USC–EE REPORT #387

Transform Image Coding

By

William K. Pratt
Harry C. Andrews

March 1970

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
USC Viterbi School of Engineering
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Suite 400
Los Angeles, CA 90089-2564 U.S.A.

March 1970

USCEE Report 387

TRANSFORM IMAGE CODING

Final Report

by

William K. Pratt
Harry C. Andrews

for

California Institute of Technology
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91103

Contract No. 952312

Electronic Sciences Laboratory
University of Southern California
University Park
Los Angeles, California 90007

Preface

This report comprises the final report for a study entitled, "Transform Processing and Coding of Images," performed by the Electronic Sciences Laboratory of the University of Southern California for the Jet Propulsion Laboratory under JPL Contract 952312. Mr. Thomas Rindfleisch of JPL served as project director for the study. This report supplants the interim report USCEE No. 341 entitled "Transform Processing and Coding of Images," published in March, 1969. Pertinent introductory material from that report is included in the present report for completeness.

Transform Image Coding

1. Introduction

The basic goal of digital image coding is the development of a coding technique that permits the representation, and subsequent recovery, of an image by a minimal number of code bits [1-3]. In some applications virtually no image distortion is permitted in the coding process, while in other applications a controlled amount of distortion is allowable in the achievement of a substantial bit reduction. In general, when redundancy is removed from a data source, the compressed data is more sensitive to the effect of channel errors. One of the restrictions in selecting a coding method, therefore, is that the compressed data must not be overly sensitive to channel errors.

In 1967 a new technique of image coding, called Fourier transform coding, was developed at the University of Southern California [4-6]. Another related method, called Hadamard transform coding,was discovered at USC in 1968 [7-8]. Since then investigations have been made into the applications of other mathematical transforms for image coding. Out of these studies has emerged the generalized technique of transform image coding [9-11].

1.1 Image Transform Coding

Figure 1-1 contains a block diagram of the image transform coding system. In operation a two-dimensional transform is taken of the brightness samples of an image, or subsection of an image, on a line by line basis. The resultant transform samples are then operated upon by a sample selector that selects which samples are to be transmitted on the basis of magnitude or position in the plane. Those samples that are to be transmitted are quantized and coded. At the receiver, the data is decoded to reconstruct the transform domain, and an inverse transform is taken to reconstruct the original image.

A bandwidth reduction is achieved simply by not transmitting all of the transform domain samples. Those samples that are not transmitted are generally of such low magnitude that they contribute little in the image reconstruction.

There are two basic forms of sample selection--zonal sampling and threshold sampling--that can be employed. In zonal sampling, only those transform samples that lie within a certain geometric region in the transform domain are selected for transmission. The basic problem with zonal sampling is that in certain pictures many large magnitude samples may lie without the zonal region and will, therefore, not be transmitted. In order

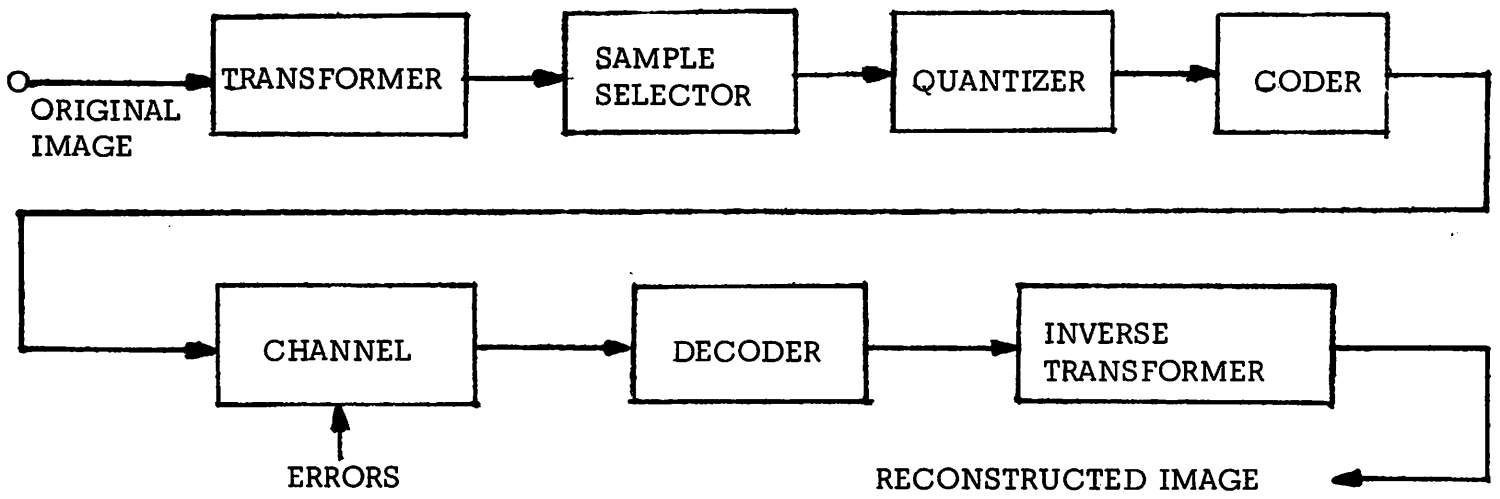


Figure 1-1. Generalized Transform Coding of Images

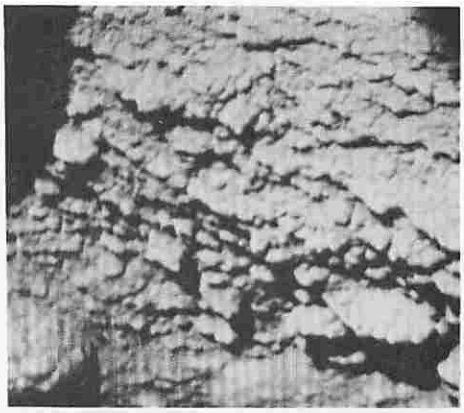
to avoid such errors it is possible to establish a threshold level on the magnitude of transform domain samples such that if the transform sample magnitude is greater than the threshold it will be selected, and the sample will be deleted if it falls below the threshold. With threshold coding it is necessary to code the location in the transform domain of a selected sample as well as its value.

The major advantage of image transform coding other than its potential for bandwidth compression is the tolerance to channel errors that transform coding affords. An intuitive justification for transmitting the transform of an image rather than the spatial representation of the image is that for many transforms the channel noise introduced in the image transform tends to be distributed evenly over the entire reconstructed image. Consequently, the channel noise is manifested as a low spatial frequency error in reconstruction. Experimental evidence indicates that the eye is more sensitive to the high frequency discrete errors caused by channel errors in the spatial domain than it is to the same number of errors in the transform domain.

1.2 Original Images

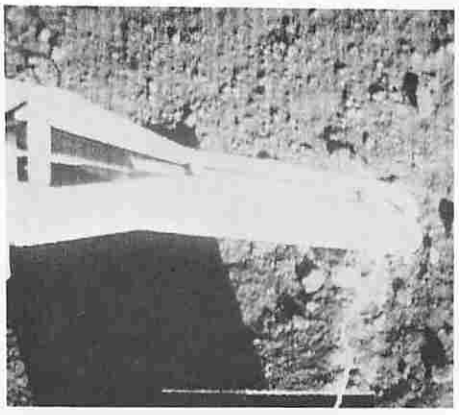
Figure 1-2 contains photographs of the five original images that have been used as test images for the evaluation of image transform coding. These images contain 256 by 256 elements quantized to 64 grey levels. The images were read from magnetic tape, displayed on a Hewlett-Packard Model 1300 cathode ray tube display, and photographed with Polaroid Type 47 film.

a. Surveyor footpad



b. Moonscape

c. Surveyor experimental box



d. Surveyor boom

e. Girl

Figure 1-2 Original Test Images

2. Image Transformation

In this section consideration is given to the mathematical formulation of image transforms. The characteristics and properties of the Fourier, Hadamard, Haar, Karhunen-Loeve, and a class of transitional transforms are briefly developed. Experimental results are presented.

2.1 Formulation

An image may be represented by an array of intensity components or samples over the image surface by two dimensional sampling. For the present discussion an image array will be considered to be a square array of N^2 intensity samples described by the function $f(x, y)$ over the image coordinates (x, y).

Conceptually, there are two major types of image transforms which shall be called transforms of the first and second kind. A transform of the first kind maps a two dimensional image array of dimension $N \times N$ into a one dimensional vector of dimension $1 \times N^2$ according to the relation

$$F(w) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)\, a(x, y, w) \qquad (2\text{-}1)$$

for $w = 0, 1, 2, \cdots, N^2-1$

where a(x, y, w) is the forward transform kernel of the first kind. A reverse transform of the first kind is defined as

$$\hat{f}(x,y) = \sum_{w=0}^{N^2-1} F(w)\, b(x,y,w) \qquad (2\text{-}2)$$

for $x, y = 0, 1, 2, \cdots, N-1$

where b(x, y, w) is the reverse transform kernel of the first kind. A transform of the second kind maps an image array of dimension $N \times N$ into a two dimensional array of the same dimension as defined by

$$F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)\, a(x,y,u,v) \qquad (2\text{-}3)$$

for $u, v = 0, 1, 2, \cdots, N-1$

where a(x, y, u, v) is the forward transform of the second kind. The corresponding reverse transformation is given by

$$\hat{f}(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v)\, b(x,y,u,v) \qquad (2\text{-}4)$$

for $x, y = 0, 1, 2, \cdots, N-1$

where b(x, y, u, v) is the reverse transform kernel of the second kind. For transforms of the first and second kind, when the function $\hat{f}(x,y)$ resulting from the reverse transform operation is equivalent to the original image, f(x, y), the reverse transform is

called an inverse transform. Transforms of the first and second kind are said to be orthogonal if the following conditions are met*

Transforms of the first kind:

$$\sum_{w} a(x,y,w)\, a^{*}(\alpha, B, w) = \delta(x-\alpha, y-B) \tag{2-5a}$$

$$\sum_{w} b(x,y,w)\, b^{*}(\alpha, B, w) = \delta(x-\alpha, y-B) \tag{2-5b}$$

$$\sum_{x}\sum_{y} a(x,y,w)\, a^{*}(x,y,\theta) = \delta(w-\theta) \tag{2-5c}$$

$$\sum_{x}\sum_{y} b(x,y,w)\, b^{*}(x,y,\theta) = \delta(w-\theta) \tag{2-5d}$$

Transforms of the second kind:

$$\sum_{u}\sum_{v} a(x,y,u,v)\, a^{*}(\alpha, B, u, v) = \delta(x-\alpha, y-B) \tag{2-6a}$$

$$\sum_{u}\sum_{v} b(x,y,u,v)\, b^{*}(\alpha, B, u, v) = \delta(x-\alpha, y-B) \tag{2-6b}$$

$$\sum_{x}\sum_{y} a(x,y,u,v)\, a^{*}(x,y,\theta,\varphi) = \delta(u-\theta, v-\varphi) \tag{2-6c}$$

$$\sum_{x}\sum_{y} b(x,y,u,v)\, b^{*}(x,y,\theta,\varphi) = \delta(u-\theta, v-\varphi) \tag{2-6d}$$

---

*The limits of summation are eliminated in subsequent equations unless required for clarity.

