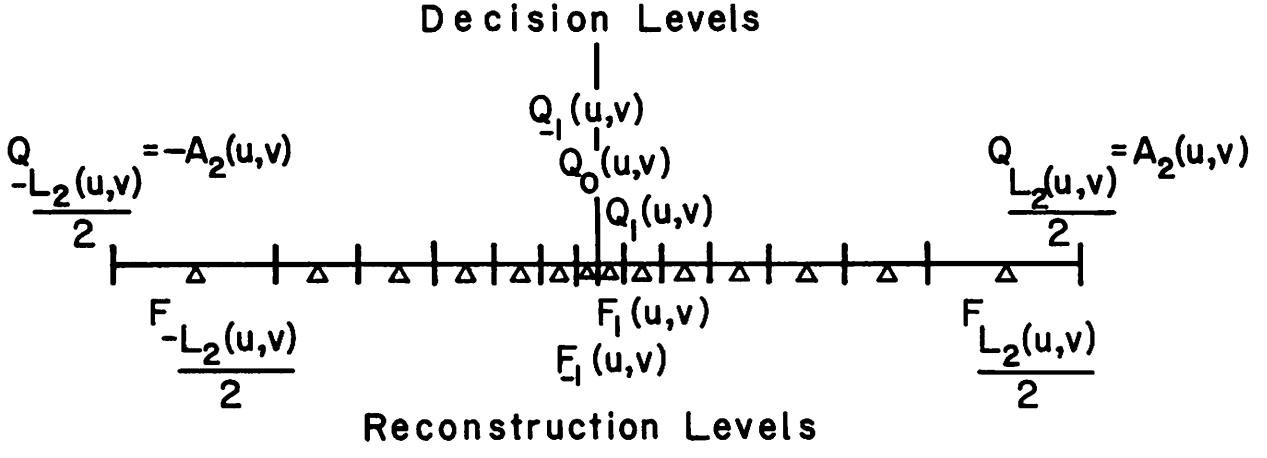


Figure 5-15. Decision and Reconstruction Levels of ac Samples.

$$\epsilon_T = \frac{1}{N^2}\left\{\frac{1}{12 L_1^2}\left(\int_0^{A_1} p_1^{\frac{1}{3}}(F)\,dF\right)^3 + \sum_u \sum_v \frac{2}{3 L_2^2(u,v)}\right.$$
$$(u,v)\neq(0,0)$$
$$S(u,v)=1$$
$$\left.\cdot\left(\int_0^{A_2(u,v)} p_2^{\frac{1}{3}}(F)\,dF\right)^3 + \sum_u \sum_v \int_\infty^\infty F\, p_2(F)\,dF\right\} \quad (5\text{-}23)$$
$$S(u,v)=0$$

Example 5-1: Suppose the probability density of dc samples is Rayleigh and ac samples is Gaussian. Then

$$\epsilon_T = \frac{1}{N^2}\left\{\frac{1}{12 L_1^2}\left(\int_0^{A_1}\left[\frac{F}{\alpha^2}\exp\left(-\frac{F^2}{2\alpha^2}\right)\right]^{\frac{1}{3}} dF\right)^3\right.$$
$$+ \sum_u \sum_v \frac{2}{3 L_2^2(u,v)}\left(\int_0^{A_2(u,v)}\left[\frac{1}{\sqrt{2\pi}\,\sigma(u,v)}\exp\left(-\frac{F^2}{2\sigma^2(u,v)}\right)\right]^{\frac{1}{3}} dF\right)^3$$
$$(u,v)\neq(0,0)$$
$$S(u,v)=1$$
$$\left. + \sum_u \sum_v \int_{-\infty}^{\infty} \frac{F^2}{\sqrt{2\pi}\,\sigma(u,v)}\exp\left(-\frac{F^2}{2\sigma^2(u,v)}\right)dF\right\} \quad (5\text{-}24)$$
$$(u,v)\neq(0,0)$$
$$S(u,v)=0$$

where $2\alpha^2$ and $\sigma^2(u,v)$ are variances of Rayleigh and Gaussian densities. Equation (5-24) can be simplified and rewritten as

$$\epsilon_T = \frac{1}{N^2}\left\{ \frac{1}{12 L_1^2 \alpha^2}\left[\int_0^{A_1} F^{\frac{1}{3}} \exp\left(-\frac{F^2}{6\alpha^2}\right) dF \right]^3 \right.$$

$$+ \sum_u \sum_v \frac{4\sqrt{3}\,\pi\sigma^2(u,v)}{L_2^2(u,v)}\left[\operatorname{erf}\frac{A_2(u,v)}{\sqrt{3}}\right]^3 + \sum_u \sum_v \sigma^2(u,v)\Big\} \tag{5-25}$$

$(u,v)\neq(0,0)$, $S(u,v)=1$; $S(u,v)=0$

where $\operatorname{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{y^2}{2}\right) dy$ (5-26)

Example 5-2: Suppose the probability density of dc samples is Rayleigh and ac samples is two-sided exponential. Then

$$\epsilon_T = \frac{1}{N^2}\left\{ \frac{1}{12 L_1^2}\left(\int_0^{A_1}\left[\frac{F}{\alpha^2} \exp\left(-\frac{F^2}{2\alpha^2}\right)\right]^{\frac{1}{3}} dF \right)^3 \right.$$

$$+ \sum_u \sum_v \frac{2}{3 L_2^2(u,v)}\left(\int_0^{A_2(u,v)} \left[\frac{\beta(u,v)}{2} \exp(-\beta(u,v)F)\right]^{\frac{1}{3}} dF \right)^3$$

$(u,v)\neq(0,0)$, $S(u,v)=1$

$$+ \sum_u \sum_v \int_{-\infty}^{\infty} \frac{\beta(u,v)}{2} F^2 \exp(-\beta F)\, dF \Big\} \tag{5-27}$$

$s(u,v)=0$

where $2\alpha^2$ and $\frac{2}{\beta^2(u,v)}$ are variances of Rayleigh and two-side exponential densities respectively. Equation (5-27) can be simplified

and rewritten as

$$e_T = \frac{1}{N^2}\left\{\frac{1}{12 L_1^2 \alpha^2}\left[\int_0^{A_1} F^{\frac{1}{3}} \exp\left(-\frac{F^2}{6\alpha^2}\right) dF\right]^3\right.$$

$$+ \sum_{\substack{u\ v \\ (u,v)\neq(0,0) \\ S(u,v)=1}} \frac{9}{\beta^2(u,v) L_2^2(u,v)} \left[1 - \exp\left(-\frac{\beta(u,v) A_2(u,v)}{3}\right)\right]^3$$

$$\left. + \sum_{\substack{u\ v \\ S(u,v)=0}} \frac{2}{\beta^2(u,v)} \right\} \qquad (5\text{-}28)$$

5.2.2 Calculation of Optimal Mean Square Error - Exact Method

The total mean square error shown in equation (5-23) is valid if the number of quantization levels is large. This is always not the case. In some applications as few as two quantization levels are assigned to transform samples. This has led to the necessity of deriving a more accurate equation which is applicable to any number of quantization levels. The following paragraphs present a derivation of the optimal mean square error which could be applied to any quantization level.

As denoted in equation (5-18) the total mean square error may be written as

$$\epsilon_T = \frac{1}{N^2}\left\{ \epsilon(0,0) + \sum_{\substack{u\ v \\ (u,v)\neq(0,0) \\ S(u,v)=1}} \epsilon(u,v) + \sum_{\substack{u\ v \\ S(u,v)=0}} E\left[F^2(u,v)\right] \right\} \qquad (5\text{-}29)$$

where $\epsilon(0,0)$ and $\epsilon(u,v)$ are the mean square errors of dc and ac samples respectively. Now suppose the dc samples are quantized into L_1 levels by a set of decision levels, Q_K, and a set of reconstruction levels, F_K, as shown in Figure 5-14. Then $\epsilon(0,0)$ may be written as

$$\epsilon(0,0) = \sum_{K=1}^{L_1} \int_{Q_{K-1}}^{Q_K} (F-F_K)^2\, p_1(F)\, dF \qquad (5\text{-}30)$$

In order to minimize $\epsilon(0,0)$ one may take the partial derivative of equation (5-30) with respect to Q_j and F_j and set the results to zero, i.e.,

$$\frac{\partial \epsilon(0,0)}{\partial Q_j} = (Q_j - F_j)^2\, p_1(Q_j) - (Q_j - F_{j+1})^2 p_1(Q_j) = 0 \qquad (5\text{-}31)$$

and

$$\frac{\partial \epsilon(0,0)}{\partial F_j} = 2\int_{Q_{j-1}}^{Q_j} (F-F_j)\, p_1(F)\, dF = 0 \qquad j=1,2,\ldots,L_1 \qquad (5\text{-}32)$$

Equation (5-31) and (5-32) may be rewritten as

$$F_{j+1} = 2Q_j - F_j \qquad (5\text{-}33)$$

and

$$\int_{Q_{j-1}}^{Q_j} (F-F_j)\, p_1(F)\, dF \qquad j=1,2,\ldots,L_1 \qquad (5\text{-}34)$$

With the method suggested by Max (21), one can find the optimal decision levels Q_K, $K=0,1,..,L_1$, and reconstruction levels F_K, $K=1,2,\ldots,L_1$, that satisfy equations (5-33) and (5-34). With the help of equation (5-34), equation (5-30) can be simplified as

$$\epsilon(0,0) = \int_0^{\infty} F^2\, p_1(F)\, dF - \sum_{K=1}^{L_1} F_K^2 \int_{Q_{K-1}}^{Q_k} p_1(F)\, dF \qquad (5\text{-}35)$$

where $Q_{L_1} = A_1 = \infty$ in this case. Similarly, suppose the ac samples are quantized into $L_2(u,v)$ levels by a set of decision levels, $Q_K(u,v)$, and a set of reconstruction levels, $F_K(u,v)$, as shown in figure (5-15). Then the ac mean square error, $\epsilon(u,v)$, may be written as

$$\epsilon(u,v) = 2 \sum_{K=1}^{\frac{L_2(u,v)}{2}} \int_{Q_{K-1}(u,v)}^{Q_K(u,v)} [F-F_K(u,v)]^2 p_1(F)\, dF \qquad (5\text{-}36)$$

The optimal decision and reconstruction levels may then be calculated by a method similar to the dc case which will satisfy the following equations

$$F_{j+1}(u,v) = 2Q_j(u,v) - F_j(u,v) \qquad j=1,2,\ldots,L_2(u,v) \qquad (5\text{-}37)$$

$$\int_{Q_{j-1}(u,v)}^{Q_j(u,v)} [F-F_j(u,v)]\, p_2(F)\,dF=0 \qquad j=1,2,\ldots,L_2(u,v) \qquad (5\text{-}38)$$

With the help of equation (5-38) the optimal mean square error of ac samples then can be written as

$$\epsilon_T = \frac{1}{N^2}\Bigg\{ \int_0^\infty F^2\, p_1(F)\, dF - \sum_{K=1}^{L_1} F_K^2 \int_{Q_{K-1}(0,0)}^{Q_K(0,0)} p_1(F)\, dF$$

$$+2\sum_{\substack{u\ v\\ (u,v)\neq(0,0)\\ S(u,v)=1}} \Bigg[\int_0^\infty F^2\, p_2(F)\,dF - \sum_{K=1}^{\frac{L_2(u,v)}{2}} F_K^2(u,v) \int_{Q_{K-1}(u,v)}^{Q_K(u,v)} p_2(F)\, dF \Bigg]$$

$$+ \sum_{\substack{u\ v\\ S(u,v)=0}} E[\, F^2(u,v)] \Bigg\} \qquad (5\text{-}40)$$

Example 5-3: Suppose dc and ac probability densities of transform domain samples are Rayleigh and Gaussian respectively. Then the minimal total mean square error is

$$\epsilon_T = \frac{1}{N^2}\Bigg\{\int_0^{\infty} \frac{F^3}{\alpha^2} \exp\Big(-\frac{F^2}{2\alpha^2}\Big)dF - \sum_{K=1}^{L_1} F_K^2 \int_{Q_{K-1}(0,0)}^{Q_K(0,0)} \frac{F}{\alpha^2}\exp\Big(-\frac{F^2}{2\alpha^2}\Big)\,dF$$

$$+ 2 \sum_{\substack{u \\ (u,v)\neq(0,0)}} \sum_{\substack{v \\ S(u,v)=1}} \Bigg[\int_0^{\infty} \frac{F^2}{\sqrt{2\pi}\,\sigma(u,v)} \exp\Big(-\frac{F^2}{2\sigma^2(u,v)}\Big)dF$$

$$- \sum_{K=1}^{\frac{L_2(u,v)}{2}} F_K^2(u,v) \int_{Q_{K-1}(u,v)}^{Q_K(u,v)} \frac{1}{\sqrt{2\pi}\,\sigma(u,v)} \exp\Big(-\frac{F^2}{2\sigma^2(u,v)}\Big)dF\Bigg]$$

$$+ \sum_{\substack{u \\ S(u,v)=0}} \sum_{v} \sigma^2(u,v)\Bigg\} \qquad (5\text{-}41)$$

where $2\alpha^2$ and $\sigma^2(u,v)$ are the variances of Rayleigh and Gaussian densities respectively. By carrying out the integrals, equation (5-41) can be simplifed as

$$\epsilon_T = \frac{1}{N^2}\Bigg\{2\alpha^2 - \sum_{K=1}^{L_1} F_K^2(0,0)\Big[\exp\Big(-\frac{Q_{K-1}^2(0,0)}{2\alpha^2}\Big) - \exp\Big(-\frac{Q_K^2(0,0)}{2\alpha^2}\Big)\Big]$$

$$+ \sum_{\substack{u \\ (u,v)\neq(0,0)}} \sum_{v} \sigma^2(u,v) - 2 \sum_{\substack{u \\ (u,v)\neq(0,0) \\ S(u,v)=1}} \sum_{v} \sum_{K=1}^{\frac{L_2(u,v)}{2}} F_K^2(u,v)\Big[\operatorname{erf}\Big(\frac{Q_K(u,v)}{\sigma(u,v)}\Big)$$

$$- \operatorname{erf}\Big(\frac{Q_{K-1}(u,v)}{\sigma(u,v)}\Big)\Big]\Bigg\} \qquad (5\text{-}42)$$

Example 5-4: Suppose dc and ac probability densities of transform domain samples are Rayleigh and two-sided exponential respectively. Then the minimal total mean square error is

$$\epsilon_T = \frac{1}{N^2}\left\{ \int_0^{\infty} \frac{F^3}{\alpha^2} \exp\left(-\frac{F^2}{2\alpha^2}\right) dF - \sum_{K=1}^{L_1} F_K^2 \int_{Q_{K-1}(0,0)}^{Q_K(0,0)} \frac{F}{\alpha} \exp\left(-\frac{F^2}{2\alpha^2}\right) dF \right.$$

$$+ \sum_{u} \sum_{v} \left[\int_0^{\infty} \frac{\beta(u,v) F^2}{2} \exp\left(-\beta(u,v) F\right) dF \right.$$

$$(u,v) \neq (0,0)$$

$$S(u,v)=1$$

$$\left. \left. - \sum_{K=1}^{\frac{L_2(u,v)}{2}} F_K^2(u,v) \int_{Q_{K-1}(u,v)}^{Q_K(u,v)} \frac{\beta(u,v)}{2} \exp\left(-\beta(u,v) F\right) dF \right] \right.$$

$$\left. + \sum_{u} \sum_{v} \frac{2}{\beta^2(u,v)} \right\} \qquad (5\text{-}43)$$

$$S(u,v)=0$$

where $2\alpha^2$ and $\frac{2}{\beta^2(u,v)}$ are the variances of Rayleigh and two-sided exponential densities respectively. Equation (5-70) can be simplified and rewritten as

$$\epsilon_T = \frac{1}{N^2}\left\{ 2\alpha^2 - \sum_{K=1}^{L_1} F_K^2(0,0)\left[\exp\left(-\frac{Q_{K-1}^2(0,0)}{2\alpha^2}\right) - \exp\left(-\frac{Q_K^2(0,0)}{2\alpha^2}\right)\right]\right.$$

$$+ \sum_{\substack{u\ v \\ (u,v)\neq(0,0)}} \frac{2}{\beta^2(u,v)} - \sum_{\substack{u\ v \\ (u,v)\neq(0,0) \\ S(u,v)=0}} \sum_{K=1}^{\frac{L_2(u,v)}{2}} F_K^2(u,v)\left[\exp(-\beta(u,v)Q_{K-1}(u,v))\right.$$

$$\left.\left. - \exp(-\beta(u,v)Q_K(u,v))\right]\right\} \qquad (5\text{-}44)$$

5.3 Coding Effects

The coding techniques for quantized transform samples are quite different for threshold sampling and zonal sampling. In threshold sampling the locations of samples exceeding the threshold varies from image to image. Therefore, it is necessary to code the position of samples exceeding the threshold level. Position coding is not necessary for zonal sampling since the location of samples to be coded is known apriori.