A forward transform kernel of the second kind is said to be separable if it can be written as

$$a(x, y, u, v) = a_1(x, u)\, a_2(y, v) \tag{2-7}$$

A separable two dimensional transform can be computed in two steps. First, a one dimensional transform is taken along each row of the image, $f(x, y)$, yielding

$$F(u, y) = \sum_{x=0}^{N-1} f(x, y)\, a_1(x, u) \tag{2-8}$$

Next, a second one dimensional transform is taken along each column of $F(u, y)$ giving

$$F(u, v) = \sum_{y=0}^{N-1} F(u, y)\, a_2(y, v) \tag{2-9}$$

The transformation kernel is called separable symmetric if

$$a(x, y, u, v) = a_1(x, u)\, a_1(y, v) \tag{2-10}$$

For ease of implementation, the separable symmetric property is desirable.

It is often useful to express two dimensional transforms in matrix notation. For example, with a forward transform kernel of the second kind that is separable symmetric let:

$[f]$ = image matrix, $f(x, y)$

$[F]$ = transformed image matrix, $F(u, v)$

$[A]$ = transform matrix, $A(\alpha, \beta)$

Then by matrix multiplication

$$[F] = [A][f][A] \tag{2-11}$$

Now pre- and post-multiplication of each side of $[F]$ by a reverse transform matrix, $[B]$, gives

$$[\hat{f}] \equiv [B][F][B] = [B][A][f][A][B] \tag{2-12}$$

where $[\hat{f}]$ is, in general, an approximation of $[f]$. If the reverse transform matrix is the inverse matrix $[A]^{-1}$ of $[A]$, then

$$[\hat{f}] = [A]^{-1}[A][f][A][A]^{-1} \tag{2-13}$$

But

$$[A]^{-1}[A] = [A][A]^{-1} = [I] \tag{2-14}$$

where $[I]$ is the identify matrix. Hence

$$[\hat{f}] = [f] = [A]^{-1}[F][A]^{-1} \tag{2-15}$$

Thus, $f(x, y)$ and $F(u, v)$ can be expressed as two dimensional transform pairs if $[A]$ has an inverse. If $[A]$ is a unitary matrix, then by definition

$$[B] \equiv [A]^{-1} = [A]^{*T} \qquad \underline{\text{unitary matrix}} \qquad (2\text{-}16)$$

where $[A]^*$ is the complex conjugate matrix of $[A]$ and $[A]^T$ is the matrix transpose of $[A]$. If in addition $[A]$ is symmetric

$$[B] \equiv [A]^{-1} = [A]^* \qquad \underline{\text{symmetric unitary matrix}} \qquad (2\text{-}17)$$

A real, unitary matrix is called an orthogonal matrix. For such a matrix

$$[B] \equiv [A]^{-1} = [A]^T \qquad \underline{\text{orthogonal matrix}} \qquad (2\text{-}18)$$

Finally, if $[A]$ is a symmetric orthogonal matrix, then

$$[B] \equiv [A]^{-1} = [A] \qquad \underline{\text{symmetric orthogonal matrix}} \qquad (2\text{-}19)$$

If the forward transformation matrix is constrained to be unitary, then the transformation can be interpreted as a decomposition of the image data into a generalized two dimensional spectrum. Each spectral component in the transform domain corresponds to the amount of energy of the spectral function within the original image. In this context the concept of frequency may now be generalized to include transformations of functions other than sine and cosine waveforms. This type of generalized spectral analysis is useful in the investigation of specific decompositions which are best suited for particular classes of images.

The following paragraphs contain an analysis of the Fourier, Hadamard, Haar, transitional, and Karhunen-Loeve transformations with particular emphasis on their applicability to image processing.

2.2 Fourier Transform

The discrete Fourier transform with and without efficient computational algorithms, has long been used for signal analysis [12]. Only recently have Fourier transform methods been utilized for image coding [4-6].

The two dimensional Fourier transform of an image field, f(x, y), may be expressed as

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \exp\left\{ - \frac{2\pi i}{N} (ux + vy) \right\} \qquad (2\text{-}20)$$

The inverse Fourier transform which reconstructs the original image is given by

$$f(x,y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{y=0}^{N-1} F(u,v) \exp\left\{ \frac{2\pi i}{N} (ux + vy) \right\} \qquad (2\text{-}21)$$

Since the transform kernels are separable and symmetric the two dimensional transform can be computed as two sequential one dimensional transforms.

The terms u and v are called the spatial frequencies of the image in analogy with time series analysis. When the Fourier transform relationship is expressed in the form given by Equation (2-20) the origin, or zero spatial frequency term appears in the corner of the transform plane. For display purposes it is convenient to shift the origin to the center of the transform domain. This is easily accomplished by multiplying the image by the function $(-1)^{x+y}$ before the transformation [13].

Even though f(x, y) is a real positive function, its transform, F(u, v), is in general complex. Thus, while the image contains N^2 components, the transform contains $2N^2$ components, the real and imaginary, or magnitude and phase components of each spatial frequency. However, since f(x, y) is a real positive function, F(u, v) exhibits a property of conjugate symmetry [13]. Specifically,

$$F(u, v) = F^*(-u, -v) \tag{2-22}$$

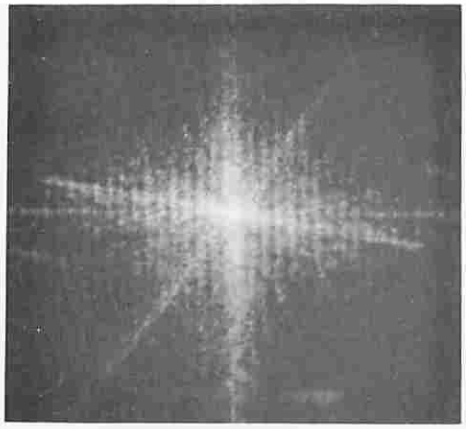
As a result of the conjugate symmetry property of the Fourier transform it is only necessary to transmit the samples of one half of the transform plane; the other half can be reconstructed from the half plane samples transmitted*. Hence, the Fourier transform of an image

*A reconstruction of the original can be obtained from the half plane transform samples directly by a Hilbert filtering technique [13].

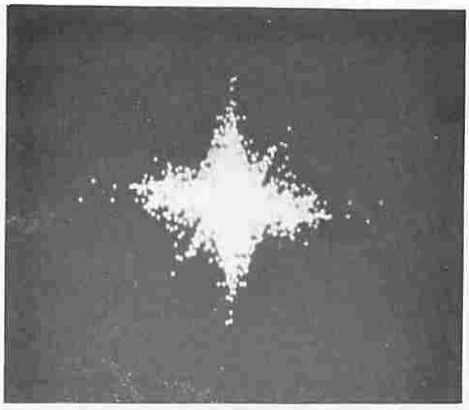
can be described by N^2 data components.

The two dimensional Fourier transform of an image is essentially a Fourier series representation of a two dimensional field. For the Fourier series representation to be valid the field must be periodic. Thus, the original image must be considered to be periodic horizontally and vertically. The right side of the image therefore abuts the left side and the top and bottom of the image are adjacent. Spatial frequencies along the coordinate axes of the transform plane arise from these transitions. Although these are false spatial frequencies from the standpoint of being necessary for representing the image within the image boundary, they do not impair reconstruction. On the contrary, these spatial frequencies are required to reconstruct the sharp boundaries of the image.

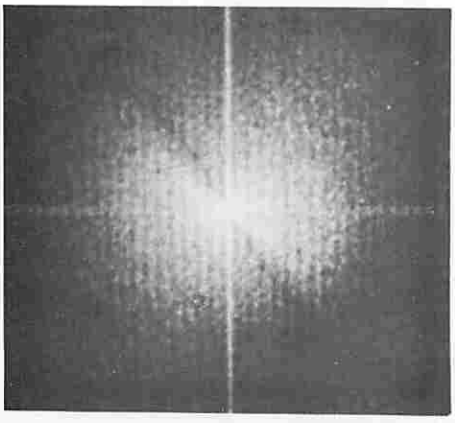
Figure 2-1 presents displays of the Fourier transforms in shifted form of two of the original test scenes. The logarithm of the magnitude of each transform is displayed rather than the magnitude itself in order to reduce the dynamic range of the display. In addition, a threshold display is presented in which all the absolute values above the threshold are set to white and all others are made black. Such a display gives a graphic illustration of the heavy concentration of energy around the origin (center of photograph) of the Fourier transform.



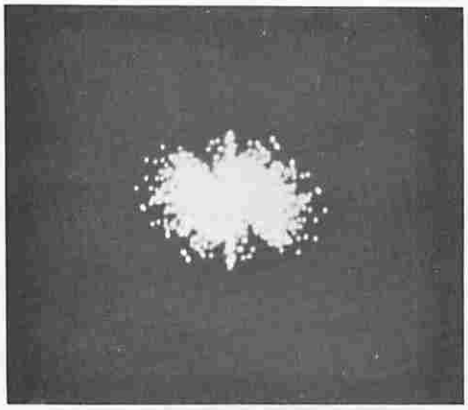
a. Logarithm of the magnitude of the Surveyor box transform



b. Threshold display of the Surveyor box transform



c. Logarithm of the magnitude of the moonscape transform



d. Threshold display of the moonscape transform

Figure 2-1 Fourier Transforms

2.3 Hadamard Transform

The Hadamard transform, also known as the Walsh transform, is based upon the Hadamard matrix which is a square array of plus and minus ones whose rows and columns are orthogonal to one another [14-16]. If [H] is an N by N Hadamard matrix, then the product of N and its transpose is

$$[H][H]^T = N[I] \tag{2-23}$$

If [H] is a symmetric Hadamard matrix, then Equation (2-23) reduces to

$$[H]\,[H] = N[I] \tag{2-24}$$

A Hadamard matrix multiplied by the normalization factor $\frac{1}{\sqrt{N}}$ is an orthonormal matrix.

The lowest order Hadamard matrix is the Hadamard matrix

$$[H_2] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2-25}$$

It is known that if a Hadamard matrix of order N exists ($N > 2$), then $N \equiv 0 \pmod 4$. The existence of a Hadamard matrix for every value of N satisfying this requirement has not been shown, but constructions are available for nearly all permissible values of N up to 200. The simplest construction is for a Hadamard matrix

of order $N = 2^n$ where n is an integer. In this case if $[H_N]$ is a Hadamard matrix of order N, the matrix

$$[H_{2N}] = \begin{bmatrix} H_N & H_N \\ H_N & -H_N \end{bmatrix} \tag{2-26}$$

is a Hadamard matrix of order 2N.

A frequency interpretation can be given to the Hadamard matrix generated from the core matrix of Equation (2-25). Along each row of the Hadamard matrix the frequency is called the number of changes in sign. Harmuth has coined the word "sequency" to designate the number of sign changes [17]. It is possible to construct a Hadamard matrix of order $N = 2^n$ that has frequency components at every integer from 0 to N-1.

This frequency interpretation of the rows of a Hadamard matrix leads one to consider the rows to be equivalent to rectangular waves ranging between ± 1 with a sub-period of $\frac{1}{N}$ units. Such functions are called Walsh functions [18-22] and are further related to the Rademacher functions [23]. Thus, in this context the Hadamard matrix merely performs the decomposition of a function by a set of rectangular waveforms rather than the sine-cosine waveforms associated with the Fourier transform.

For symmetric Hadamard matrices of order $N = 2^n$, the two dimensional Hadamard transform may be written in series

form as

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)(-1)^{p(x,y,u,v)} \tag{2-27}$$

where $p(x,y,u,v) \equiv \sum_{i=0}^{n-1} (u_i x_i + v_i y_i)$. The terms u_i, v_i, x_i, and y_i are the binary representations of u, v, , x, and y respectively. For example,

$$(u)_{\text{DECIMAL}} = (u_{n-1}\ u_{n-2} \cdots u_1\ u_0)_{\text{BINARY}}$$

where $u_i \in \{0,1\}$.

Another series representation exists for a Hadamard matrix in "ordered" form in which the sequency of each row is larger than the preceding row. By this representation

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)\ (-1)^{q(x,y,u,v)} \tag{2-28}$$

where

$$q(x,y,u,v) = \sum_{i=0}^{n-1} [g_i(u)\ x_i + g_i(v)\ y_i]$$

and

$$\begin{aligned} g_0(u) &\equiv u_{n-1} \\ g_1(u) &\equiv u_{n-1} + u_{n-2} \\ g_2(u) &\equiv u_{n-2} + u_{n-3} \\ &\vdots \\ g_{n-1}(u) &\equiv u_1 + u_0 \end{aligned}$$

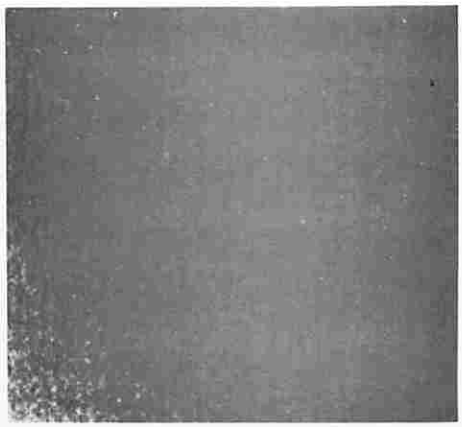
The two dimensional Hadamard transform may be computed in either natural or ordered form with an algorithm analogous to the fast Fourier transform computer algorithm.

Figure 2-2 presents the ordered Hadamard transforms of two test scenes. The origin of the transform domain is now in the lower left corner and the axes are now spatial sequencies as opposed to spatial frequencies. Notice that as in the Fourier case, the image energy tends to concentrate itself heavily in the lowest spatial sequency areas providing the potential for large bandwidth reductions. Again both a logarithmic and threshold display are provided for dynamic range purposes.

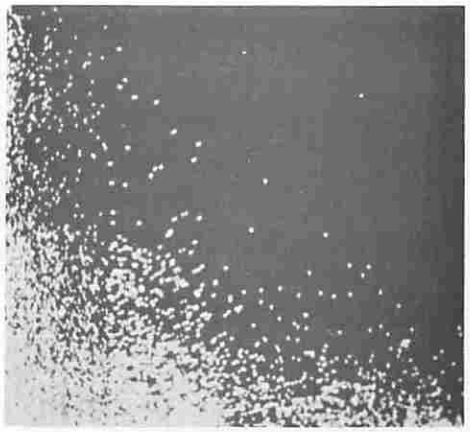
2.4 Transitional Transforms

In related work [24-27] it has been shown that a class of rapidly implementable orthogonal transformations exists for matrices composed of Kronecker products of smaller core matrices. In fact, both the Hadamard and Fourier transforms have been shown to be subsets of this much larger class of Kronecker transform matrices. A class of transformations exists for which the Hadamard transform is a limiting case, and these transforms will now be investigated as to their image processing potential.

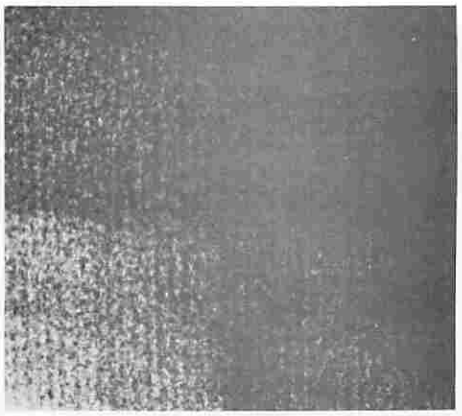
The transformation resulting from performing the Kronecker operation of the core matrix



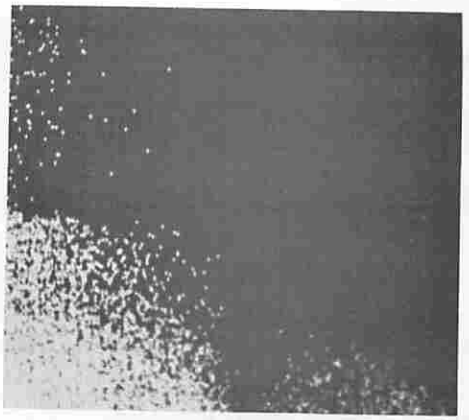
a. Logarithm of the magnitude of the Surveyor box transform



b. Threshold display of the Surveyor box transform



c. Logarithm of the magnitude of the moonscape transform



d. Threshold display of the moonscape transform

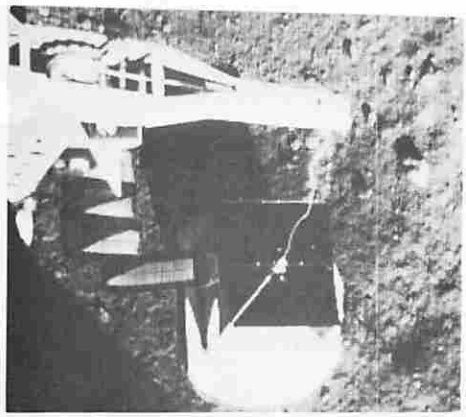
Figure 2-2 Hadamard Transforms

$$[H] = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \tag{2-29}$$

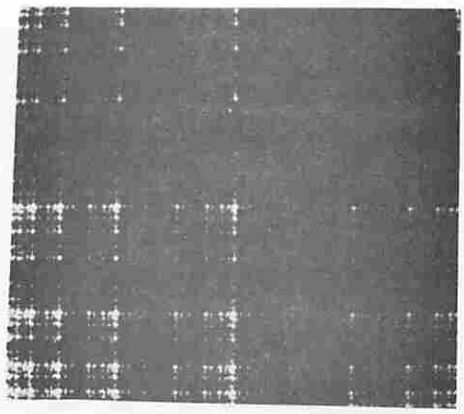
with itself n times results in a matrix whose row and column entries, indexed by x and u respectively, can be described by the following equation

$$H_n(x,u) = \left(\frac{\sin\theta}{\cos\theta} \right)^{\sum_{r=0}^{n-1} u_r \oplus x_r} (-1)^{\sum_{r=0}^{n-1} u_r x_r} \tag{2-30}$$

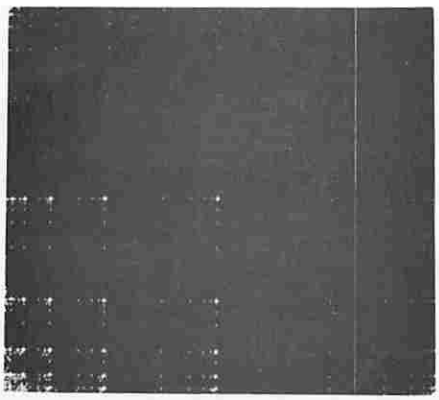
The u_r and x_r variables are the bits in the binary representation of the column and row indexes respectively. It is evident that while the Hadamard transform has received considerable attention (often under the name of the discrete Walsh transform) it is important to note that this transform is the limiting case of the powers of two Kronecker transforms presented above. As θ varies between 0° and 45° the transforms vary from a diagonal matrix to the Hadamard matrix at 45°. In the process of varying θ over this interval, the transformations have ranged from having all of their energy on the diagonal at 0° to uniform energy spread at 45° (Hadamard case). Figure 2-3 presents examples of the transitional transforms of the Surveyor box test scene for four different values of θ. Notice that functions of the magnitude are displayed in all cases because even for the $\theta = 0^\circ$ diagonal case, negative signs on the diagonal are possible. It is evident from



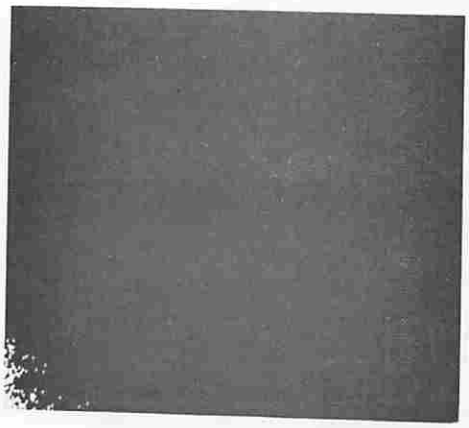
a. $\Theta = 0^{\circ}$, diagonal transform, magnitude display, max. value = 63



b. $\Theta = 15^{\circ}$, threshold display max. value = 1,200



c. $\Theta = 30^{\circ}$, threshold display max. value = 6,611



d. $\Theta = 45^{\circ}$, Hadamard transform, threshold display, max. value = 11,486

$$\left[H\right] = \begin{bmatrix} \cos\Theta & \sin\Theta \\ \sin\Theta & -\cos\Theta \end{bmatrix}$$

Figure 2-3 Transitional Transforms

this figure that the transform which computes the image into the fewest significant coefficients is the Hadamard transform.