5.3.1 Threshold Coding

There are a variety of ways that position coding could be employed. The simplest method conceptually would be to code the coordinates of each significant samples. However, higher coding efficiency can be obtained by coding the number of non-significant

samples between significant samples. This scheme, called run length coding, is implemented as follows:

(1) The first sample along each line is coded regardless of its magnitude. A position code bits of all zeros or all ones affixed to the amplitude provides a line synchronization code group;

(2) The amplitude of the second run length code word is the coded amplitude of the next significant sample. The position code is the binary count of the number of samples of the significant sample from the previous significant sample;

(3) If a significant sample is not encountered after scanning the maximum run length of samples, the position and amplitude code bits are set to all ones to indicate a maximum run length.

The advantage of including a line synchronization code group is that it becomes unnecessary to code the line number and, also, it prevents the propagation of channel errors over more than one line. A simple code to implement this run length coding procedure is given as follows:

position	amplitude	
X X . . . X X	Y Y . . . Y Y	
1 1 1 1 1 1	0 0 0 0 0 0	first sample of a line below threshold
0 0 . . . 0 0	Y Y . . . Y Y	first sample of a line above threshold
0 0 . . . 0 1	Y Y . . . Y Y	run length = 1
0 0 . . . 1 0	Y Y . . . Y Y	run length = 2
.	.	
.	.	
.	.	
1 1 1 1 1 0	Y Y . . . Y Y	run length = 2^N-2 (where N is number of position code bits)
1 1 1 1 1 1	1 1 1 1 1 1	pseudo-run of length 2^N-2

This run length coding procedure for transform threshold coding has been tested for the GIRL, COUPLE and MOONSCENE. Table 5-3 shows the relationship between sample reduction and average code bits per pixel for the slant transform of the GIRL image with the amplitude of each significant sample quantized and coded into six bits. It can be seen that with a 4:1 sample reduction the best number of position code bits is four, and with 6:1 and 12:1 sample reductions the best number of position code bits are five and six respectively. Figure 5-16 shows a plot of sample reduction versus average number of code bits per pixel for the slant transformed GIRL image with various numbers of position code bits. It can be seen that the number of position code bits changes with the sample reduction in order to

	Number of Position Code Bits				
Sample Reduction	3	4	5	6	7
4:1	3.028	2.768	2.850	3.052	3.289
6:1	2.511	2.062	1.989	2.077	2.215
12:1	1.993	1.378	1.152	1.127	1.162

Table 5-3 Relationship Between Sample Reduction and Average Bits Per Pixel of the Slant Transform Threshold Coded GIRL Image.

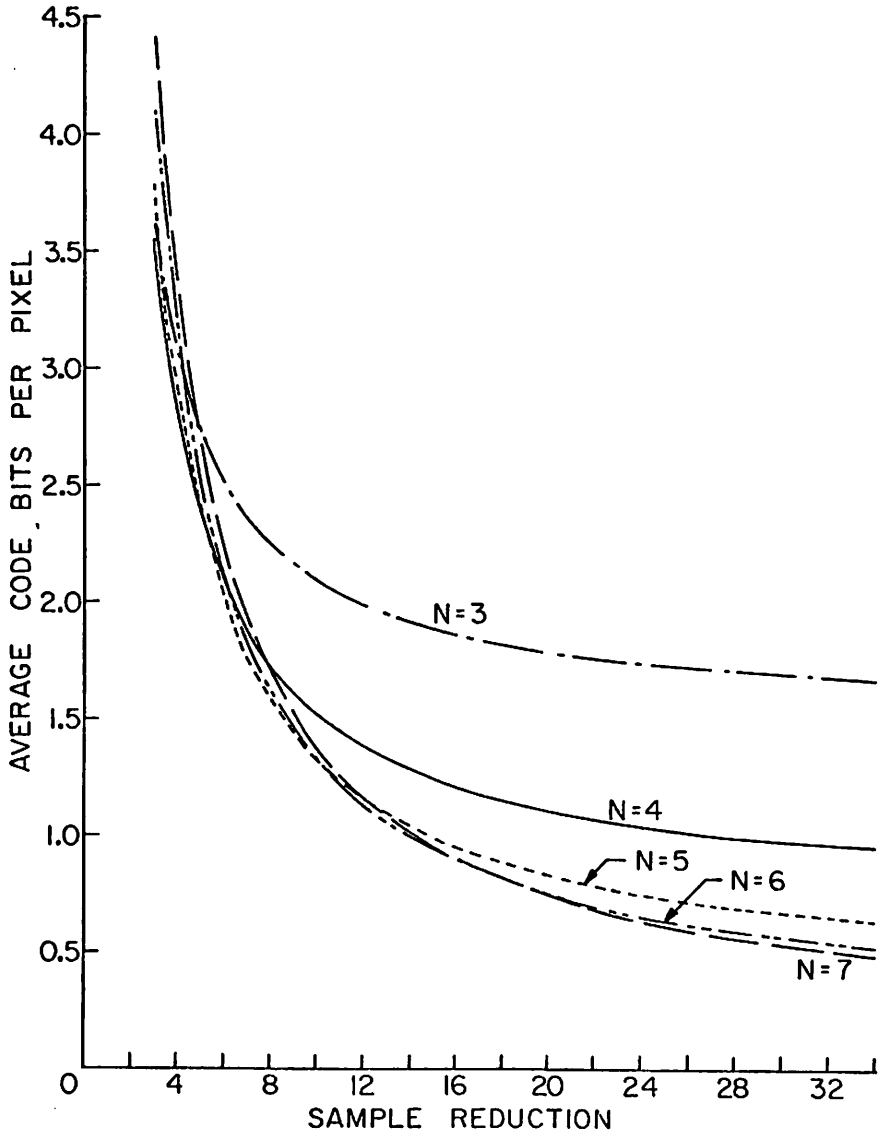


Figure 5-16. Sample Reduction Versus Average Code Bits Per Pixel for Slant Transform Threshold Sampling of GIRL Image in 16 x 16 Pixel Blocks.

achieve the least average number of code bits per pixel. For a sample reduction factor of less than 5:1 the best number of position code bits is four, and for a sample reduction of from 5:1 to 9.5:1 the best number of position code bits is five, etc.

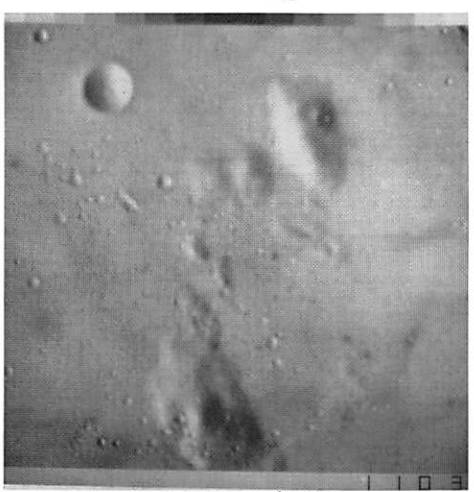
Figure 5-17 shows the quantized and coded versions of Figure 5-5. A two dimensional slant transform was taken of the image pixels repeatedly over subsections of an image in 16 x 16 pixel blocks. A threshold was assigned to make the sample reduction, and then these reduced samples were optimally quantized and coded. Following the decoding, an inverse slant transform was taken to reconstruct the original image. The quantization scheme used was the one suggested by Max [21] where dc decision and reconstruction levels were obtained by solving equations (5-33) and (5-34) and ac decision and reconstruction levels were obtained by solving equations (5-37) and (5-38). The dc probability density p_1 in equation (5-34) and the ac probability density p_2 in equation (5-38) were assumed to be Rayleigh and Gaussian, respectively. It was found experimentally that, for the optimal reconstruction of test images, the variances of p_1 and p_2 were 670 and 200, respectively. For the runlength coding and decoding parts of the experiments a constant code word length of six bits was assigned to the amplitude of each significant transform sample and a constant code word length of four bits was assigned to the position. As expected, quantization increases the mean square

1.99 bits/pixel

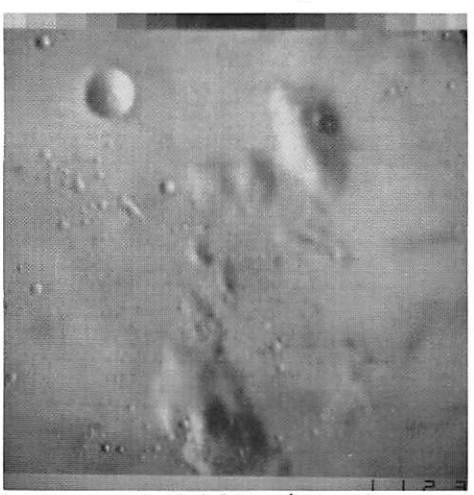
1.15 bits/pixel

1.99 bits/pixel

1.15 bits/pixel



1.99 bits/pixel



1.15 bits/pixel

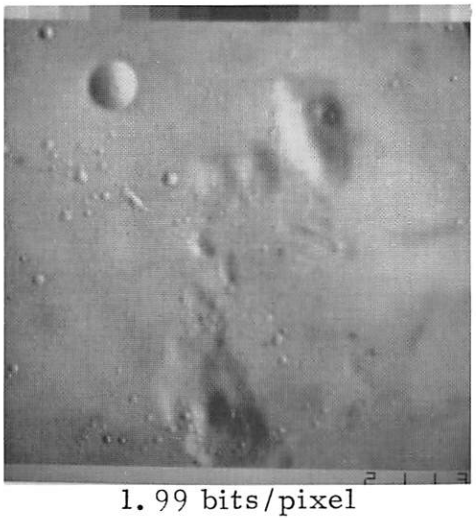
Figure 5-17. Slant Transform Threshold Sampling in 16 x 16 Pixel Blocks, Quantized Transform.

1.99 bits/pixel

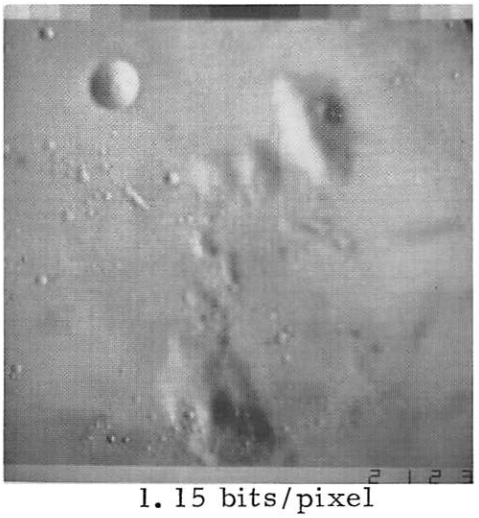
1.15 bits/pixel

1.99 bits/pixel

1.15 bits/pixel



1.99 bits/pixel



1.15 bits/pixel

Figure 5-18. Hadamard Transform Threshold Sampling in 16 x 16 Pixel Blocks, Quantized Transform.

error; but, subjectively, the reconstructed images appear to have little visible degradation. Figure 5-18 illustrates the quantized and coded pictures of the Hadamard transform. It is easily seen that the slant transform in figure 5-17 performs better than those in figure 5-18.

Table 5-4 exhibits the mean square errors of the slant transformed GIRL image with sample reductions of 4:1, 6:1 and 12:1. The corresponding thresholds and average code bits per pixel are also included. Figure 5-19 shows a plot of sample reduction versus mean square error between the original and its reconstructions for the slant transform threshold sampled GIRL image. Since the dc samples are much larger than the ac samples, the former are essentially all quantized and, therefore, the mean square quantization error is nearly constant. For ac samples, as the sample reduction factor increases, the number of significant samples decrease and, therefore, the mean square quantization error decreases accordingly. The mean square error after the sample reduction (unquantized) and total mean square error are almost linearly related to the sample reduction factor, which is expected for the slant transform threshold sampled or coded images.

5.3.2 Zonal Coding and Bit Allocation

The quantization levels and code bit assignment for each

Sample Reduction	Threshold	# Of Bits Per Pixel	Unquantized M. S. E.	M. S. E. of dc Samples	M. S. E. of ac Samples	Total M. S. E.
4:1	6.77	2.768	0.166%	0.0039%	0.172%	0.342%
6:1	9.98	1.989	0.299%	0.0039%	0.127%	0.430%
12:1	18.91	1.152	0.681%	0.0039%	0.090%	0.775%

Table 5-4 Mean Square Error Versus Sample Reduction Factor of the Slant Transform Threshold Coded GIRL Image (Transform is performed in 16 x 16 pixel blocks).

$$\text{M. S. E.} = \sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\left[f(j,k)-\hat{f}(j,k)\right]^2 \Big/ \sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\left[f(j,k)\right]^2$$

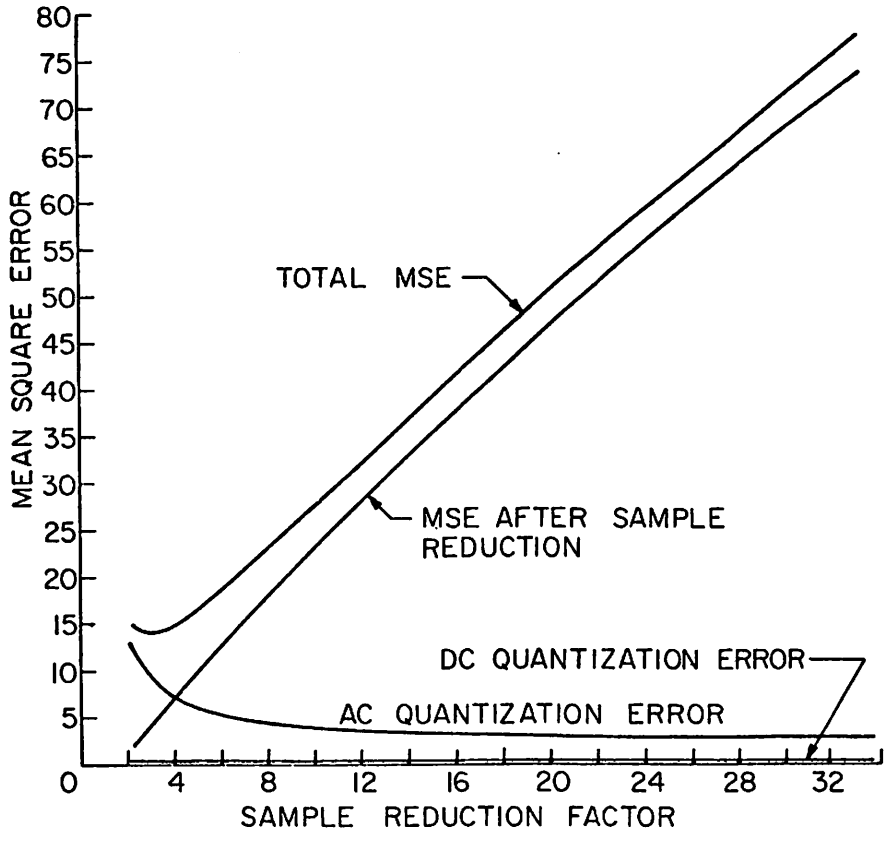


Figure 5-19. Mean Square Error Versus Sample Reduction Factor for the Slant Transform Threshold Sampling of GIRL in 16 x 16 Pixel Blocks.

significant dc or ac sample have been kept constant in the previous section. For zonal sampling the quantization levels and code bit assignment can be varied according to the statistics of the transform domain samples. Since the magnitude of the transform domain samples appears to be in descending order from low to high frequencies, as can be seen in figure 5-2, the variation of quantization levels and the allocation of code bits will certainly result in the most efficient coding. The total mean square error between the original image and the reconstructed image can easily be calculated . Suppose a constant word length code of length

$$L_C(u,v) = 2^{N_B(u,v)} \tag{5-45}$$

is assigned to each quantization level and a total of

$$N_B = \sum_{\substack{u \\ S(u,v)=1}} \sum_{v} N_B(u,v) \tag{5-46}$$

bits are allocated for transform domain samples. Then the minimum total mean square error in equation (5-42) becomes

$$\epsilon_T = \frac{1}{N^2}\Bigg\{ \int_0^{\infty} F^2 \, p_1(F)\, dF - \sum_{K=1}^{2^{N_B(0,0)}} F_K \int_{Q_{K-1}(0,0)}^{Q_K(0,0)} p_1(F)\, dF$$

$$+ 2 \sum_{\substack{u \\ (u,v)\neq(0,0) \\ S(u,v)=1}} \sum_{v} \Bigg[\int_0^{\infty} F^2 p_2(F)\, dF - \sum_{K=1}^{2^{[N_B(u,v)-1]}} F_K^2(u,v) \int_{Q_{K-1}(u,v)}^{Q_K(u,v)} p_2(F)\, dF \Bigg]$$

$$+ \sum_{\substack{u \\ S(u,v)=0}} \sum_{v} E[F^2(u,v)] \Bigg\} \qquad (5\text{-}47)$$

Let the probability densities of the dc and ac samples be modelled as Rayleigh and Gaussian densities respectively. Equation (5-47) then can be written as

$$\epsilon_T = \frac{1}{N^2}\Bigg\{ 2\alpha^2 - \sum_{K=1}^{2^{N_B(0,0)}} F_K^2(0,0)\left[\exp\left(-\frac{Q_{K-1}^2(0,0)}{2\sigma^2}\right) - \exp\left(-\frac{Q_K^2(0,0)}{2\sigma^2}\right)\right]$$

$$+ \sum_{\substack{u \\ (u,v)\neq(0,0)}} \sum_{v} \sigma^2(u,v) - \sum_{\substack{u \\ (u,v)\neq(0,0) \\ S(u,v)=1}} \sum_{v} \sum_{K=1}^{2^{[N_B(u,v)-1]}} F_K^2(u,v)\left[\operatorname{erf}\left(\frac{Q_K(u,v)}{(u,v)}\right)\right.$$

$$\left. - \operatorname{erf}\left(\frac{Q_{K-1}(u,v)}{(u,v)}\right)\right]\Bigg\} \qquad (5\text{-}48)$$

where $\operatorname{erf} x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{y^2}{2}\right) dy$. It should be noted that the constant code word length $L_C(u,v)$ in equation (5-45) is the same as

the quantization levels L(u, v).

The bit assignment $N_B(u, v)$ for each transform domain sample has been based upon an algorithm of rate distortion theory [19]. The number of bits is selected according to the relation

$$N_B(u, v) = \ell n \lceil V_F(u, v) \rceil - \ell n \lceil D \rceil \qquad (5\text{-}49)$$

where $V_F(u, v)$ is the variance of a transform domain sample and D is proportional to the mean square error of the coding process. A plot of rate versus distortion for various transforms is shown in appendix C. Figure 5-20 illustrates a typical assignment of code bits for the slant transform zonal coding in 16 x 16 pixel blocks.

8	8	8	7	7	7	5	5	4	4	4	4	4	4	4	4
8	8	7	5	5	5	3	3	3	3	3	3	2	2	2	2
8	7	6	4	4	4	3	3	2	2	2	2	2	2	2	2
7	5	4	3	2	2	2	2	0	0	0	0	0	0	0	0
7	5	4	2	2	2	2	2	0	0	0	0	0	0	0	0
7	5	4	2	2	2	2	2	0	0	0	0	0	0	0	0
5	3	3	2	2	2	0	0	0	0	0	0	0	0	0	0
5	3	3	2	2	2	0	0	0	0	0	0	0	0	0	0
4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5-20. Typical Bit Assignments for the Slant Transform Zonal Coding in 16 x 16 Pixel Block.

The performance of the transform coding system can again be evaluated in terms of the mean square error between the original image, which is statistically described by a Markov process, and its reconstruction. Figure 5-21 contains a refinement of the work which has been done in Figure 5-10. Every sample inside the maximum variance zone is quantized, and bit allocated according to the relation shown in equation (5-49). The maximum variance zone is adjusted such that an average of 1.5 bits/pixel is used to code the samples inside the zone. Again from the figure it is seen that the Karhunen-Loeve transform provides the minimum mean square error and the slant transform results in only a slightly greater error. By comparing figures 5-21 and 5-10 it is easily seen that the variation of quantization levels and the allocation of code bits reduces the mean square error for all transforms (with the exception of 4 x 4 block size).

Figure 5-22 shows the optimally quantized and bit allocated versions of figure 5-11. The quantization scheme used was again a nonlinear quantization rule where dc decision and reconstruction levels were obtained by solving equations (5-33) and (5-34) and ac decision and reconstruction levels were obtained by solving equations (5-37) and (5-38). The variance of dc samples, $V_F(0,0)$, is defined in equation (5-15). In the process of finding $V_F(0,0)$, it was determined experimentally that the variance of pixels along the row (or

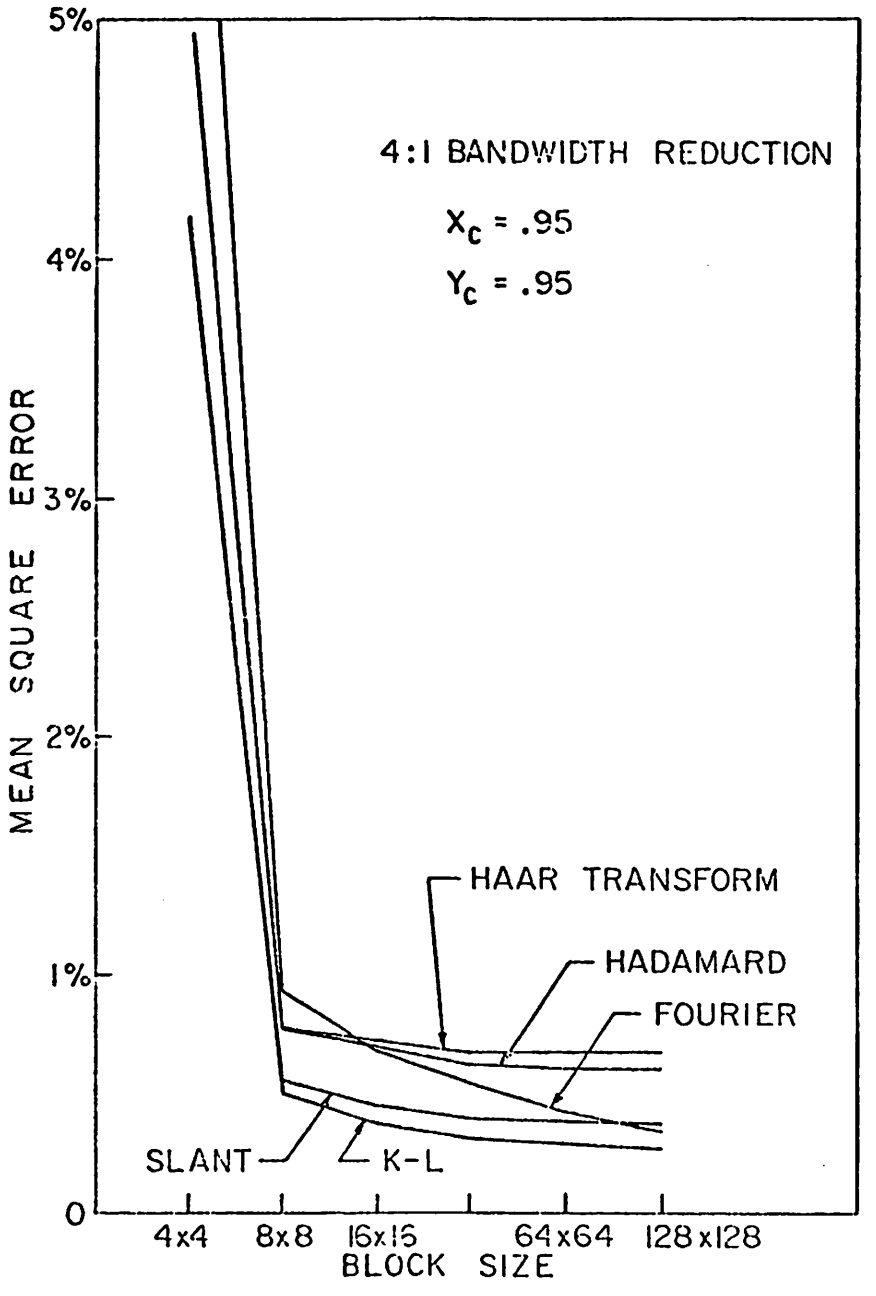


Figure 5-21. Mean Square Error Performance of Image Transforms as a Function of Block Size.

column) in equation (5-13) (or equation (5-14))is 24, i.e., $\sigma_j^2 = \sigma_k^2 = 24$. The variance of ac samples, $V_F(u,v)$, is also defined in the variance matrix $[V_F]$ where the experimental result of σ_j^2 or σ_k^2 is dependent upon the bandwidth reduction which can be summarized as follows:

Average Code Bits	σ_j^2 in equation (5-14) or σ_j^2 in equation (5-15)
0.5	5.75
1.0	6.03
1.5	6.30
2.0	6.58

From these figures it can be seen that the quantization and the bit allocation improve the quality of picture substantially. For purposes of comparison the Hadamard transform zonal coded pictures are also shown in figure 5-23. It is easily seen that the slant transform performs better than the Hadamard transform subjectively.

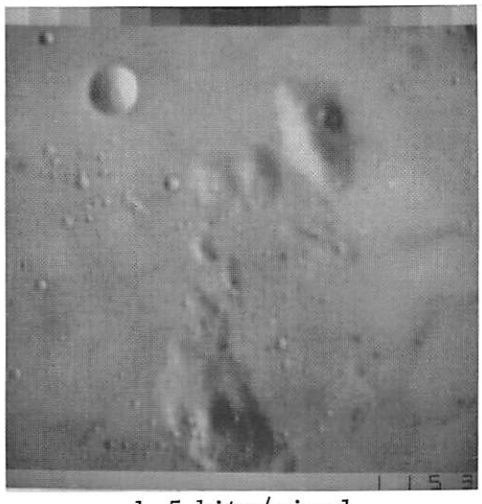
Figure 5-24 contains a plot of the mean square error of the optimally quantized and bit allocated slant transform GIRL image for threshold and zonal codings as a function of average code bits per pixel. It is seen that threshold coding results in a better mean square error for an average code bit per pixel of 1.9 or less. For an average code bit per pixel of greater than 1.9 the zonal coding appears to be favorable in the mean square sense.

1.5 bits/pixel

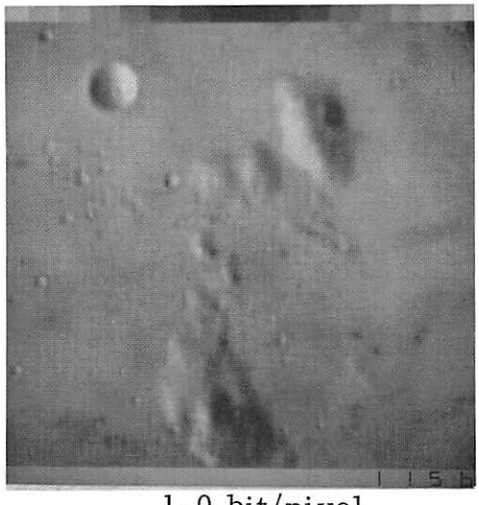
1.0 bit/pixel

1.5 bits/pixel

1.0 bit/pixel



1.5 bits/pixel



1.0 bit/pixel

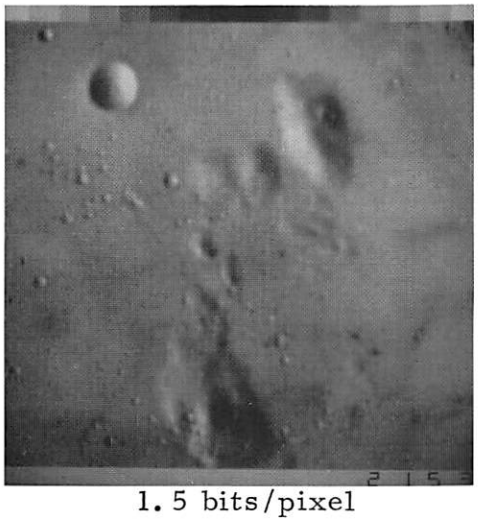
Figure 5-22. Slant Transform Zonal Sampling in 16 x 16 Pixel Blocks, Quantized Transform.

1.5 bits/pixel

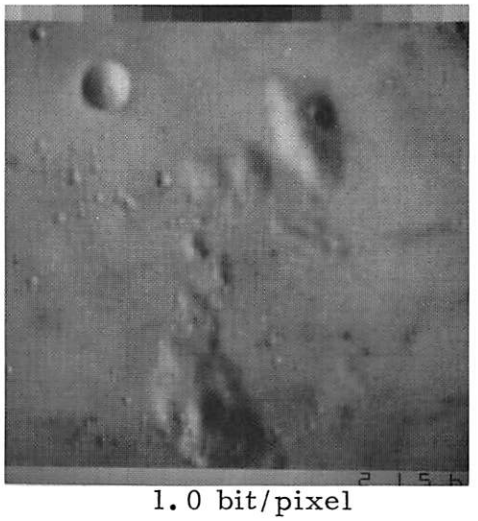
1.0 bit/pixel

1.5 bits/pixel

1.0 bit/pixel



1.5 bits/pixel



1.0 bit/pixel

Figure 5-23. Hadamard Transform Zonal Sampling in 16 x 16 Pixel Blocks, Quantized Transform.