2.5 Haar Transform

The Haar transform [28] is another transformation that, like the Hadamard or Walsh transform, requires no multiplications. The Haar matrix consists of plus and minus ones as well as zeros and is non-symmetric, orthogonal but not orthonormal (unless multiplied by the proper diagonal matrix). The Haar matrix can be likened to a sampling system in which various rows sample the input with finer and finer resolution increasing in powers of two. An 8×8 orthonormal Haar matrix is shown below:

$$[H] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix} \qquad (2\text{-}31)$$

The Haar transform is defined for data of resolution equal to a power of two, and the matrix is factorable into a product of matrices with a large number of zero entries [25]. Consequently,

a fast algorithm also exists for this transform. The number of computer operations required for a vector matrix multiplication is given by 2(N-1) as compared to the $2N \log_2 N$ requirement of Fourier or Hadamard, which itself is a considerable savings over the normal vector-matrix multiplication requirement of N^2 operations. As with the Walsh functions, the Haar functions can be generalized to contain entries of roots of unity other than ± 1. Watari [29] has described the generalized Haar system and has shown that it is possible to preserve some of the original Haar convergence properties. The extension to matrix factorization is straightforward and will not be pursued further. However, the number of operations necessary to implement a p^{th} order generalized Haar transform is given by a geometric progression resulting in $p(N-1)/(p-1)$. In image processing applications, the Haar transform provides a transform domain in which a type of differential energy is concentrated in localized regions. Thus there is an area in which adjacent picture element differential energy is concentrated, (the upper right quarter of the transform plane), an area in which differential energy of adjacent picture elements taken two at a time is concentrated, and in general an area in which difference energy of adjacent picture elements taken a power of two at a time is concentrated.

Figure 2-4 presents the Haar transform of the test scenes. The logarithmic results vividly display the derivative energy effect especially in the upper right quarter of the plane. Note that in the Haar transform there is also a concentration of image energy in the lower left corner or origin of the transform plane. The data point at the origin in the Fourier, Hadamard, and Haar transforms all are equal to the average energy in the original image and correspond to the row of all "ones" in the transform matrices.

2.6 Karhunen-Loeve Transform

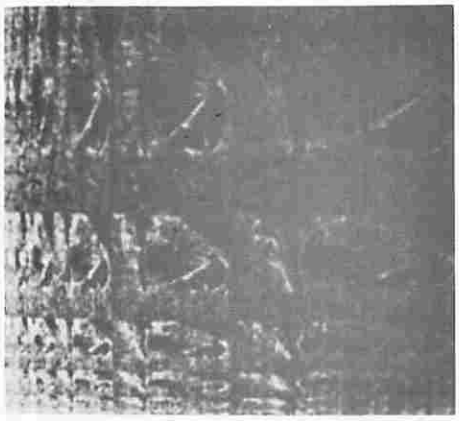
The Karhunen-Loeve transform is a special case of an eigenvector matrix transformation [30-37]. Consider a real symmetric matrix [C] of order n. The eigenvectors of [C] are column vectors $[K_i]$, $i = 1, 2, \cdots, n$ satisfying the relationship

$$[C][K_i] = \lambda_i [K_i] \tag{2-32}$$

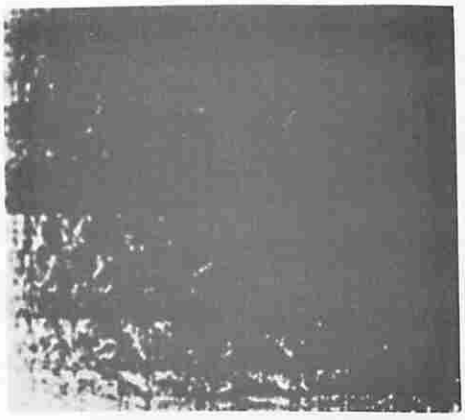
where the scalars λ_i are the eigenvalues of [C]. Let a square matrix [K], called the modal matrix of [C], be constructed from the eigenvector columns in the following manner:

$$[K] = \left[[K_i][K_2] \cdots [K_n]\right] \tag{2-33}$$

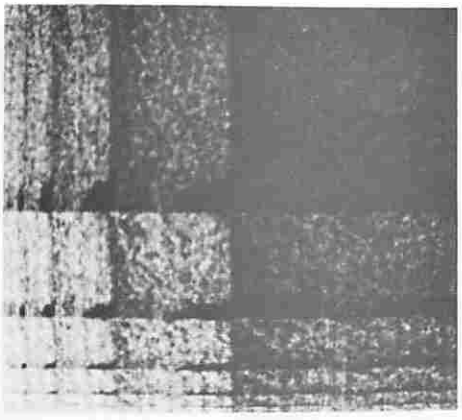
Also let the eigenvalues be located along the diagonal of a matrix



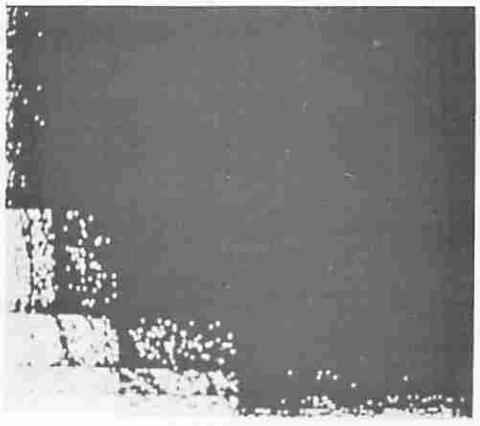
a. Logarithm of the magnitude of the Surveyor box transform



b. Threshold display of the Surveyor box transform



c. Logarithm of the magnitude of the moonscape transform



d. Threshold display of the moonscape transform

Figure 2-4 Haar Transforms

$$[E] = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Then by equation (2-32)

$$[C][K] = [K][E] \tag{2-34}$$

Now, premultiplication of equation (2-34) by $[K]^{-1}$ gives

$$[K]^{-1}[C][K] = [K]^{-1}[K][E] = [E] \tag{2-35}$$

Taking the transpose of both sides of equation (2-35) yields

$$[K]^T[C]^T\left[[K]^{-1}\right]^T = [E]^T \tag{2-36}$$

But, since [C] is a symmetric matrix and [E] is diagonal, by correspondence between equation (2-35) and (2-36)

$$[K]^{-1} = [K]^T \tag{2-37}$$

Thus, if they exist, the eigenvectors of a matrix are orthogonal. It can be easily shown [38] that when [C] is symmetric, its eigenvalues are all real quantities.

Consider now a data column vector [f] of length m. The eigenvector transform [F] of [f] is then

$$[F] = [K][f] \tag{2-38}$$

and the inverse eigenvector transform $[\hat{f}]$ of $[F]$ is

$$[\hat{f}] = [K]^T [F] = [K]^T [K][f] = [f] \tag{2-39}$$

Thus $[f]$ and $[F]$ are transform pairs of an orthogonal matrix transformation. The vector $[F]$ represents a matrix decomposition of $[f]$ into a set of orthogonal waveforms defined by $[K]$. Generally, the exact form of the orthogonal functions cannot be easily described.

If only the first q of the m columns of $[K]$ are employed in the forward and reverse transform, then the mean square error between the original and the reconstructed data vector is [30, 31]

$$\delta = \sum_{k=q+1}^{m} \lambda_k \tag{2-40}$$

Since the λ_k are monotonically decreasing in value, the error will be minimum for any q.

When the eigenvector matrix $[K]$ is composed of eigenvectors of the covariance matrix

$$[C] \equiv E\left\{ [f(i)-\overline{f(i)}] [f(j)-\overline{f(j)}] \right\} \tag{2-41}$$

for i, j = 1, 2, ..., n, of the data vectors, the resulting eigenvector matrix $[K]$ is called the Karhunen-Loeve (K-L) transform.

For a two dimensional image transform of the first kind, the forward transform kernel $a(x, y, w)$ of Equation (2-1) satisfies the equation

$$\lambda(w)\, a(x, y, w) = \sum_{x'=0}^{N-1} \sum_{y'=0}^{N-1} C\{x, x', y, y'\}\, a(x', y', w) \qquad (2\text{-}42)$$

for $w = 0, 1, 2, \cdots, N^2-1$, where $\lambda(w)$ are the eigenvalues of the covariance function $C\{x, x', y, y'\}$ of the image. The corresponding K-L transform of the second kind described by the kernel $a(x, y, u, v)$ of Equation (2-3) is found from

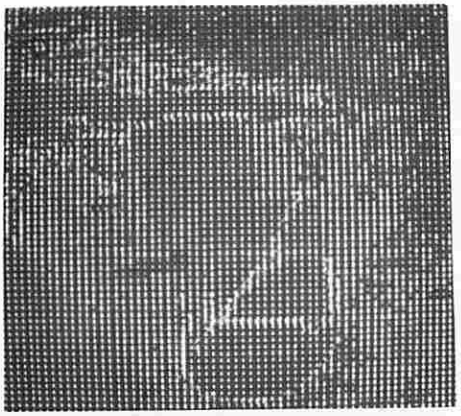
$$\lambda(u, v)\, a(x, y, u, v) = \sum_{x'=0}^{N-1} \sum_{y'=0}^{N-1} C\{x, x', y, y'\}\, a(x', y', u, v) \qquad (2\text{-}43)$$

for $u, v = 0, 1, 2, \cdots, N-1$, where $\lambda(u, v)$ are a two dimensional ordering of the eigenvalues $\lambda(w)$. If the covariance function in Equation (2-43) can be written as

$$C\{x, x', y, y'\} = C_1\{x, x'\}\, C_2\{y, y'\} \qquad (2\text{-}44)$$

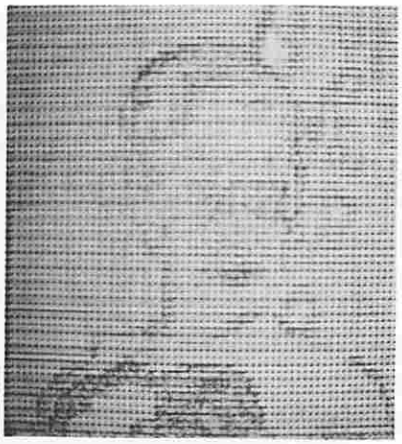
then the transform kernel $a(x, y, u, v)$ can be separated. The resulting two dimensional transform can then be computed sequentially along each row and column of the image.