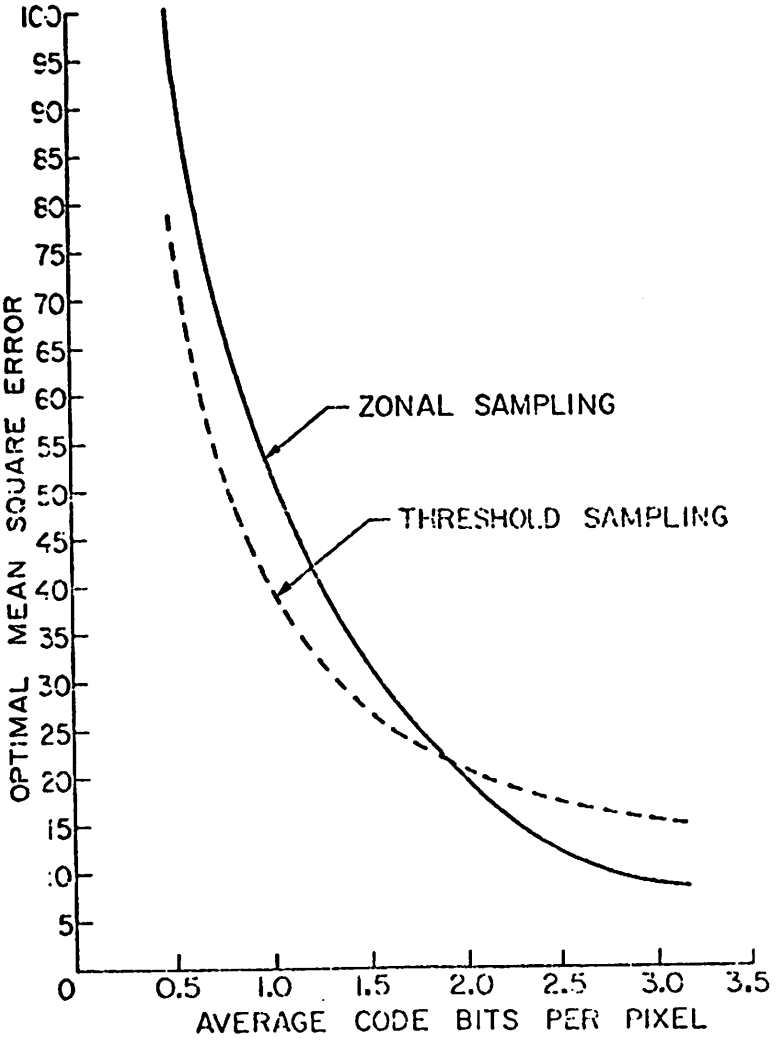


Figure 5-24. Comparison of the Optimal Threshold and Zonal Coding as a Function of Average Code Bits/ Pixel for the GIRL IMAGE.

6. EFFECT OF CHANNEL ERROR FOR SLANT TRANSFORM IMAGE TRANSMISSION

The preceding chapter has been devoted to a presentation of the slant transform for image coding in an error free channel. Since it is impossible for a practical communication system to possess such a channel, it is necessary to study the noise effects on the transform coded images. The inherent "error averaging" property of transform coding has provided a means of image coding for which channel errors are less deleterious than for conventional spatial coding of an image. This chapter presents the results of computer simulations of noise effects on the spatial and slant transform domain of the GIRL image. Simulations of noise effects are also made for the bandwidth compressed slant transform image of 1.5 bits/pixel for both threshold and zonal coding. As expected, it is found that zonal coding has the property of best noise immunity.

A binary symmetric channel is used as the noise model. Figure 6-1 illustrates a classical representation of such a communication channel where the probability of receiving an incorrect symbol is p and receiving a correct one is 1-p regardless of which symbol is transmitted.

6.1 Channel Error Effects - Without Bandwidth Compression

A major advantage of transmitting the transform rather than

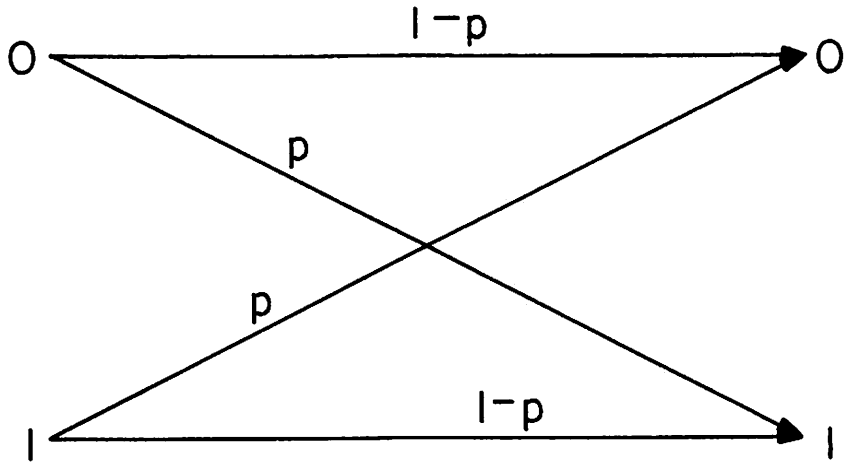


Figure 6-1. Model of a Binary Symmetric Channel.

the spatial domain of an image is the fact that the channel noise introduced in the transform domain tends to be distributed over the entire channel [14]. Since the eye is more sensitive to the "salt and pepper" effects of channel noise introduced in the spatial domain, the same channel error rate in the transform domain is not so offensive.

Figure 6-2 illustrates the effects of channel error on the spatial domain of the GIRL image with error rates of 10^{-4}, 10^{-3}, and 10^{-2}. The "salt and pepper" effect is quite evident in these pictures. Figure 6-3 illustrates the effects of channel error on the slant transform domain with the same error rates. Here the transform is performed in 16 x 16 pixel blocks and each transform sample is coded with 8 bits. It can be seen for error rates of less than 10^{-4}, the transformed image indicates little effect from the channel error. It can also be seen, however, for larger error rates the transformed image turns out to be "washed out". This can be explained by the fact that some of the bit assignment for the larger transform domain samples are reversed for the larger error rates.

6.2 Channel Error Effects - With Bandwidth Compression

Figure 6-4 and 6-5 illustrate channel error effects on the threshold and zonal coded GIRL image with a bandwidth reduction coding of 1.5 bits per pixel (No error correction has been attempted). It can be seen that threshold coding, which appears to be a bit better than zonal coding in the error free channel, is much more affected by

(a) $Pe=10^{-4}$

(b) $Pe=10^{-3}$

(c) $Pe=10^{-2}$

Figure 6-2. Spatial Domain Coding Effects of Channel Errors, 8 bits/pixel.

(a) $Pe=10^{-4}$

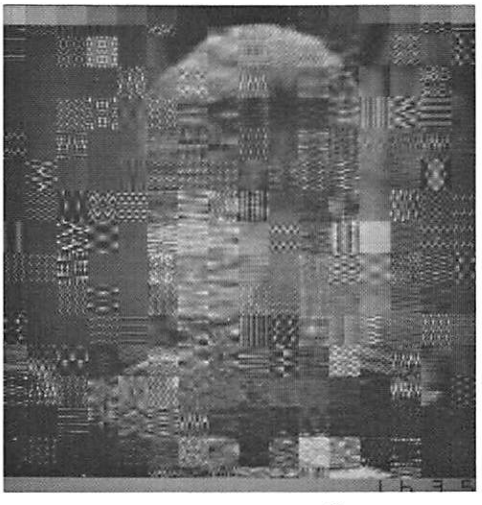
(b) $Pe=10^{-3}$

(c) $Pe=10^{-2}$

Figure 6-3. Slant Transform Coding Effects of Channel Errors, 8 bits/coefficient.

(a) $Pe=10^{-4}$

(b) $Pe=10^{-3}$



(c) $Pe=10^{-2}$

Figure 6-4. Slant Transform Threshold Coding Effects of Channel Errors, average coding of 1.5 bits/pixel.

(a) $Pe=10^{-4}$

(b) $Pe=10^{-3}$

(c) $Pe=10^{-2}$

Figure 6-5. Slant Transform Zonal Coding Effects of Channel Errors, average coding of 1.5 bits/pixel.

channel errors. This is inevitable because position coding is necessary in the threshold coding system. Once a position code bit is reversed, the entire image line thereafter will be incorrectly decoded which causes degradations in the reconstruction. Zonal coding appears to be much less affected by channel errors.

7. SLANT TRANSFORM COLOR IMAGE CODING

Figure 7-1 shows a block diagram of the slant transform color image coding system. In the system the color image is represented by three source tristimulus signals, $R(j,k)$, $G(j,k)$ and $B(j,k)$ which specify the red, green and blue content of an image pixel at spatial coordinate (j,k), according to the NTSC receiver phosphor primary system [26]. The source tristimulus signals are then converted to a new three dimensional space $Y(j,k)$, $I(j,k)$, and $Q(j,k)$ which specify the luminance and the chrominance information of the image pixel, according to the NTSC transmission primary system [26]. The converted signals then individually undergo a two dimensional slant transform over the entire image, or repeatedly over subsections of the image called blocks, resulting in three transform domain planes $F_Y(u,v)$, $F_I(u,v)$, and $F_Q(u,v)$. Next, a sample selection and quantization are performed on the three transform domains. The resultant quantized transform signals are then coded and transmitted over a channel. At the receiver, the channel output is decoded, and an inverse slant transform and inverse coordinate conversion operations are performed to reconstruct the source tristimulus signals.

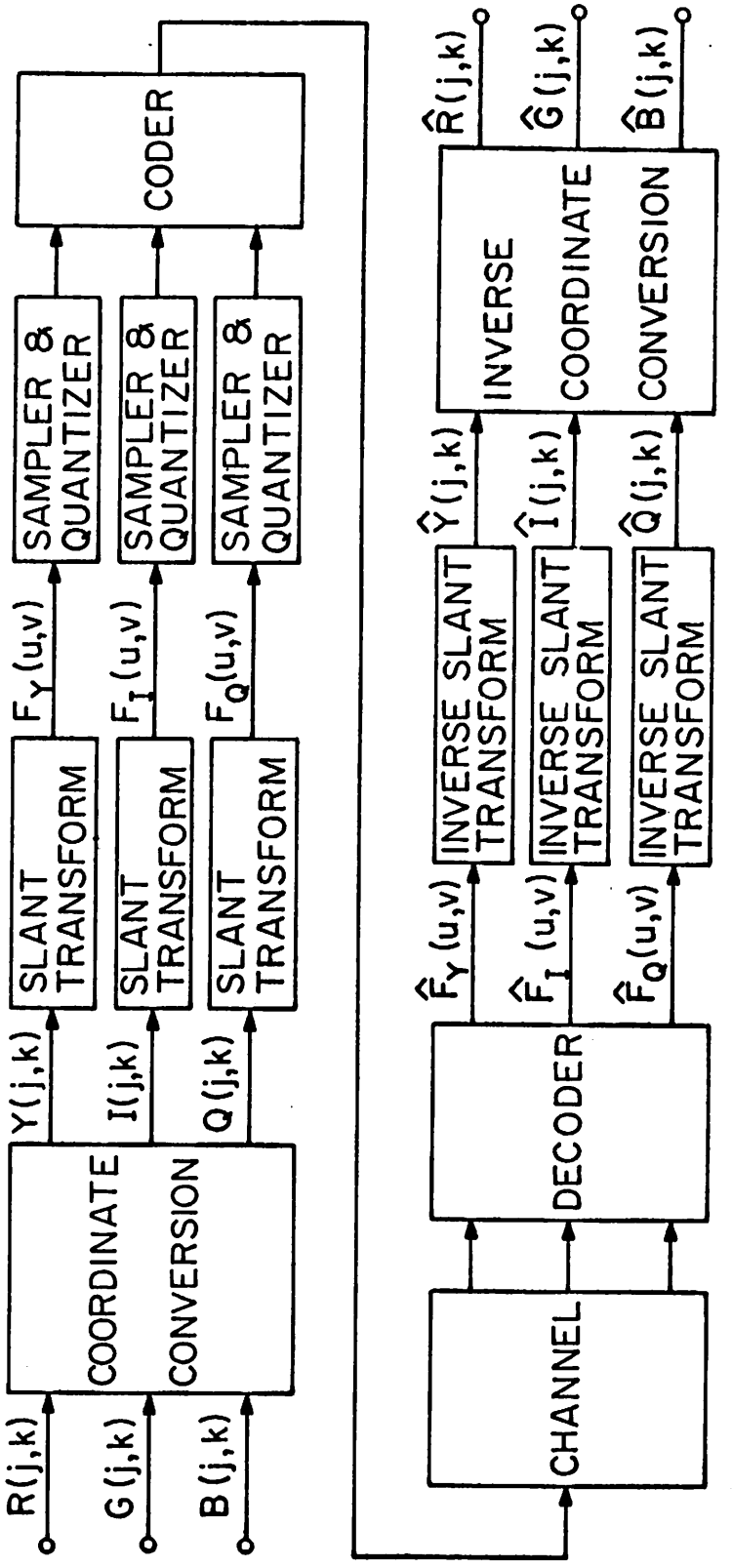


Figure 7-1. Slant Transform Color Image Coding System.

As discussed in Chapter 5, there are two methods of achieving a bandwidth reduction in transform coding system--threshold and zonal coding. The discussion in this chapter is limited to zonal color image transform coding. However, it is expected that, as in monochromatic image coding, threshold coding will perform slightly better than zonal coding. But, it is also expected that threshold coding will be affected by channel errors to a greater degree than zonal coding.

7.1 Color Image Coordinate Conversion

There are a number of coordinate systems which could be employed in color image coding and transmission [25]. The NTSC transmission primary coordinate system has been used in this study because:

(1) the NTSC transmission primary (YIQ) system is the U.S. standard for color television transmission.

(2) the (YIQ) system includes a luminance (monochrome) image signal.

(3) a previous study [24] has shown that the Y, I, and Q signals are less correlated in comparison with other standard coordinate systems so that they can be processed separately without much loss in coding performance.

Conversion of tristimulus values from the NTSC receiver phosphor primary system to NTSC transmission primary system can be mathematically expressed by the matrix equation [25]

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{7-1}$$

The inverse coordinate conversion is given by

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.621 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.106 & 1.703 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \tag{7-2}$$

7.2 Color Image Transform Coding

Consider the color image transform coding system of Figure 7-1. The color coordinate conversion of equation (7-1) can be rewritten here as

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ I &= 0.596R - 0.274G - 0.322B \\ Q &= 0.211R - 0.523G + 0.312B \end{aligned} \tag{7-3}$$

Now each of these color signal planes is separately slant transformed to produce three transform domain planes:

$$[F_Y] = [\mathcal{A}][Y][\mathcal{A}]^T$$

$$= 0.299\,[\mathcal{A}][R][\mathcal{A}]^T + 0.587[\mathcal{A}][R][\mathcal{A}]^T + 0.114$$

$$[\mathcal{A}][B][\mathcal{A}]^T$$

$$[F_I] = [\mathcal{A}][I][\mathcal{A}]^T \qquad (7\text{-}4)$$

$$= 0.596\,[\mathcal{A}][R][\mathcal{A}]^T - 0.274\,[\mathcal{A}][G][\mathcal{A}]^T - 0.322$$

$$[\mathcal{A}][B][\mathcal{A}]^T$$

$$[F_Q] = [\mathcal{A}][Q][\mathcal{A}]^T$$

$$= 0.211\,[\mathcal{A}][R][\mathcal{A}]^T - 0.523\,[\mathcal{A}][G][\mathcal{A}]^T + 0.312$$

$$[\mathcal{A}][B][\mathcal{A}]^T$$

It is apparent that the order of the color coordinate conversion and two dimensional forward transformation processes is immaterial.

Next, each transform plane undergoes a sample selection according to the "maximum variance zone" as introduced in Chapter 5. Those selected samples in each plane are then optimally quantized. Again, the quantization scheme used is the same as the monochromatic case where Rayleigh and Gaussian densities are introduced in deriving the quantization levels. The variations of quantization level are again tailored by a bit allocation matrix where the assignment of bits is proportional to the logarithm of variance of each transform domain sample and the percentage of energy

distributed in each color plane. Figure 7-2 illustrates a typical assignment of code bits for the slant transform samples of each plane in 16 x 16 sample blocks. Special attention must be paid to the quantization and coding of the F(0, 0) samples in $F_I(u, v)$ and $F_Q(u, v)$ planes. Since pixels of I(j, k) and Q(j, k) are no longer all positive, the F(0, 0) transform samples are not necessarily all positive. Therefore, a special code bit must be assigned for the sign of these samples.

At the receiver, the received samples are decoded into three transform planes and an inverse slant transform is performed in each plane to reconstruct three color signals:

$$\begin{aligned} [\hat{Y}] &= [\mathcal{S}]^T [\hat{F}_Y][\mathcal{S}] \\ [\hat{I}] &= [\mathcal{S}]^T [\hat{F}_I][\mathcal{S}] \\ [\hat{Q}] &= [\mathcal{S}]^T [\hat{F}_Q][\mathcal{S}] \end{aligned} \tag{7-5}$$

where $\hat{F}_Y$, $\hat{F}_I$, and $\hat{F}_Q$ denote three decoded transform planes at the receiver.

The inverse coordinate conversion at the final stage of the color image transform coding system, as defined by equation (7-2), is then performed to give

8	8	8	8	8	8	6	6	5	5	5	5	4	4	4	4
8	8	8	6	6	6	4	4	4	4	4	4	3	3	3	3
8	8	7	5	4	4	3	3	3	3	3	3	2	2	2	2
8	6	5	3	3	3	2	2	2	2	2	2	2	2	2	2
8	6	4	3	3	3	2	2	2	2	2	2	0	0	0	0
8	6	4	3	3	3	2	2	2	2	2	2	0	0	0	0
6	4	3	2	2	2	0	0	0	0	0	0	0	0	0	0
6	4	3	2	2	2	0	0	0	0	0	0	0	0	0	0
5	4	3	2	2	2	0	0	0	0	0	0	0	0	0	0
5	4	3	2	2	2	0	0	0	0	0	0	0	0	0	0
5	4	3	2	2	2	0	0	0	0	0	0	0	0	0	0
5	4	3	2	2	2	0	0	0	0	0	0	0	0	0	0
4	3	2	2	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	2	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	2	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	2	0	0	0	0	0	0	0	0	0	0	0	0

(a) $F_Y(u, v)$

8	8	6	4	4	4	3	3	2	2	2	2	2	2	2	2
8	5	4	3	2	2	2	2	0	0	0	0	0	0	0	0
6	4	3	2	2	2	0	0	0	0	0	0	0	0	0	0
4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(b) $F_I(u, v)$

8	6	5	3	3	3	2	2	2	2	2	2	0	0	0	0
6	4	3	2	2	2	0	0	0	0	0	0	0	0	0	0
5	3	2	2	0	0	0	0	0	0	0	0	0	0	0	0
3	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(c) $F_Q(u, v)$

Figure 7-2 Typical Bit Assignments for the Slant Transform Zonal Coding of Color Images.

$$\hat{R} = \hat{Y} + 0.956\hat{I} + 0.621\hat{Q}$$
$$\hat{G} = \hat{Y} - 0.272\hat{I} - 0.647\hat{Q} \qquad (7\text{-}6)$$
$$\hat{B} = \hat{Y} - 1.106\hat{I} + 1.703\hat{Q}$$

The energy compaction properties of $F_Y(u,v)$, $F_Z(u,v)$, and $F_Q(u,v)$ can be statistically evaluated if the covariance function of R, G, B is known. Consider the case in which the second order statistical variations of the tristimulus values R, G, B are spatially identical and described by covariance matrices $[C_{f_j}]$ and $[C_{f_k}]$ as denoted in equations (5-14) and (5-15). Suppose the covariance matrix of the source tristimulus value is given by

$$[C_{RGB}] = \begin{bmatrix} \sigma_R^2 & C_{RG} & C_{RB} \\ C_{RG} & \sigma_G^2 & C_{GB} \\ C_{RB} & C_{GB} & \sigma_B^2 \end{bmatrix} \qquad (7\text{-}7)$$

where σ_R^2, σ_G^2, σ_B^2 are the variances of the source tristimulus values, and C_{RG}, C_{RB}, C_{GB} are the correlations between pairs of the source tristimulus values. Then it can easily be shown that the covariance matrix of the Y, I, Q signals is

$$[C_{YIQ}] = \begin{bmatrix} \sigma_Y^2 & C_{YI} & C_{YQ} \\ C_{YI} & \sigma_I^2 & C_{IQ} \\ C_{YQ} & C_{IQ} & \sigma_Q^2 \end{bmatrix} \qquad (7\text{-}8)$$

where $\sigma_Y^2 = 0.0894\,\sigma_R^2 + 0.3446\,\sigma_G^2 + 0.0130\,\sigma_B^2$

$$+ 0.3510\,C_{RG} + 0.0682\,C_{RB} + 0.1338\,C_{GB}$$

$$\sigma_I^2 = 0.3552\,\sigma_R^2 + 0.0751\,\sigma_G^2 + 0.1037\,\sigma_B^2$$

$$- 0.3266\,C_{RG} - 0.3838\,C_{RB} + 0.1765\,C_{GB}$$

$$\sigma_Q^2 = 0.0445\,\sigma_R^2 + 0.0640\,\sigma_G^2 + 0.0973\,\sigma_B^2$$

$$- 0.1068\,C_{RB} + 0.1317\,C_{RG} - 0.1579\,C_{GB}$$

$$C_{YI} = 0.1728\,\sigma_R^2 - 0.1608\,\sigma_G^2 - 0.0367\,\sigma_B^2$$

$$+ 0.2678\,C_{RG} - 0.0283\,C_{RB} - 0.2193\,C_{GB}$$

$$C_{YQ} = 0.0631\,\sigma_R^2 - 0.1485\,\sigma_G^2 + 0.0356\,\sigma_B^2$$

$$+ 0.0482\,C_{RG} + 0.1179\,C_{RB} + 0.1540\,C_{GB}$$

$$C_{IQ} = 0.1258\,\sigma_R^2 + 0.0693\,\sigma_G^2 - 0.1005\,\sigma_B^2$$

$$- 0.2086\,C_{RG} + 0.1180\,C_{RB} - 0.040\,C_{GB}$$

From equation (5-15) the covariance matrices of rows and columns of each transform plane are

$$[C_{F_j}] = [\mathscr{A}][C_{f_j}][\mathscr{A}]^T \qquad (7\text{-}9)$$

$$[C_{F_k}] = [\mathscr{A}][C_{f_k}][\mathscr{A}]^T \qquad (7\text{-}10)$$

Therefore, the variance of the transform planes may be written as

$$\left[V_{F_Y}\right] = \sigma_Y^2 \left[V_{F_j}\right]\left[V_{F_k}\right]^T \tag{7-11}$$

$$\left[V_{F_I}\right] = \sigma_I^2 \left[V_{F_j}\right]\left[V_{F_k}\right]^T \tag{7-12}$$

$$\left[V_{F_Q}\right] = \sigma_Q^2 \left[V_{F_j}\right]\left[V_{F_k}\right]^T \tag{7-13}$$

where $\left[V_{F_j}\right]^T = \left[C_{F_j}(0,0) \quad C_{F_j}(1,1) \quad \ldots \quad C_{F_j}(N-1,N-1)\right]$

$$\left[V_{F_k}\right]^T = \left[C_{F_k}(0,0) \quad C_{F_k}(1,1) \quad \ldots \quad C_{F_k}(N-1,N-1)\right]$$

A summary of the energy distribution between the color signal planes of R, G, B and Y, I, Q, for the GIRL and COUPLE test images, is given in Table 7-1. It can easily be seen that the YIQ system provides a better energy compaction in comparison with the RGB system.

In order to optimally design the slant transform image coder it is necessary to specify some analytic measure of color image fidelity. Unfortunately, there exists no standard fidelity measures. As a rational alternative, the design procedure selected has been to design the transform domain quantization system to minimize the mean square error between the Y,I,Q and $\hat{Y},\hat{I},\hat{Q}$ color

Test Image	Coordinate System	Percentage of σ_1^2	Percentage of σ_2^2	Percentage of σ_3^2
GIRL	RGB	45.14	35.41	19.45
	YIQ	78.32	17.54	4.14
COUPLE	RGB	51.55	31.09	17.36
	YIQ	84.84	13.81	1.35

Table 7-1 Energy Distribution of Color Planes.

planes. From eq. (5-40) the minimal mean square error in each transform plane can be written as

$$\epsilon_{T_i} = \frac{1}{N^2} \left\{ \int_0^{\infty} F^2 \, p_{1_i}(F) \, dF - \sum_{K=1}^{L_{1_i}} F_K^2(0,0) \int_{Q_{K-1}(0,0)}^{Q_K(0,0)} p_{1_i}(F) \, dF \right.$$

$$+ 2 \sum_{\substack{u \\ (u,v)\neq(0,0)}} \sum_{\substack{v \\ S_i(u,v)=1}} \left[\int_0^{\infty} F^2 \, p_{2_i}(F) \, dF - \sum_{K=1}^{L_{2_i}(u,v)/2} F_K^2(u,v) \int_{Q_{K-1}(u,v)}^{Q_K(u,v)} p_{2_i}(F) \, dF \right]$$

$$\left. + \sum_{\substack{u \\ S_i(u,v)=0}} \sum_{v} E\left[F_i^2(u,v)\right] \right\} \tag{7-11}$$

where ϵ_{T_i} is the mean square error in each transform plane;

$S_i(u,v)$ is a sampling function in each transform plane;

p_{1_i} and p_{2_i} are probability density functions of dc and ac samples;

L_{1_i} and $L_{2_i}(u,v)$ are numbers of dc and ac quantization levels;

$Q_K(0,0)$ and $Q_K(u,v)$ are dc and ac decision levels;

$F_K(0,0)$ and $F_K(u,v)$ are dc and ac reconstruction levels;

and $F_i(u,v)$ are transform samples of a plane.

The total optimal relative mean square error then may be defined as

$$\epsilon_T = \frac{\epsilon_{T_1} + \epsilon_{T_2} + \epsilon_{T_3}}{3} \qquad (7\text{-}12)$$

Figure 7-3 contains a plot of the mean square error versus the average code bit assignments of $F_Y(u,v)$ and a fixed total average code of 2 bits/pixel for the GIRL image. The optimal average bit allocation for this test image is the maximum point on an envelope which is constructed by drawing through each peak point of the curves shown. The value is found to be: 1.25 for $F_Y(u,v)$, 0.55 for $F_I(u,v)$, and 0.20 for $F_Q(u,v)$. The optimum scale does not change appreciably for the other image.

7.3 Experiment Results

A computer simulation has been performed to subjectively evaluate the performance of the slant transform color image coding system.

Figure 7-4 contains black-and-white photographs of the R, G, B components of the original images shown in figure 1-2. It can be seen that the R, G, B pixels are highly correlated. Figure 7-5 illustrates black-and-white versions of the Y, I, Q planes of the same images. It is clearly seen that the degree of correlation among these planes is much less than those in R, G, B planes.

The energy compaction properties of the slant transformed

Y, I and Q planes can be seen from pictures of the three transform domain samples. Figure 7-6 shows the slant transform planes, $F_Y(u,v)$, $F_I(u,v)$, and $F_Q(u,v)$, where the transform is performed in 16 x 16 pixel blocks over the entire plane. Again, in these pictures, a logarithm has been taken for each transform sample in order to compact the dynamic range of the transform samples. It can easily be seen that most of the significant samples in each plane are located around the maximum variance zone. It can also be seen that the energy distribution in $F_I(u,v)$ and $F_Q(u,v)$ planes is quite small in comparison with that of $F_Y(u,v)$.

To illustrate the bandwidth reduction capability of the slant transform for color image coding, two sets of experiments have been performed for both the GIRL and COUPLE images. The first set of experiments results in an average coding of 2 bits/pixel where the transform samples of Y,I, and Q are coded with 1.2, 0.54, and 0.26 bits/pixel respectively. The second set of experiments results in an average coding of 3 bits/pixel where the transform samples of Y, I, and Q are coded, respectively, with 2.0, 0.6, and 0.4 bits/pixel. The corresponding reproductions of Y,I,Q and R,G,B for the first set of experiments are shown in figure 7-7 and 7-8. Figure 7-9 and 7-10 show the reproductions of the color images with channel error rates of Pe = 0, Pe = 10^{-4} and Pe = 10^{-2}. It can be seen that even with an average coding of 2 bits/pixel and channel error rate

Total Average Code Bits	Color Plane	Average Code Bits	MSE Without Quantization	MS Quantization Error	Total MSE of Each Color Plane
2 Bits/Pixel	Y	1.20	0.0458%	0.0254%	0.0712%
	I	0.54	0.0328%	0.0133%	0.0461%
	Q	0.26	0.0141%	0.0057%	0.0198%
3 Bits/Pixel	Y	2.00	0.0163%	0.0136%	0.0299%
	I	0.60	0.0327%	0.0102%	0.0429%
	Q	0.40	0.0127%	0.0043%	0.0170%

Table 7-2 Mean Square Error Between the YIQ GIRL and the Slant Transform Coded $\hat{Y}\hat{I}\hat{Q}$ GIRL.

$$\text{M.S.E.} = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left[f(j,k) - \hat{f}(j,k) \right]^2; \quad 0 \le f(j,k) \le 1$$

of $Pe = 10^{-4}$, the result can still be considered as a good quality reconstruction. Table 7-2 exhibits the mean square errors between the original Y,I,Q planes and slant transform coded Y,I,Q planes for the GIRL image. The reason that the mean square errors of YIQ is presented rather than the mean square errors of RGB is that the bandwidth compression has been made only to the YIQ signals.

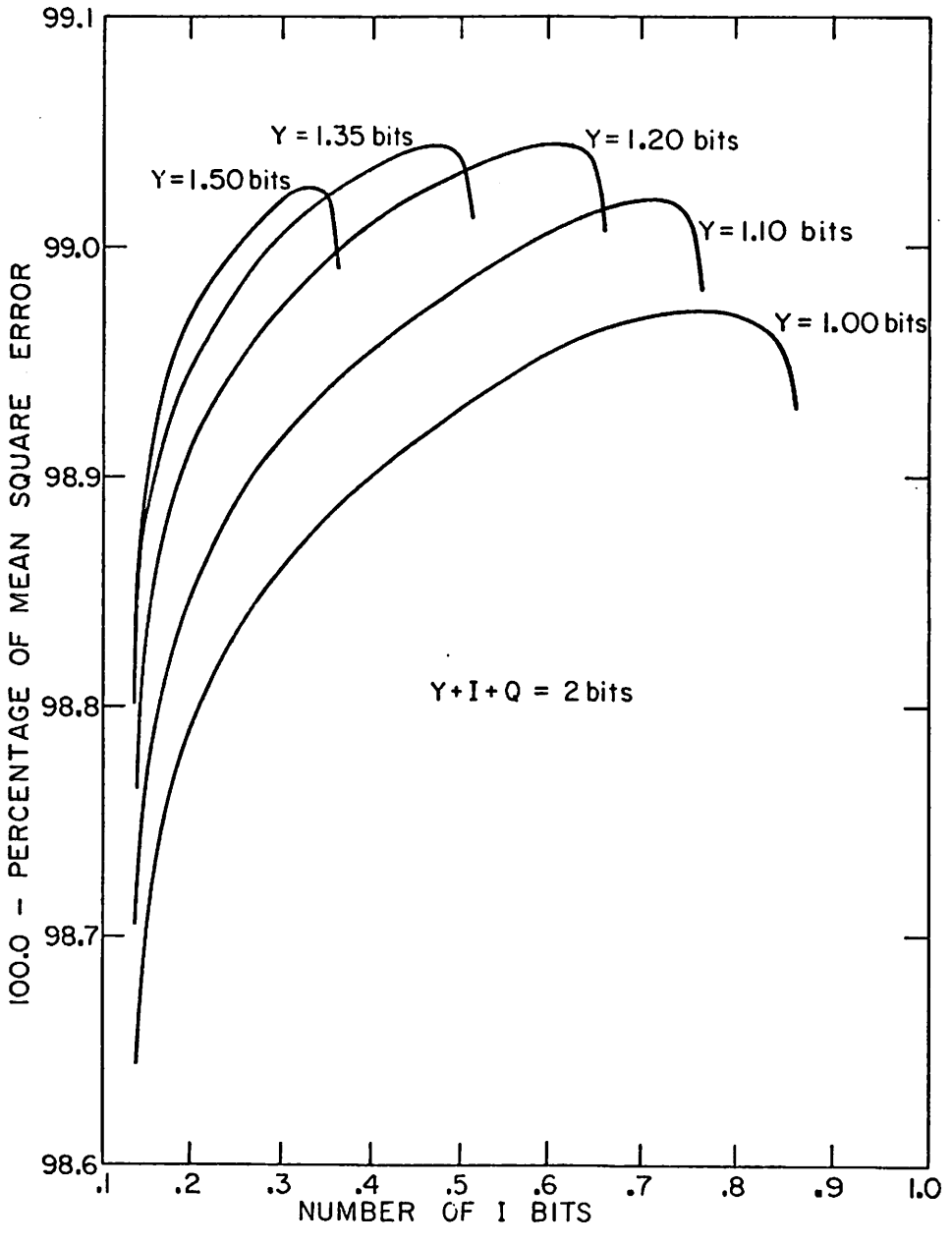


Figure 7-3. Mean Square Error for Various Color Plane Bit Assignments.