Figure 2-5 contains photographs of an image that has been Karhunen-Loeve transformed in 4×4 element blocks. The 16×16



a. Logarithm of the magnitude of the Surveyor box transform

b. Inverse transform of transform of Surveyor box



c. Logarithm of the magnitude of the girl transform

d. Inverse transform of transform of girl

$$A = 1.$$
$$B = .8$$
$$C = .6$$
$$D = E = F = 0$$

Figure 2-5 Karhunen-Loeve Transforms in 4 x 4 Element Blocks

component correlation matrix from which the K-L transform matrix was derived is shown in Figure 2-6.

2.7 Summary

For image coding the desirable properties of a mathematical transform are that the transform redistribute the image energy to as few transform domain samples as possible, and furthermore that the transform be easily computable. The Fourier and Hadamard transforms fulfill both requirements, and will be analyzed in greater detail in subsequence sections.

None of the transitional transforms, other than the Hadamard transform, provide a compact distribution of energy. These transforms will not be considered further for image coding, but it should be noted that the transitional transforms may be useful for dimensionality reduction for pattern recognition applications.

The Haar transform possesses an extremely fast computational algorithm. However, the peculiar spatial sampling procedure--sampling in pairs of elements--does not appear to be particularly useful for image coding, and therefore the Haar transform will not be considered further. The Haar transform may find some usefulness, however, for digital edge enhancement since the transform domain is a mapping of the spatial

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

a. Element Array

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	A	B	C	D	B	C	D	E	C	D	E	F	D	E	F	G
2	B	A	B	C	C	B	C	D	D	C	D	E	E	D	E	F
3	C	B	A	B	D	C	B	C	E	D	C	D	F	E	D	E
4	D	C	B	A	E	D	C	B	F	E	D	C	G	F	E	D
5	B	C	D	E	A	B	C	D	B	C	D	E	C	D	E	F
6	C	B	C	D	B	A	B	C	C	B	C	D	D	C	D	E
7	D	C	B	C	C	B	A	B	D	C	B	C	E	D	C	D
8	E	D	C	B	D	C	B	A	E	D	C	B	F	E	D	C
9	C	D	E	F	B	C	D	E	A	B	C	D	B	C	D	E
10	D	C	D	E	C	B	C	D	B	A	B	C	C	B	C	D
11	E	D	C	D	D	C	B	C	C	B	A	B	D	C	B	C
12	F	E	D	C	E	D	C	B	D	C	B	A	E	D	C	B
13	D	E	F	G	C	D	E	F	B	C	D	E	A	B	C	D
14	E	D	E	F	D	C	D	E	C	B	C	D	B	A	B	C
15	F	E	D	F	E	D	C	D	D	C	B	C	C	B	A	B
16	G	F	F	D	F	F	D	C	E	D	C	B	D	C	B	A

b. Correlation Matrix

Figure 2-6 Correlation Matrix Model for 4 x 4 Element Karhunen-Loeve Transform

differential energy of the original image.

The Karhunen-Loeve transform provides the best compaction of image energy for natural images. The major difficulty associated with the use of the Karhunen-Loeve transform for image coding is the great amount of computation involved. First, the image correlation matrix must be estimated or modeled. Next, the correlation matrix must be diagonalized to determine its eigenvalues and eigenvectors. Finally, the transform itself must be taken. In general , there is no fast computational algorithm for the transform. In those applications in which the amount of computation is not of principal concern, the Karhunen-Loeve transform may find practical application. Furthermore, since the K-L transform is the optimum image transform in a mean square error sense, when sample deletion is employed, it is worthwhile to consider its performance as a standard for other image transforms.

The next two sections contain a general analysis of the Fourier, Hadamard, and Karhunen-Loeve image transforms.

3. Statistical Analysis of Image Transforms

The development of efficient quantization and coding methods for image transform samples requires an understanding of the statistical properties of the transform domain. This section presents a derivation of the first and second moments of transform samples, and also contains the development of a stochastic model for the probability density of transform samples.

The statistical analysis of image transforms is predicated on the representation of an original image as a two dimensional stochastic process, $f(x, y)$. The spatial mean

$$E\{f(x,y)\} \equiv \overline{f(x,y)} \tag{3-1}$$

and the covariance

$$E\left\{[f(x_1,y_1)-\overline{f(x_1,y_1)}]\,[f(x_2,y_2)-\overline{f(x_2,y_2)}]\right\} \equiv C\{x_1,x_2,y_1,y_2\} \tag{3-2}$$

are assumed known or at least estimateable. Appendix A describes measurements of the covariance function of an image.

3.1 Moments

For a generalized forward transform given by

$$F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)\, a(x,y,u,v) \tag{3-3}$$

the mean of the transform samples is simply the forward transform of the mean of the image samples. Thus,

$$E\{F(u,v)\} \equiv \overline{F(u,v)} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \overline{f(x,y)}\, a(x,y,u,v) \tag{3-4a}$$

For an ordered, orthonormal transform with an average value term

$$\overline{F(u,v)} = N\, \overline{f(x,y)}\, \delta(u,v) \tag{3-4b}$$

The covariance function of the transform domain samples is by definition

$$C\{u_1,u_2,v_1,v_2\} \equiv E\left\{[F(u_1,u_1)-\overline{F(u_1,v_1)}][F^*(u_2,v_2)-\overline{F(u_2,v_2)}^*]\right\} \tag{3-5}$$

Substitution of equations (3-3) and (3-4) yields

$$C\{u_1,u_2,v_1,v_2\} = E\left\{\left[\sum_{x_1}\sum_{y_1} [f(x_1,y_1)-\overline{f(x_1,y_1)}]\, a(x_1,y_1,u_1,v_1)\right]\right.$$
$$\left.\cdot\left[\sum_{x_2}\sum_{y_2} [f^*(x_2,y_2)-\overline{f(x_2,y_2)}^*]\, a^*(x_2,y_2,u_2,v_2)\right]\right\} \tag{3-6}$$

or

$$C\{u_1, u_2, v_1, v_2\} = \sum_{x_1} \sum_{y_1} \sum_{x_2} \sum_{y_2} E\Big\{[f(x_1, y_1) - \overline{f(x_1, y_1)}]$$

$$\cdot [f^*(x_2, y_2) - \overline{f(x_2, y_2)}^*]\Big\} \; a(x_1, y_1, u_1, v_1)\, a^*(x_2, y_2, u_2, v_2) \qquad (3\text{-}7)$$

The expected value of the bracketed term in the summation of Equation (3-7) is by definition the spatial domain covariance function, $C\{x_1, x_2, y_1, y_2\}$. Hence,

$$C\{u_1, u_2, v_1, v_2\} = \sum_{x_1} \sum_{x_2} \sum_{y_1} \sum_{y_2} C\{x_1, x_2, y_1, y_2\}$$

$$\cdot a(x_1, y_1, u_1, v_1)\, a^*(x_2, y_2, u_2, v_2) \qquad (3\text{-}8)$$

The variance of the transform domain samples is

$$\sigma^2(u, v) = C\{u, u, v, v\} \qquad (3\text{-}9)$$

Therefore, the general expression for the variance of transform domain samples becomes

$$\sigma^2(u, v) = \sum_{x_1} \sum_{x_2} \sum_{y_1} \sum_{y_2} C\{x_1, x_2, y_1, y_2\}$$

$$\cdot a(x_1, y_1, u, v)\, a^*(x_2, y_2, u, v) \qquad (3\text{-}10)$$

There are two special cases of interest. For an image that is statistically stationary in the spatial domain

$$C\{x_1, x_2, y_1, y_2\} = C\{x_1 - x_2,\ y_1 - y_2\} \tag{3-11}$$

If the original image is uncorrelated in the horizontal and vertical directions

$$C\{x_1, x_2, y_1, y_2\} = C_x\{x_1, x_2\}\, C_y\{y_1, y_2\} \tag{3-12}$$

and the transform domain variance can be computed as $\sigma^2(u,v) = \sigma^2(u)\ \sigma^2(v)$ provided that the transform kernel is separable. Further investigation of the variance of transform domain samples requires specification of the transform.