(a)

(b)

(c)

Figure 7-4. R, G, B Tristimulus Color Planes of the Original Images: (a) R; (b) G; (c) B.

Figure 7-5. Y, I, Q Tristimulus Color Planes of the Original Images: (a) Y; (b) I; (c) Q.

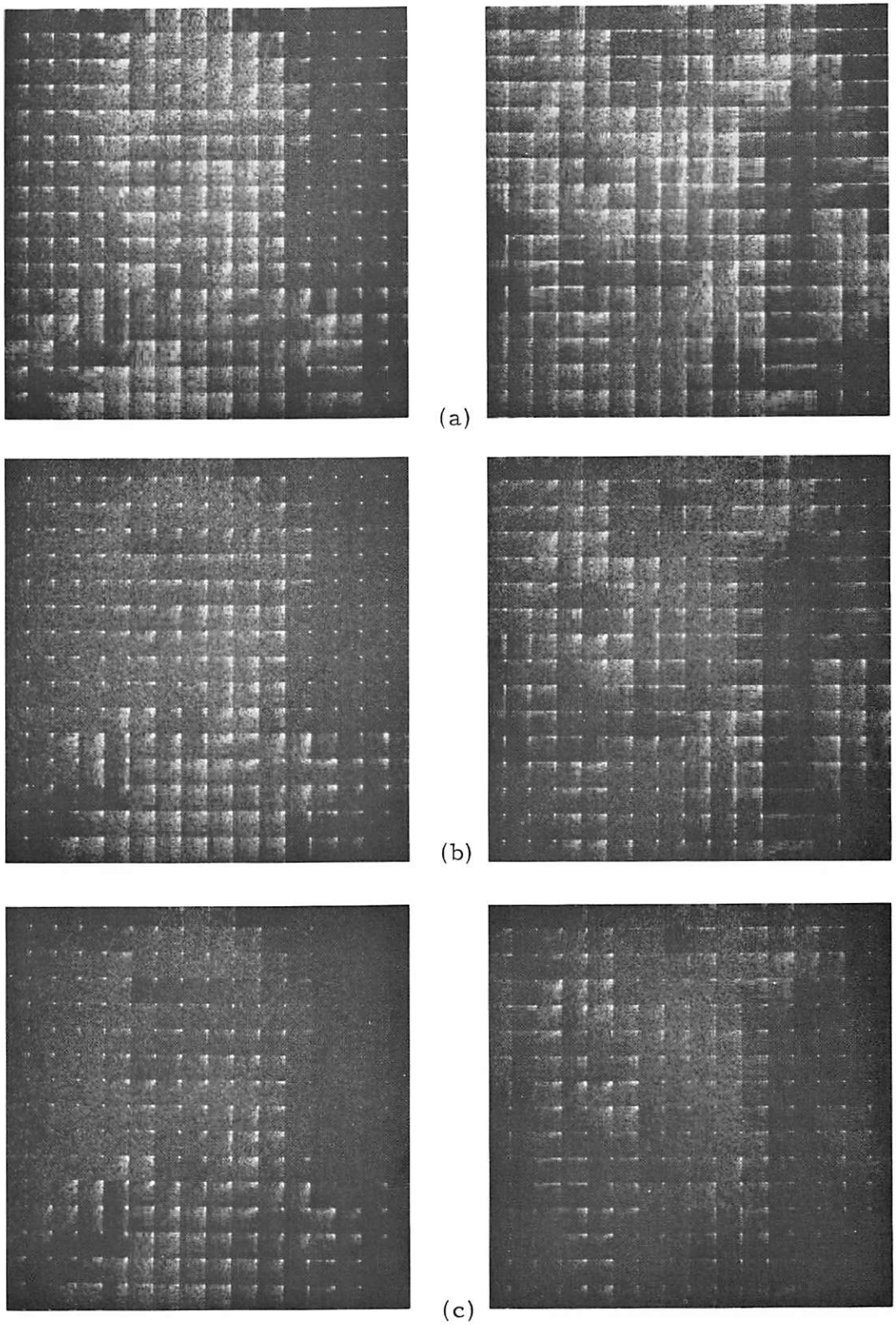


Figure 7-6. Slant Transform Domain of Y, I, Q Images: (a) F_Y; (b) F_I; (c) F_Q.

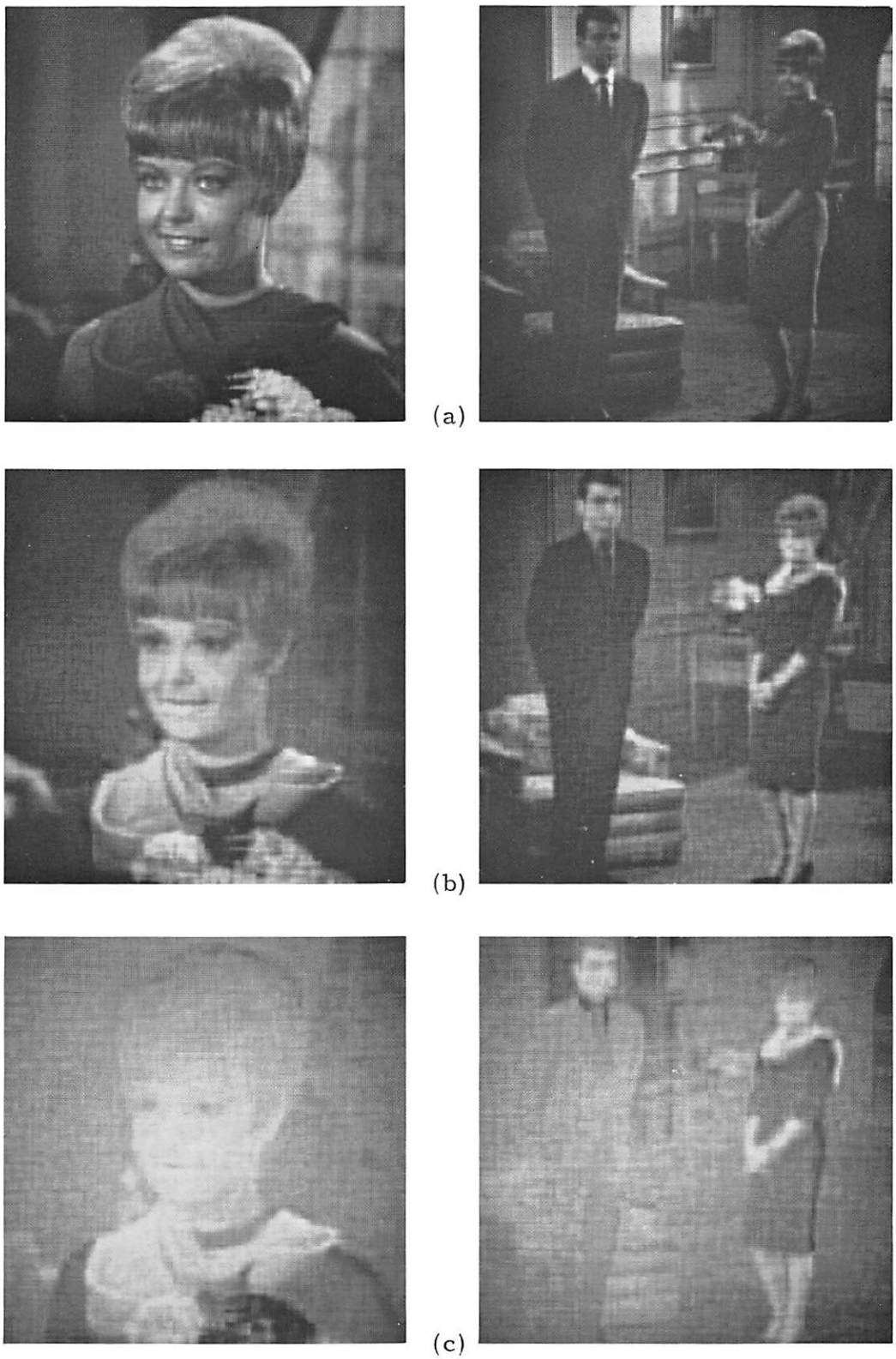
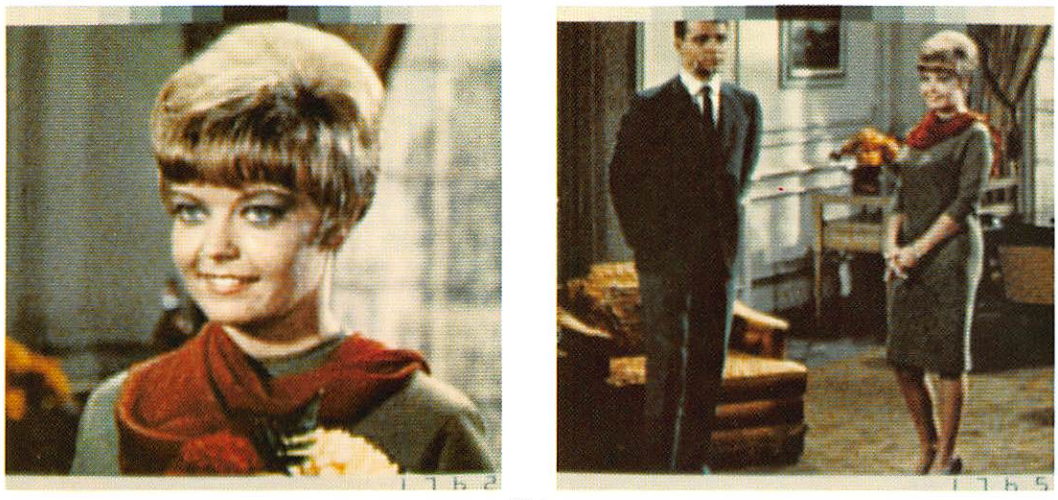


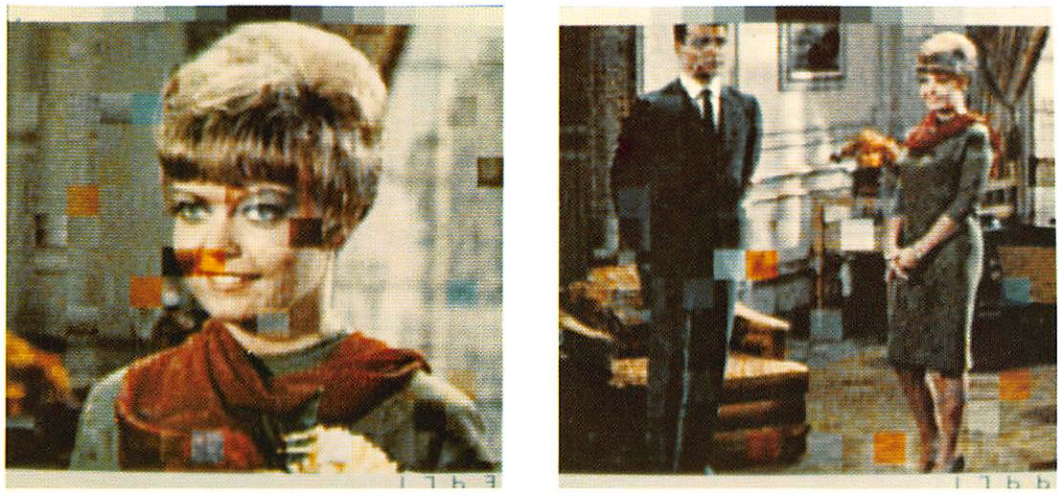
Figure 7-7. Slant Transform Zonal Coding, 2 bits/pixel: (a) $\hat{Y}$; (b) $\hat{I}$; (c) $\hat{Q}$.

Figure 7-8. Slant Transform Zonal Coding, 2 bits/pixel: (a) $\hat{R}$; (b) $\hat{G}$; (c) $\hat{B}$.

(a)



(b)



(c)

Figure 7-9. Slant Transform Zonal Coding, 2 bits/pixel. Channel Error Rates: (a) Pe=0; (b) Pe=10^{-4}; (c) Pe=10^{-2}.

Figure 7-10. Slant Transform Zonal Coding, 3 bits/pixel. Channel Error Rates: (a) Pe=0; (b) Pe=10^{-4}; (c) Pe=10^{-2}.

8. SUMMARY

This dissertation has presented a theoretical development and analysis of the two dimensional slant transform for image coding. Various transforms which possess an energy compaction property have also been briefly discussed.

The implementation of the slant transform coding system appears feasible using the fast transform algorithm developed in this dissertation. It has been found that for a slant transform of order N, the total number of operations is N log N + 2N-4, which is only slightly greater than the number of operations required for the Hadamard transform.

The statistical properties of the slant transform have been analyzed based upon the assumption that the original image is a sample of a two dimensional process with known mean and covariance. The probability density functions of the transform samples have been modelled as a Rayleigh density function for dc samples and as a Gaussian density for ac samples.

The energy compaction property of the slant transform has been exploited to achieve a sample reduction by two means: threshold sampling and zonal sampling. Threshold sampling simply entails the coding of each transform sample that exceeds a threshold

level. In zonal sampling those samples with the largest expected variance are coded. For purposes of comparison experiments have also been performed on the Fourier, Hadamard, Haar, and Karhunen-Loeve transforms. The conclusions are:

a) A significant sample reduction with slant transform coding can be obtained by threshold and zonal sampling;

b) Threshold sampling provides a better performance than zonal sampling for the same sample reduction factor. A higher sample reduction can be obtained with threshold sampling without seriously degrading the reconstruction.

c) The slant transform exhibits somewhat better performance than the Fourier, Hadamard, or Haar transforms.

d) For block sizes larger than 16 x 16, the improvement of performance will not be significant.

An analysis has been performed to determine the quantization effect of transform domain samples. A mathematical expression of the mean square error between the original image and its transform coded reconstruction has been derived.

Coding techniques for optimally quantized transform samples have been implemented and evaluated for both threshold and zonal processed samples. For threshold processed samples, a position coding technique employing runlength coding has been introduced. For zonal processed samples a maximum variance zonal coding

technique has been introduced. The maximum variance zonal coding technique appears to be much easier to implement than runlength coding, since no position coding is required.

The effect of channel errors on slant transform coded images has been simulated on a digital computer. All code bits were packed into a long string of binary data, and a random noise generator was introduced to generate bit reversals according to a specific error rate. It has been shown that channel errors in the transform domain tend to cause a small overall loss in resolution which is perferable to the "salt and pepper" errors introduced in spatial domain coding. Comparing the effect of channel errors on threshold and zonal coding techniques, it appears that zonal coding is less sensitive to channel errors. It has been shown that zonal coding can tolerate an error rate as high as $Pe = 10^{-3}$ for a bandwidth reduced image of 1.5 bits/pixel without serious image degradation.