Fourier Transform

For the Fourier transform the variance function of Equation (3-10) can be written as

$$\sigma^2(u,v) = \frac{1}{N^2} \sum_{x_1} \sum_{x_2} \sum_{y_1} \sum_{y_2} C\{x_1, x_2, y_1, y_2\} \exp\left\{ -\frac{2\pi i}{N} [u(x_1 - x_2) + v(y_1 - y_2)] \right\} \tag{3-13}$$

Consider the case for which the original image is stationary and orthogonally uncorrelated. The variance function reduces to

$$\sigma^2(u,v) = \sigma^2(u)\ \sigma^2(v) \tag{3-14}$$

where

$$\sigma^2(\rho) = \frac{1}{N} \sum_{z_1} \sum_{z_2} C\{ z_1 - z_2\} \exp\left\{ - \frac{2\pi i}{N} \rho (z_1 - z_2) \right\} \tag{3-15}$$

with ρ = u or v and $z_i = x_i$ or y_i. The coordinate variance function may then be rewritten as

$$\sigma^2(\rho) = \frac{1}{N} \sum_{z_2} \exp\left\{ \frac{2\pi i}{N} \rho z_2 \right\} \sum_{z_1} C\{ z_1 - z_2\} \exp\left\{ - \frac{2\pi i}{N} \rho z_1 \right\} \tag{3-16}$$

The second summation is the one dimensional discrete Fourier transform of the covariance function shifted by z_2. By the Fourier transform translation theorem

$$\sigma^2(\rho) = \frac{1}{N} \sum_{z_2} \exp\left\{ \frac{2\pi i}{N} \rho z_2 \right\} \exp\left\{ - \frac{2\pi i}{N} \rho z_2 \right\} G(\rho) \tag{3-17a}$$

or

$$\sigma^2(\rho) = G(\rho) \tag{3-17b}$$

where $G(\rho)$ and $C\{ z_1 \}$ are one dimensional discrete Fourier transform pairs. If the transform is over a complete image dimension, $G(\rho)$ is the discrete version of the power spectral density, $S_z(\rho)$, of the image function along one coordinate minus the average image power, $S_z(0)$, along the coordinate. Hence,

$$\sigma^2(\rho) = S_z(\rho) - S_z(0) \tag{3-18}$$

Thus, the transform domain sample variance along a coordinate direction is directly proportional to the power spectral density of the image along the corresponding orthogonal coordinate. If the original image function can be considered to be a Gauss-Markov process, the covariance function is [39]

$$C_z(z_1-z_2) = C_z(0) \exp\left\{ -\gamma |z_1-z_2| \right\} \tag{3-19}$$

where $C_z(0)$ is a scaling constant and γ is a shape constant. Then

$$S_z(\rho)-S_z(0) = C_z(0) \left[\frac{2\gamma}{\gamma^2 + \rho^2} \right] \tag{3-20}$$

For the Markov process example, the transform domain variance becomes

$$\sigma^2(u,v) = C_x(0)\, C_y(0) \quad \frac{4\alpha\beta}{(\alpha^2 + u^2)(\beta^2 + v^2)} \tag{3-21}$$

where $C_x(0)$ and $C_y(0)$ are the magnitude scaling constants and the shape constants of the spatial domain covariance function, respectively.

Hadamard transform

From Equations (3-10) and (2-20) the variance function of transform domain samples for the ordered Hadamard transform is

$$\sigma^2(u,v) = \frac{1}{N} \sum_{x_1} \sum_{x_2} \sum_{y_1} \sum_{y_2} C\{x_1, x_2, y_1, y_2\}$$

$$\cdot (-1)^{\sum_{i=0}^{n-1} \left[g_i(u)(x_{1i} + x_{2i}) + g_i(v)(y_{1i} + y_{2i}) \right]} \tag{3-22}$$

Since the Hadamard transform does not possess a sequency shifting property it is not possible to reduce Equation (3-22) to closed form.

Karhunen-Loeve Transform

The general expression for the variance of the variance of transform samples given by Equation (3-10) can be rewritten as

$$\sigma^2(u,v) = \sum_{x_2} \sum_{y_2} a^*(x_2, y_2, u, v) \sum_{x_1} \sum_{y_1} C\{x_1, x_2, y_1, y_2\}\, a(x_1, y_1, u, v) \tag{3-23}$$

For the Karhunen-Loeve transform from Equation (2-43) the second set of summations defines the transform kernel. Thus,

$$\lambda(\theta,\varphi)\, a(x_2, y_2, \theta, \varphi) = \sum_{x_1} \sum_{y_1} C\{x_1, x_2, y_1, y_2\}\, a(x_1, y_1, \theta, \varphi) \tag{3-24}$$

where $\lambda(\theta,\varphi)$ are the eigenvalues of the covariance matrix. By this equivalence

$$\sigma^2(u,v) = \sum_{x_2}\sum_{y_2} \lambda(\theta,\varphi) \quad a(x_2,y_2,\theta,\varphi)\, a^*(x_2,y_2,u,v) \tag{3-25}$$

Since the Karhunen-Loeve transform is an orthogonal transformation, from Equation (2-6c)

$$\sigma^2(u,v) = \lambda(u,v) \tag{3-26}$$

and the variance of each transform sample is equal to its corresponding eigenvalue.

3.2 Probability Densities

It would be desirable to know the probability density of transform samples for an arbitrary image transform. Unfortunately, this result is not easily obtained since the original image probability density is not usually well defined, and also, the transform operation is quite often mathematically complex. However, the transforms considered for image processing applications form a weighted sum over all of the elements in the original image. Therefore, one can evoke qualitative arguments based upon the Central Limit Theorem of statistics that the probability density of transform samples tends to be Gaussian with moments as calculated in the previous section. For the subsequent analysis, a Gaussian model is developed for the probability density of the Fourier, Hadamard, and Karhunen-Loeve transform domain samples.

Fourier transform samples are complex numbers which may be represented in real and imaginary, or magnitude and phase, form. In either case there are two components per transform sample that must be quantized. The real, $F_R(u,v)$ and imaginary, $F_I(u,v)$, components of the Fourier transform samples may be assumed to follow the same Gaussian distribution whose variance, $\sigma^2(u,v)$, is proportional to the power spectral density of the original image. Hence,

$$p\left\{F_R(u,v)\right\} = [2\pi\sigma^2(u,v)]^{-\frac{1}{2}} \exp\left\{\frac{-F_R^2(u,v)}{2\sigma^2(u,v)}\right\} \tag{3-27}$$

$$p\left\{F_I(u,v)\right\} = [2\pi\sigma^2(u,v)]^{-\frac{1}{2}} \exp\left\{\frac{-F_I^2(u,v)}{2\sigma^2(u,v)}\right\} \tag{3-28}$$

If the real and imaginary components are Gaussian, the magnitude of the Fourier transform sample, $F_M(u,v)$, is Rayleigh distributed

$$p\left\{F_M(u,v)\right\} = \frac{F_M(u,v)}{\sigma^2(u,v)} \exp\left\{\frac{-F_M^2(u,v)}{2\sigma^2(u,v)}\right\} \quad F_M(u,v) > 0 \tag{3-29}$$

and its phase, $F_P(u,v)$, is uniformly distributed

$$p\left\{F_P(u,v)\right\} = \frac{1}{2\pi} \qquad -\pi \leq F_P \leq +\pi \tag{3-30}$$

Hadamard transform samples are real, bipolar numbers which can be represented by a single component per sample. The statistical distribution of Hadamard sample components, $F_H(u,v)$, may be considered to follow a Gaussian distribution of the form

$$p\left\{F_H(u,v)\right\} = [2\pi\sigma^2(u,v)]^{-\frac{1}{2}} \exp\left\{\frac{-F_H^2(u,v)}{2\sigma^2(u,v)}\right\} \tag{3-31}$$

Karhunen-Loeve transform samples are also real bipolar numbers. The probability density of the samples may be modeled as

$$p\left\{F_K(u,v)\right\} = [2\pi\sigma^2(u,v)]^{-\frac{1}{2}} \exp\left\{\frac{-F_K^2(u,v)}{2\sigma^2(u,v)}\right\} \tag{3-32}$$

When the variance function, $\sigma^2(u,v)$, is not known for a particular image, or class of images, to be transformed, the function can usually be modeled without seriously affecting the quantization process. From examination of the Fourier, Hadamard, and Karhunen-Loeve transforms of a typical image, it can be deduced that the variance function should be a maximum at the origin in the transform domain, be circularly symmetric, and decrease in magnitude monotonically toward the higher spatial frequencies. A two dimensional function processing these

characteristics is the Gaussian shaped curve described by

$$\sigma^2(u, v) = S^2 \exp\left\{ - \frac{u^2 + v^2}{p/2} \right\} \tag{3-33}$$

where S is an amplitude scaling constant and p is a spread control constant. Another useful function for modeling of the variance function is

$$\sigma^2(u, v) = \frac{S^2}{(u^2 + \alpha^2)(v^2 + \beta^2)} \tag{3-34}$$

where S is an amplitude scaling constant and α and β are spread control constants. This model holds exactly for the Fourier transform if the original image can be considered as a Gauss-Markov process source.

4. Generalized Transform Coding

The basic premise of image transform coding 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 element-to-element correlation of natural images, for many image transforms, the energy in the transform domain tends to be clustered in a relatively few number of transform samples. This property can be exploited to achieve a sample reduction compared to conventional spatial domain coding.

There are two methods of obtaining a sample reduction by transform coding--zonal sampling and threshold sampling. In zonal sampling the image reconstruction is made with a subset, usually the lowest spatial coefficients, of the transform domain samples. Those samples which are employed in the reconstruction are chosen before the transformation on the basis of expected energy. With threshold sampling the reconstruction is made with a subset of the largest magnitude transform domain samples.

This section presents a discussion of the performance of the Karhunen-Loeve, Fourier, and Hadamard transforms for zonal and threshold sampling in the transform domain. The three transforms are compared on the basis of minimum mean square error. Experimental results are presented for a subjective comparison.