The studies of the slant transform for monochrome image coding have also been extended to color images. First, a coordinate conversion from RGB to YIQ has been made in order to compact the image energy between color planes. Next, the slant transform coding technique has been applied to each color plane. It has been shown that relatively large bandwidth reductions may be obtained in the I and Q planes without seriously effecting the color image reconstruction.

The conclusions to be drawn from experiments that have been performed in monochrome and color image coding are:

(a) A bandwidth reduction of 1 bit/pixel and 1.5 bits/pixel can be made for a monochrome image by threshold and zonal coding techniques respectively without seriously degrading the reconstruction quality.

(b) For color images, a total coding of about 2 to 3 bits/pixel can be realized while maintaining good quality reconstructions.

Appendix

SLANT TRANSFORM PROGRAMS

This appendix presents the programs of the one-dimensional forward and inverse slant transform of size N=256. The transform of sizes other than 256 can be obtained by a minor modification of these programs.

(a) Forward Slant Transform:

```
C       THIS PROGRAM PERFORMS ONE DIMENSIONAL
C       SLANT TRANSFORM OF A 256 BY 256 IMAGE
C
        DIMENSION A(256), B(256), C(256)
        N=256
        EN=N
        ENN=SQRT(EN)
        B1=1./SQRT(5.)
        A1=3.*B1
        DO 79 M=1,N
        CALL DSKIO(C, 1024, M, 1, 4)
C
        K1=N/4
        DO 40 II=1,K1
        IIA=4*(II-1)
        DO 11 I=1,4
        J=IIA+I
 11     B(I)=C(J)
C
        A(1)=B(1)+B(4)
        A(2)=B(2)+B(3)
        A(3)=B(1)-B(4)
        A(4)=B(2)-B(3)
        B(1)=A(1)+A(2)
        B(2)=A1*A(3)+B1*A(4)
        B(3)=A(1)-A(2)
        B(4)=B1*A(3)-A1*A(4)
```

```
      DO 40 I=1,4
      J=IIA+I
 40   C(J)=B(I)
C
      K=8
      L=3
      GO TO 49
 41   K=16
      L=4
      GO TO 49
 42   K=32
      L=5
      GO TO 49
 43   K=64
      L=6
      GO TO 49
 44   K=128
      L=7
      GO TO 49
 45   K=256
      L=8
 49   KK=K/2
      K2=N/K
      LL=L-1
      SUM1=0.
      DO 61 I=1,LL
 61   SUM1=SUM1+2.**(2*(I-1))
      SUM2=SUM1+2.**(2*LL)
      A2=FLOAT(KK)/SQRT(SUM2)
      B2=SQRT(SUM1)/SQRT(SUM2)
      DO 65 II=1,K2
      IIB=K*(II-1)
      DO 51 I=1,K
      J=IIB+I
 51   B(I)=C(J)
      DO 59 I=1,KK
      DO 59 L2=1,2
      T=0.
      DO 57 L3=1,2
      I1=KK*(L3-1)+I
      IF (L3+L2-4) 55,53,55
 53   T=T-B(I1)
      GO TO 57
 55   T=T+B(I1)
```

```
 57   CONTINUE
      IF (IAND(I, 1)) 52, 54, 52
 52   L1=2*(I-1)
      I2=L1+L2
      GO TO 59
 54   L1=2*I+1
      I2=L1-L2
 59   A(I2)=T
C
      DD=A(4)
      A(4)=A(3)
      A(3)=A(2)
      A(2)=DD
C
      E2=B2*A(2)+A2*A(3)
      F2=A2*A(2)-B2*A(3)
      A(2)=E2
      A(3)=A(4)
      A(4)=F2
C
 63   DO 65 I=1, K
      J=IIB+I
 65   C(J)=A(I)
      IF(K-16)41, 42, 67
 67   IF(K-64) 43, 44, 68
 68   IF(K-256) 45, 71, 71
 71   DO 75 I=1, N
 75   C(I)=C(I)/ENN
      CALL DSKIO(C, 1024, M, 0, 5)
 79   CONTINUE
      STOP
      END
```

(b) Inverse Slant Transform:

```
C     THIS PROGRAM PERFORMS ONE DIMENSIONAL
C     INVERSE SLANT TRANSFORM OF A 256 BY 256
C     IMAGE
C
      DIMENSION A(256), B(256), C(256)
      N=256
      S=SQRT(FLOAT(N))
      B1=1./SQRT(5.)
      A1=3.*B1
```

```
      DO 179 M=1,N
      CALL DSKIO(C,1024,M,1,4)
      K=256
      L=8
      GO TO 149
   22 K=128
      L=7
      GO TO 149
   33 K=64
      L=6
      GO TO 149
   44 K=32
      L=5
      GO TO 149
   55 K=16
      L=4   GO TO 149
  141 K=8
      L=3
  149 KK=K/2
      K2=N/K
      LL=L-1
      SUM1=0.
      DO 161 I=1,LL
  161 SUM1=SUM1+2.**(2*(I-1))
      SUM2=SUM1+2.**(2*LL)
      A2=FLOAT(KK)/SQRT(SUM2)
      B2=SQRT(SUM1)/SQRT(SUM2)
      DO 165 II=1,K2
      IIQ=K*(II-1)
      DO 151 I=1,K
      J=I+IIQ
  151 B(I)=C(J)
C
      E2=B2*B(2)+A2*B(4)
      F2=A2*B(2)-B2*B(4)
      B(2)=E2
      B(4)=B(3)
      B(3)=F2
      JK=KK+1
      JKK=JK+1
      A(1)=B(1)+B(3)
      A(2)=B(2)+B(4)
      A(JK)=B(1)-B(3)
      A(JKK)=B(2)-B(4)
      DO 159 I=3,KK
```

```
      IJ=2*I-1
      DO 159 L2=1,2
      T=0.
      DO 157 L3=1,2
      I1=L3-1+IJ
      IF(L2+L3-4) 155,153,155
  153 T=T-B(I1)
      GO TO 157
  155 T=T+B(I1)
  157 CONTINUE
      I2=KK*(L2-1)+I
  159 A(I2)=T
C
      JJK=KK+4
      DO 160 IC=JJK,K,2
  160 A(IC)=-A(IC)
C
      DO 165 I=1,K
      J=I+IIQ
  165 C(J)=A(I)
      IF(K-128) 66,33,22
   66 IF(K-32) 77,55,44
   77 IF(K-8) 99,105,141
  105 K1=N/4
      DO 144 II=1,K1
      IIQ=4*(II-1)
      DO 111 I=1,4
      J=I+IIQ
  111  B(I)=C(J)
      A(1)=B(1)+B(3)
      A(2)=B(1)-B(3)
      A(3)=A1*B(2)+B1*B(4)
      A(4)=B1*B(2)-A1*B(4)
      B(1)=A(1)+A(3)
      B(2)=A(2)+A(4)
      B(3)=A(2)-A(4)
      B(4)=A(1)-A(3)
      DO 144 I=1,4
      J=I+IIQ
  144 C(J)=B(I)/S
      CALL DSKIO(C,1024,M,0,5)
  179 CONTINUE
   99 STOP
      END
```

Appendix B

PHOTOGRAPHIC DENSITY IMAGE REPRESENTATION

Results of an experiment of using the photographic density rather than intensity for the slant transform image coding system are presented in this appendix. The original digital image, $f(j,k)$, defined as an array of samples of a continuous two dimensional intensity pattern of light, has been converted to the photographic density by [36]

$$f'(j,k) = \log\{f(j,k)\} \qquad j,k = 0,1,2,\ldots,N-1$$

The samples of this conversion then undergo a two dimensional slant transformation repeated over subsections of images in 16 x 16 pixel blocks. The resultant transform samples are then quantized, coded, and transmitted over a computer simulated channel. At the receiver the received samples are decoded, inverse transformed, and reconverted back to the photographic intensity by

$$f(j,k) = 10^{f'(j,k)}$$

Figure A shows the result of this experiment. Comparing this with figure 4-35(a) it appears that the conversion of the photographic density does not have any improvement for the slant transform image coding system subjectively.

1.5 bits/pixel

Figure A. Slant Transform Zonal Coding of GIRL Image; transform was performed in 16 x 16 pixel blocks to the photographic density rather than intensity.

Appendix C

RATE DISTORTION MEASURE FOR TRANSFORM CODING

The rate distrotion function of information theory has proven to be a useful measure of the performance of source coding methods [19]. It has been shown that for a Gaussian source of independent symbols and a mean square error fidelity criterion, the minimum information rate R(D) that can be achieved while maintaining a fixed distortion D is given parametrically by [19]

$$R(D) = \frac{1}{2N} \sum_{i=1}^{N} \max \left\{ 0, \log \left(\frac{\sigma_i}{\theta} \right) \right\} \qquad \text{(C-1)}$$

$$D = \frac{1}{N} \sum_{i=1}^{N} \min (\theta, \sigma_i) \qquad \text{(C-2)}$$

where σ_i^2 is the variance of the i th sample and N represents the number of symbols in a block. This result can be applied to transform coding by treating the transform coefficients as being independent (note: the coefficients are quantized and coded separately), and by observing the probability density of the samples is well modelled by a Gaussian density. Thus, the factor σ_i in eqs. (C-1) and (C-2) can be assumed to be the standard deviation of the transform coefficients as given by eq. (4-17) or (4-18).

Figure C-1 illustrates rate versus distortion curves for a first order Markov process, with correlation coefficient $\rho = 0.95$ and $N = 16$. The curves show that the Karhunen-Loeve transform gives the best rate over the whole range of distortion, while the slant transform result is very close to the Karhunen-Loeve transform. Figure C-2 contains rate versus distortion curves for the same Markov process with $N = 256$. The curves show that the Fourier transform tends to become the Karhunen-Loeve transform for a large size data vector.

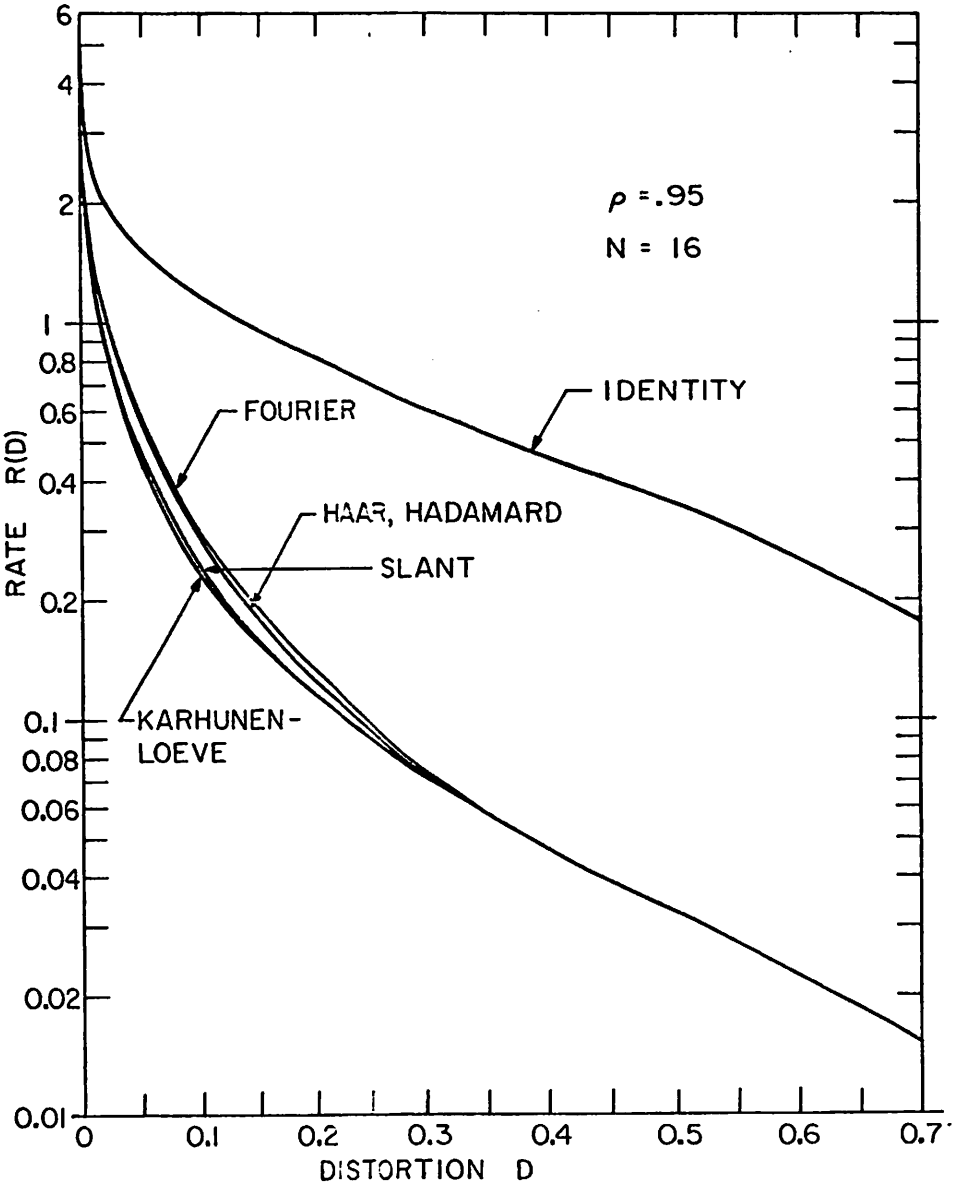


Figure C-1. Rate Versus Distortion for Various Transforms for a First-Order Markov Process, N=16.

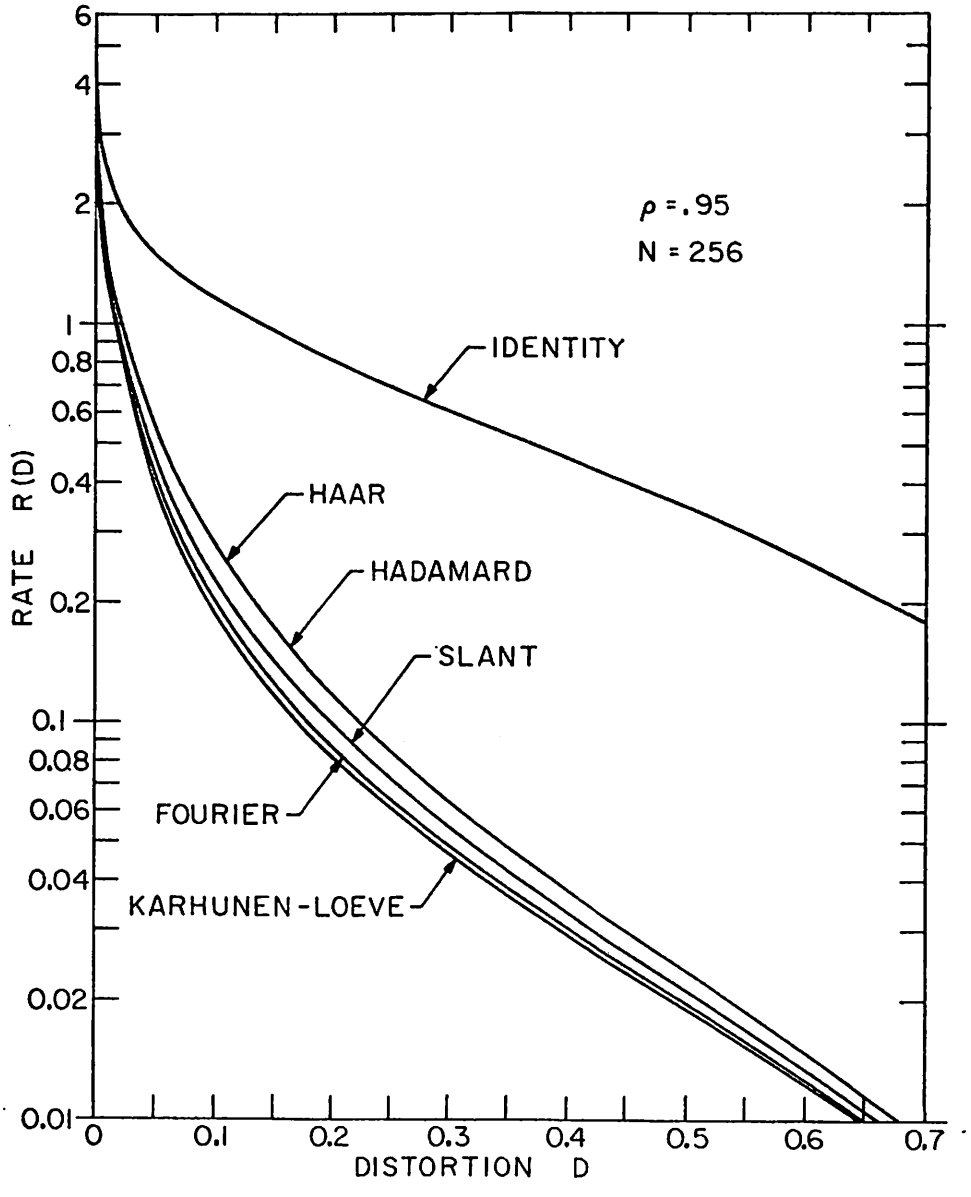


Figure C-2. Rate Versus Distortion for Various Transforms for a First-Order Markov Process, N=256.