4.1 Generalized Zonal Sampling

Optimum Zonal Sampling

Consider an image transform of the first kind. With zonal sampling the image is reconstructed with the first M of the N^2 transform samples. Thus, the reconstructed image becomes

$$\hat{f}(x, y) = \sum_{w=0}^{M-1} F(w)\, b(x, y, w) \tag{4-1}$$

The mean square error is then given by

$$\mathcal{E}_s = \frac{1}{N^2} E\left\{ \sum_x \sum_y [f(x, y) - \hat{f}(x, y)]^2 \right\} \tag{4-2}$$

or

$$\mathcal{E}_s = \frac{1}{N^2} \sum_x \sum_y E\{f^2(x, y)\} - \frac{2}{N^2} \sum_x \sum_y E\{f(x, y)\, \hat{f}(x, y)\}$$

$$+ \frac{1}{N^2} \sum_x \sum_y E\{\hat{f}^2(x, y)\} \tag{4-3}$$

The first term above is the spatial domain autocorrelation function $R(0, 0, 0, 0)$. The other terms may be evaluated by substitution of the reverse transforms yielding

$$\delta_s = R(0,0,0,0) - \frac{2}{N^2} \sum_x \sum_y E \left\{ \left[\sum_{w=0}^{N^2-1} F(w)\, b(x,y,w) \right] \cdot \left[\sum_{w'=0}^{M-1} F(w')\, b(x,y,w') \right] \right\} + \frac{1}{N^2} \sum_x \sum_y E \left\{ \left[\sum_{w=0}^{M-1} F(w)\, b(x,y,w) \right] \cdot \left[\sum_{w'=0}^{M-1} F(w')\, b(x,y,w') \right] \right\} \tag{4-4}$$

Expanding the series and changing the order of summation gives

$$\delta_s = R(0,0,0,0) - \frac{2}{N^2} E \left\{ \underbrace{\sum_{w=0}^{N^2-1} \sum_{w'=0}^{M-1}}_{w \neq w'} F(w)F(w') \sum_x \sum_y b(x,y,w)\, b(x,y,w') \right\}$$

$$+ \frac{1}{N} E \left\{ \sum_{w=0}^{M-1} \sum_{w'=0}^{M-1} F(w)\, F(w') \sum_x \sum_y b(x,y,w)\, b(x,y,w') \right\} \tag{4-5}$$

By the orthogonality of the kernel $b(x,y,w)$

$$\delta_s = R(0,0,0,0) - \frac{2}{N^2} E \left\{ \sum_{w=0}^{N^2-1} \sum_{w'=0}^{M-1} F(w)\, F(w')\, \delta(w-w') \right\}$$

$$+ \frac{1}{N^2} E \left\{ \sum_{w=0}^{M-1} \sum_{w'=0}^{M-1} F(w)\, F(w')\, \delta(w-w') \right\} \tag{4-6}$$

Thus,

$$\delta_s = R(0,0,0,0) - \frac{1}{N^2} \sum_{w=0}^{M-1} E\{ F^2(w) \} \tag{4-7a}$$

or

$$\delta_s = C(0,0,0,0) - \frac{1}{N^2} \sum_{w=0}^{M-1} \sigma^2(w) \tag{4-7b}$$

For a given number, M, of samples to be included in the transformation, the mean square error will be small if the variance of each of the transform samples is large. The transform which minimizes the mean square error is the Karhunen-Loeve transform of the first kind in which the eigenvectors are arranged in correspondence with the eigenvalues in descending order [30, 31].

For an image transform of the second kind, with zonal filtering, the reconstructed image is given by

$$\hat{f}(x,y) = \underbrace{\sum_{u}^{N-1} \sum_{v}^{N-1}}_{u,v \in M(u,v)} F(u,v)\, b(x,y,u,v) \tag{4-8}$$

where the transform domain indices are members of a set determined by a mask function M(u, v). By an analysis similar to that for an image transform of the first kind, it is found that the mean square error is of the form

$$\delta_S = C(0,0,0,0) - \frac{1}{N^2} \underbrace{\sum_{u=0}^{N-1} \sum_{v=0}^{N-1}}_{u,v \in M(u,v)} \sigma^2(u,v) \tag{4-9}$$

The transform which minimizes the mean square error is the Karhunen-Loeve transform of the second kind for which the mask function corresponds to the index pairs u, v that have the largest eigenvalues.

Karhunen-Loeve Transform

For the Karhunen-Loeve transform of the first kind, the minimum mean square error becomes

$$\delta_S = C(0,0,0,0) - \frac{1}{N^2} \sum_{w=0}^{M-1} \lambda(w) \qquad (4\text{-}10)$$

where $\lambda(w)$ represents the eigenvalues of the covariance matrix of the image. The operational procedure for performing zonal sampling with the Karhunen-Loeve transform of the first kind is simply to compute and code only the first M components of the transform which are subsequently to be used in the inverse transform.

The minimum mean square reconstruction error for zonal sampling of the Karhunen-Loeve transform of the second kind is

$$\delta_S = C(0,0,0,0) - \frac{1}{N^2} \underbrace{\sum_{u=0}^{N-1} \sum_{v=0}^{N-1}}_{u,v \in M(u,v)} \lambda(u,v) \qquad (4\text{-}11)$$

It should be noted that the mask function is generally not a simple rectangle in the transform domain. Exhibit 4-1 shows the ordering of the eigenvalues for a Karhunen-Loeve transform of the second kind. In this example the image covariance function is separable, and the vertical and horizontal element correlation is the same. Therefore, the eigenvalues corresponding to each line and column of the image are identical. The eigenvalue products give the same eigenvalues that would be obtained for a Karhunen-Loeve transform of the first kind. However, there is no simple and general ordering between $\lambda(u,v)$ and $\lambda(w)$. Thus, the eigenvalue ordering must be determined experimentally for a given Karhunen-Loeve transform. Figure 4-1 shows 16 by 16 element sampling masks for a 4:1 sample reduction for two values of the image covariance function.

If a rectangular mask function

$$u,v \in M(u,v) \quad \text{if } u \leq u_c ; \; v \leq v_c \qquad \underline{\text{rectangular mask}} \tag{4-12}$$

is employed for ease of implementation, the performance of the operation will not be optimum, but the degradation will usually not be too serious. Two other simple mask functions that could be employed are listed below:

		3.103	.559	.209	.129	$\lambda(u)$
		1	2	3	4	u →
3.103	1	9.630	1.735	.648	.400	
.559	2	1.735	.313	.117	.072	
.209	3	.648	.117	.044	.027	
.129	4	.400	.072	.027	.017	
$\lambda(v)$	v ↓					

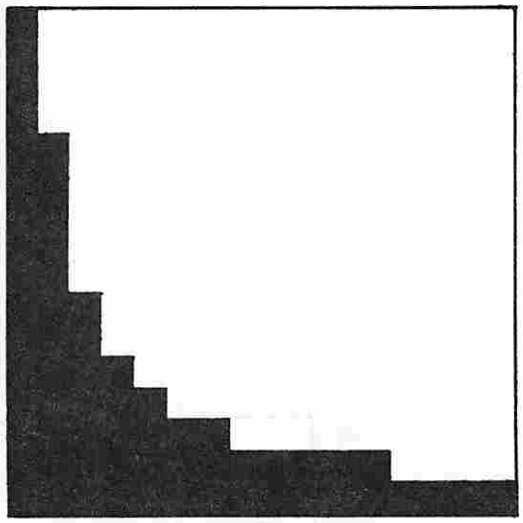
a. Eigenvalue products $\lambda(u,v) = \lambda(u)\,\lambda(v)$

w	1	2	3	4	5	6	7	8
(u,v)	1,1	1,2	2,1	1,3	3,1	1,4	4,1	2,2
$\lambda(w)$	9.630	1.735	1.735	.648	.648	.400	.400	.313

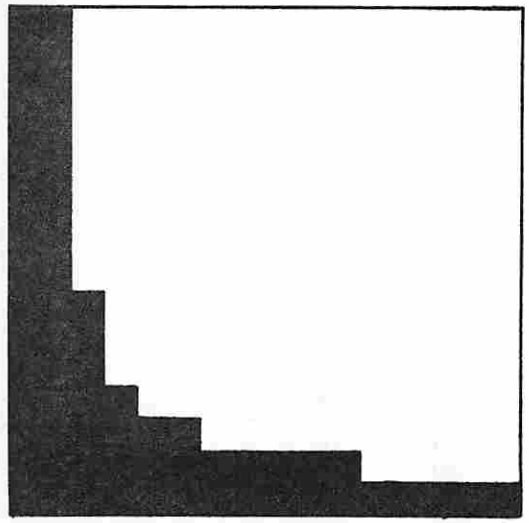
w	9	10	11	12	13	14	15	16
(u,v)	3,2	2,3	4,2	2,4	3,3	4,3	3,4	4,4
$\lambda(w)$	.117	.117	.072	.072	.044	.027	.027	.017

b. Ordering of eigenvalue products

Exhibit 4-1. Karhunen-Loeve Transform Zonal Sampling Mask Generation



a. XC = 0.9
YC = 0.9



b. XC = 0.95
YC = 0.90

$$\text{Image covariance function} -- \frac{C\{x_1, x_2, y_1, y_2\}}{C\{0, 0, 0, 0\}} = (XC)^{|x_1 - x_2|}\,(YC)^{|y_1 - y_2|}$$

Figure 4-1 Karhunen-Loeve Zonal Sampling Masks

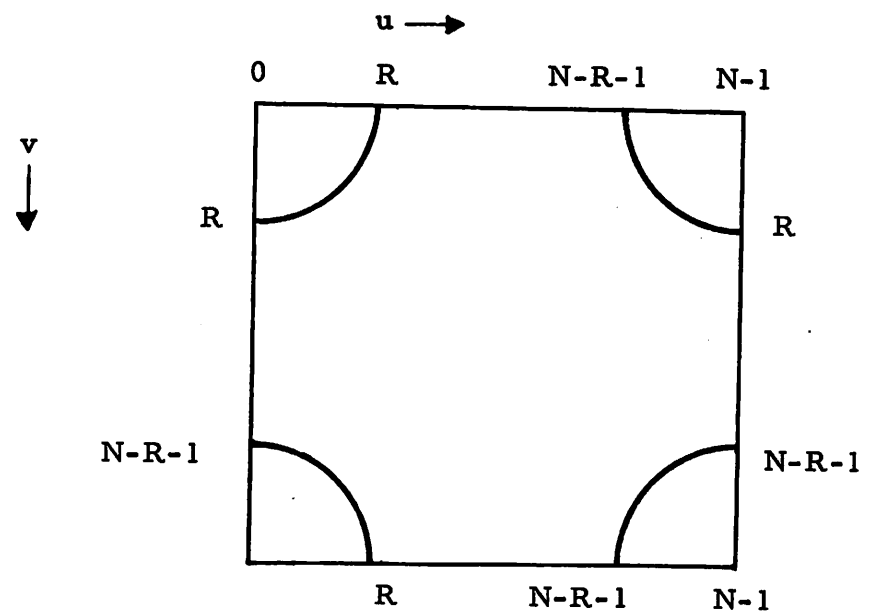
$$u, v \in M(u,v) \quad \text{if} \quad \frac{u^2}{u_c^2} + \frac{v^2}{v_c^2} \leq 1 \qquad \underline{\text{elliptical mask}} \tag{4-13}$$

$$u, v \in M(u,v) \quad \text{if } uv \leq K \qquad \underline{\text{hyperbolic mask}} \tag{4-14}$$

The hyperbolic mask most closely resembles the optimum mask determined by ordering the eigenvalues of the image covariance function.