Appendix D

AN APPROXIMATE METHOD OF MINIMIZING THE QUANTIZATION ERROR AND FINDING THE DECISION LEVELS OF THE DC AND AC TRANSFORM SAMPLES

This appendix presents a method of minimizing the quantization error and finding the decision levels of the dc and ac transform samples. The method described is similar to the Panter and Dite [20]. Let F and F(u,v) represent the dc and ac transform samples respectively. Suppose F is quantized into L_i levels by a set of decision levels, Q_K, and a set of reconstruction levels, F_k, as shown in figure 5-14. Then the dc quantization error can be written as

$$\epsilon_1 = \sum_{K=1}^{L_1} \int_{Q_{K-1}}^{Q_K} [F-F_K]^2 \, p_1(F) \, dF \qquad (D-1)$$

where p_1 is the probability density of the random variable F. In order to minimize ϵ_1 as shown in (D-1) consider

$$\epsilon_K = \int_{Q_{K-1}}^{Q_K} (F-F_K)^2 \, p_1(F) \, dF \qquad (D-2)$$

Suppose $p_1(F)$ is nearly constant over the region of integral such that

$$p_1(F) = p_1\left(\frac{Q_K + Q_{K-1}}{2}\right)$$

then

$$\epsilon_K = \frac{p_1\left(\frac{Q_K + Q_{K-1}}{2}\right)}{3}\left[(Q_K - F_K)^3 + (F_K - Q_{K-1})^3\right] \qquad \text{(D-3)}$$

Differentiating ϵ_K with respect to F_K and setting the result to zero gives

$$\frac{d\epsilon_K}{dF_K} = p_1\left(\frac{Q_K + Q_{K-1}}{2}\right)\left[-(Q_K - F_K)^2 + (F_K - Q_{K-1})^2\right] = 0 \qquad \text{(D-4)}$$

or

$$F_K = \frac{Q_K + Q_{K-1}}{2} \qquad \text{(D-5)}$$

Let $Q_K - F_K = F_K - Q_{K-1} = \Delta F_K$ (D-6)

Then ϵ_K may be written as

$$\epsilon_K = \frac{2p_1(F_K)\,\Delta F_K^3}{3} \qquad \text{(D-7)}$$

Substituting this into equation (D-1) gives

$$\epsilon_1 = \frac{2}{3}\sum_{K=1}^{L_1} X_K^3 \qquad \text{(D-8)}$$

where $W_K = \left[p_1(F_K)\,\Delta F_K\right]^{1/3}$

If the number of quantization levels is sufficiently large so that the definition of the integral is applicable, then

$$\sum_{K=1}^{L_1} X_K = \sum_{K=1}^{L_1} \left[p_1(F_K)\right]^{1/3} \Delta F_K = \frac{1}{2} \int_0^A \left[p_1(F)\right]^{1/3} dF = K_1 \tag{D-9}$$

where A_1 is the maximum decision level of F and K_1 is a constant since the integral is a function of only its limit. The minimization of equation (D-1) is now reduced to the minimization of equation (D-8) subject to the constraint of equation (D-9). Using Lagrange's method it is easily found that ϵ_1 is minimum when

$$X_1 = X_2 = \cdots = X_{L_1} = \frac{K_1}{L_1} \tag{D-10}$$

By substituting equations (D-10) and (D-9) into (D-8) the minimum quantization error becomes

$$\epsilon_1 = \frac{1}{12 L_1^2} \left(\int_0^{A_1} \left[p_1(F)\right]^{1/3} dF \right)^3 \tag{D-11}$$

Suppose now that each ac sample F(u, v) is quantized into $L_2(u, v)$ levels by a set of decision levels, $Q_K(u, v)$, and a set of reconstruction levels, $F_K(u, v)$, as shown in figure 5-15. Then the ac quantization error can be written as

$$\epsilon_2 = \sum_{K=\frac{-L_2(u,v)}{2}}^{\frac{L_2(u,v)}{2}} \int_{Q_{K-1}(u,v)}^{Q_K(u,v)} [F(u,v)-F_K(u,v)]^2 p_2\{F(u,v)\}\, dF(u,v) \qquad \text{(D-12)}$$

where p_2 is the probability density of $F(u,v)$. Following a similar method as introduced in equations (D-2) to (D-8), the minimization of equation (D-12) may be reduced to minimize

$$\epsilon_2 = \frac{2}{3} \sum_{K=\frac{-L_2(u,v)}{2}}^{\frac{L_2(u,v)}{2}} [\, Y_K(u,v)]^3 \qquad \text{(D-13)}$$

subject to the condition

$$\sum_{K=\frac{-L_2(u,v)}{2}}^{\frac{L_2(u,v)}{2}} Y_K = \frac{1}{2} \int_{-A_2(u,v)}^{A_2(u,v)} p_2\{F(u,v)\}\, dF(u,v) = K_2(u,v) \qquad \text{(D-14)}$$

where $Y_K = \left[p_2[F_K(u,v)]\right]^{1/3} \Delta F_K(u,v)$; $A_2(u,v)$ is the maximum decision level of $F(u,v)$, and $K_2(u,v)$ is a constant. Using Lagrange's method it is easily found that ϵ_2 is minimum when

$$Y_{-\frac{L}{2}} = Y_{-\frac{L}{2}+1} = \cdots = Y_{-1} = Y_1 = \cdots = Y_{\frac{L}{2}-1} = Y_{\frac{L}{2}} = \frac{K_2(u,v)}{L_2(u,v)} \qquad \text{(D-15)}$$

Substituting equations (D-15) and (D-14) into equation (D-13) the

minimum ac quantization error becomes

$$\epsilon_2 = \frac{1}{12\, L_2^2(u,v)} \left(\int_{-A_2(u,v)}^{A_2(u,v)} [\, p_2(F)\, dF\,]^{1/3} \right)^3 \qquad \text{(D-16)}$$

A method of obtaining decision levels for dc samples may be obtained by writing

$$Q_K = 2\,\Delta F_1 + 2\,\Delta F_2 + \cdots + 2\Delta F_K \qquad \text{(D-17)}$$

or

$$Q_K = \frac{2K_1}{L_1} \left[p_1^{-1/3}(F_1) + p_1^{-1/3}(F_2) + \cdots + p_1^{-1/3}(F_K) \right] \qquad \text{(D-18)}$$

The series may be approximated by an integral

$$Q_K = C \int_0^{\frac{K A_1}{L_1}} [\, p_1(F)\, (F)]^{-1/3}\, d\, F \qquad \text{(D-19)}$$

where $K = 0, 1, 2, \ldots, L_1$ and C is a constant of proportionality so chosen that when $K = L_1$, $Q_K = A_1$. Hence

$$Q_K = \frac{A_1 \int_0^{\frac{KA_1}{L_1}} [\, p_1(F)\,]^{-1/3}\, dF}{\int_0^{A_1} [\, p_1(F)]^{-1/3}\, dF} \qquad K = 0, 1, \ldots, L_1 \qquad \text{(D-20)}$$

The decision levels for ac sample may then be obtained in a similar way which could be written as follows:

$$Q_{K\,(u,v)\neq(0,0)}(u,v) = \frac{A_2(u,v)\int_0^{\frac{2KA_2(u,v)}{L_2(u,v)}} [p_2(F)]^{-1/3}\,dF}{A_2(u,v)\int_0^{} [p_2(F)]^{-1/3}\,dF} \qquad K = 0,1,\ldots,\frac{L_2(u,v)}{2} \tag{D-21}$$

REFERENCES

1. H. C. Andrews and W. K. Pratt, "Television Bandwidth Reduction by Encoding Spatial Frequencies", Journal of the SMPTE, vol. 77, No. 2, December, 1968, pp. 1279-1281.

2. H. C. Andrews and W. K. Pratt, "Fourier Transform Coding of Images", Hawaii International Conference on System Science, January, 1968, pp. 677-679.

3. H. C. Andrews, J. Kane, and W. K. Pratt, "Hadamard Transform Image Coding", Proceedings of the IEEE, vol. 57, No. 1, January, 1969.

4. A. Habibi and P. A. Wintz, "Image Coding by Linear Transformation and Block Quantization", IEEE Transactions on Communication Technology, vol. Com-19, No. 1, February, 1971, pp. 50-63.

5. H. C. Andrews, "Computer Techniques in Image Processing", Academic Press, New York, 1970.

6. H. Enomoto and K. Shibata, "Orthogonal Transform Coding System for Television Signals", Journal of the Institute of TV Engineers of Japan, vol. 24, No. 2, February 1970, pp. 99-108.

7. H. C. Andrews, "Fourier Coding of Images", University of Southern California, USCEE Report No. 271, June, 1968.

8. W. K. Pratt, "Spatial Transform Coding of Color Images", IEEE Transactions on Communication Technology, vol. COM-19, No. 6, December, 1971.

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

10. J. J. Y. Huang and P. M. Schutheiss, "Block Quantization of Correlated Gaussian Random Variables", IEEE Transactions on Communication System, vol. CS-11, No. 3, September, 1963, pp. 289-296.

11. L. M. Goodman, "A Binary Linear Transformation for Redundancy Reduction", Proceedings IEEE, vol. 55, No. 3, March, 1967, pp. 467-468.

12. C. A. Andrews, J. M. Davies, and G. R. Schwarz, "Adaptive Data Compression", Proceedings IEEE, vol. 55, No. 3, March, 1967, pp. 267-277.

13. L. A. Pipes, "Matrix Methods in Engineering", Prentice-Hall, Englewood Cliffs, New Jersey, 1963.

14. W. K. Pratt and H. C. Andrews, "Transform Image Coding", USCEE Report 387, March 1970.

15. A. Papoulis, "Probability, Random Variables, and Stochastic Processes", McGraw-Hill Inc., 1965.

16. K. L. Chung, "A Course in Probability Theory", Harcourt, Brace & World, Inc., 1968.

17. L. Breiman, "Probability", Addison-Wesley Publishing Co., 1968.

18. W. D. Ray and R. M. Driver, "Further Decomposition of the Karhunen-Loeve Series, Representation of a Stationary Random Process", IEEE Transactions on Information Theory, vol. IT-16, No. 6, November, 1970, pp. 663-668.

19. J. Pearl, H. C. Andrews, and W. K. Pratt, "Performance Measures for Transform Data Coding", IEEE Transactions on Communications, June, 1972, pp. 411-415.

20. P. F. Panter and W. Dite, "Quantization Distortion in Pulse-Count Modulation with Nonlinear Spacing of Levels", IRE Proceedings, January 1951, pp. 44-48.

21. Joel Max, "Quantizing for Minimum Distortion", IRE Transactions on Information Theory, March 1960.

22. W. K. Pratt, L. R. Welch, and W. H. Chen, "Slant Transform for Image Coding", Walsh Domain Conference, Washington D.C., April, 1972.

23. W. K. Pratt, "Walsh Functions in Image Processing and Two-Dimensional Filtering", Walsh Domain Conference, Washington D.C., April 1972.

24. W. K. Pratt, "Spatial Transform Coding of Color Images", IEEE Transactions on Communication Technology, vol. COM-19, No. 6, December 1971, pp. 980-992.

25. W. K. Pratt, "Digital Image Coding and Transmission", USCEE Report 403, June 1971.

26. D. G. Fink, Ed., "Television Engineering Handbook", New York, McGraw-Hill, 1957.

27. R.W.G. Hunt, "The Reproduction of Colour", John Wiley and Sons, New York, 1967.

28. G. Wyszecki and W. S. Stiles, "Color Science", New York, Wiley, 1967.

29. A. Haar, "Zur Theorie des Orthogonalen Funktionen-Systeme", Inaugural dissertation, Math. annals, vol. 69 (1910), pp. 331-371 and vol. 71 (1912), pp. 33-53.

30. C. Wateri, "A Generalization of Haar Functions", Tohoku Mathematical Journal, vol. 8 (1956), pp. 286-290.

31. H. C. Andrews and K. L. Caspari, "A Generalized Technique for Spectral Analysis", IEEE Transactions on Computers, Vol. C-9, No. 1, January 1970, pp. 16-25.

32. W. K. Pratt, "Stop-Scan Edge Detection Systems of Television Bandwidth Reduction", USCEE Report 131T.

33. W. K. Pratt, "A Comparison of Digital Image Transforms", The Mervin J. Kelly Communication Conference, University of Missouri, Rolla, October, 1970.

34. W. H. Chen and W. K. Pratt, "Slant Transform for Color Image Coding", Walsh Domain Conference, Washington D.C., April 1973.

35. W. T. Wintringham, "Color Television and Colorimetry", Proceedings IRE, vol. 39, October, 1951, pp. 1135-1171.

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)	2a. REPORT SECURITY CLASSIFICATION
Image Processing Institute, Electronic Sciences Laboratory, University of Southern California, University Park, Los Angeles, California 90007	UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE

SLANT TRANSFORM IMAGE CODING

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

Technical Report, May 1973

5. AUTHOR(S) (First name, middle initial, last name)

Wen-Hsiung Chen

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
May 1973	167	35

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
F08606-72-C-0008	USCEE Report 441
b. PROJECT NO. ARPA Order No. 1706	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	None

10. DISTRIBUTION STATEMENT

Approved for release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

A slant transform matrix consisting of basis vectors which resemble typical lines of an image has been developed. A fast transform algorithm based on the matrix decomposition has also been presented. The transform has been proven to be superior, from the standpoint of image quality, to other transforms possessing fast computational algorithms.

The statistical properties of the slant transform have been analyzed by introducing probability density and covariance models for the transform samples. The bandwidth reduction capability of the slant transform has been investigated by several test images. Two methods of achieving bandwidth reduction have been presented, namely, threshold and zonal coding. Studies have indicated that the average coding of a monochrome image can be reduced from 8 bits/pixel to 1 bit/pixel or 1.5 bits/pixel for the threshold and zonal coding, respectively, without seriously degrading the image quality. Studies have also indicated that zonal coding has an extremely high noise immunity, and can be practically implemented. The average coding of a color image can be reduced from 24 bits/pixel to 2 bits/pixel by zonal coding while preserving good quality reconstruction.

****

14. Key Words: Image coding, Slant transform

DD FORM 1 NOV 65 1473

UNCLASSIFIED

Security Classification

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