Fourier Transform

Zonal sampling with the Fourier transform consists of sampling the lowest spatial frequencies in the transform domain. For the Fourier transform defined by Equation (2-20), the lowest spatial frequencies lie in the four zones shown below:



Hadamard Transform

With an ordered Hadamard transform, zonal sampling consists of sampling the transform domain samples with the low sequencies. These samples lie within a circular quadrant about the origin in the transform domain.

4.2 Generalized Threshold Sampling

Zonal sampling in the transform domain will provide small mean square error reconstructions of good subjective quality if the actual magnitude of a transform domain sample does not differ greatly from the standard deviation $\sigma(w)$ or $\sigma(u,v)$. The difficulty with zonal sampling is that in most natural images there are many high spatial frequency samples lying outside the sampling zone that are of significant magnitude. In threshold sampling rather than determining a priori which transform domain samples are to be coded, the selection is made after the transform has been taken on a particular image. A threshold level is established a priori, or perhaps adaptively, and only those samples whose magnitudes are greater than the threshold are coded. If the threshold level is chosen a priori, based upon the probability density of the transform samples, the actual sample reduction factor for a particular image will be variable. As an alternative procedure the threshold level could be chosen so that a given number of

transform domain samples would be coded for a particular image.

If transform domain threshold coding is to be employed, the major question of interest is: What is the optimum image transform? In general the best transform is the transform which maps the image energy into the fewest transform domain samples. For a checkerboard image of half black and half white elements in each direction, the Hadamard transform is a very efficient transform since the image can be represented by only two transform domain samples. For natural images, the image can only be defined statistically, not deterministically. In such instances the optimum (minimum mean square error) transform is the transform for which the smallest number of samples have the largest variances. As stated previously, for a given class of images, this transform is the Karhunen-Loeve transform. Thus, it is expected that the Karhunen-Loeve transform would exhibit the best minimum mean square error performance for threshold coding for a given class of images. However, for a particular image of the class, another transform could provide better performance.

4.3 Image Block Size Considerations

For either zonal or threshold sampling in the transform domain, consideration must be given to the size of the image block. From the standpoint of image energy it is best to make the image block size as large as possible in order to derive benefit from all element-to-element correlations within the image. For natural images, however, the correlation between elements separated by over 10 to 20 elements is usually relatively small (see Appendix A). Therefore, little is lost in taking the image transform over smaller size blocks. This point is illustrated by Figure 4-2 which contains a plot of the percentage of transform domain energy contained in the lowest one-fourth of the transform domain samples for a one dimensional Karhunen-Loeve transform as a function of block size. In this example the image covariance function is modeled as a Gauss-Markov process dependent only upon the adjacent element correlation factor, XC. As indicated in Figure 4-2 about 90% of the image energy is contained in the lowest one-fourth of the Karhunen-Loeve transform coefficients for a block size of 16 by 16 elements. The percentage of energy contained in the low pass zone increases rather slowly for larger size blocks and decreases much more rapidly for smaller size blocks. It appears that a block size of about 16 by 16 elements

$$Y = \sum_{u=1}^{N/4} \frac{\lambda_N(u)}{N}$$

$$\frac{C\{x_1, x_2\}}{C\{0,0\}} = (XC)^{|x_1 - x_2|} \text{ --- normalized image covariance function}$$

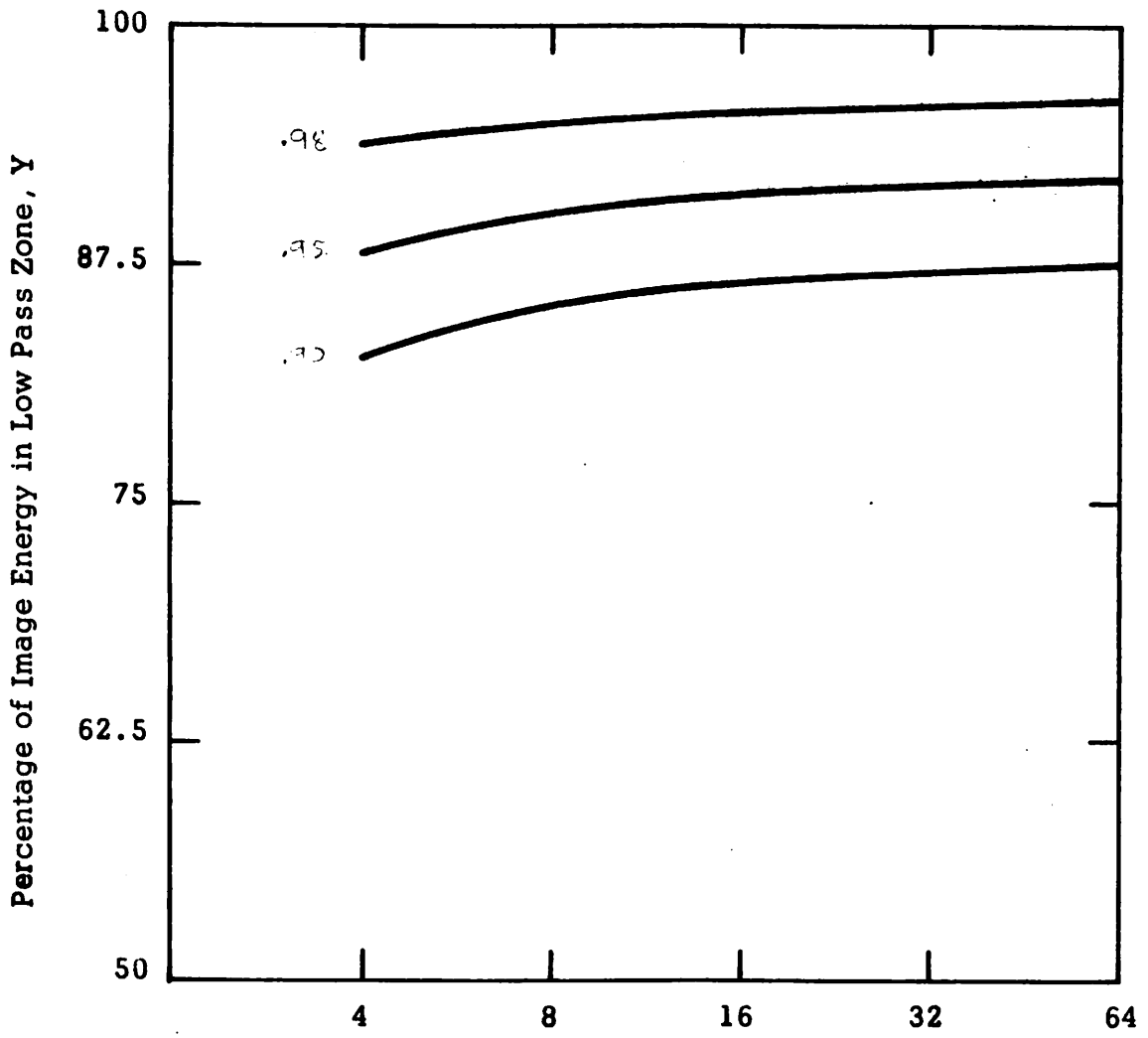


Figure 4-2 Effect of Image Block Size for Karhunen-Loeve Transform

is a good compromise between maximizing the amount of image compression possible and simplifying the implementation of the transform coder.

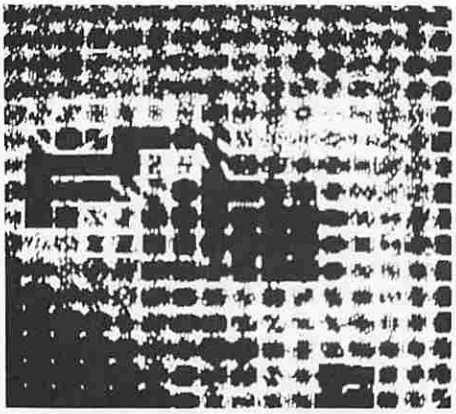
4.4 Comparison of Image Transforms

A series of experiments has been conducted to determine the image coding performance of the Fourier, Hadamard, and Karhunen-Loeve transforms for natural images. As a result of the computational requirements of the Karhunen-Loeve transform, the image block size was limited to 16 by 16 elements.*

Figure 4-3, 4-4, and 4-5 contain displays of the Fourier, Hadamard, and Karhunen-Loeve transforms. It should be noted that there is no apparent grid structure in the reconstructed image despite the block processing.

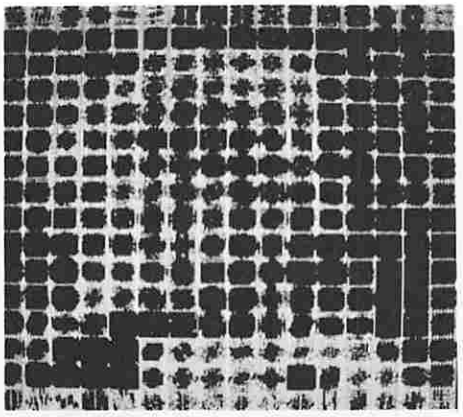
Figures 4-6 to 4-10 illustrate the effect of zonal low pass filtering for the three transforms. In Figures 4-6 and 4-8 the filter pass band for the Fourier and Hadamard transforms is a circular zone in the transform domain. For sample reduction factors greater than 4:1 the 16 by 16 element grid structure becomes apparent because many of the transform samples that correspond to

*Examples of Fourier and Hadamard transform coding in larger size blocks are presented in Section 6.



a. Threshold display of Surveyor box transform

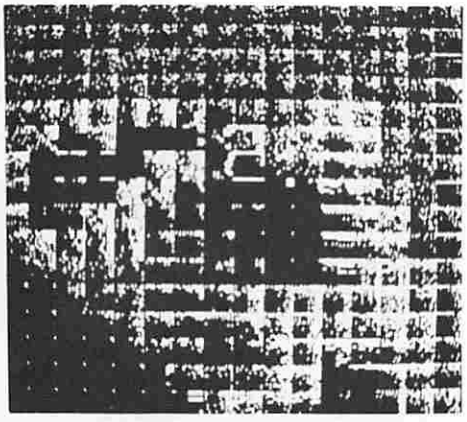
b. Inverse transform of transform



c. Threshold display of Girl transform

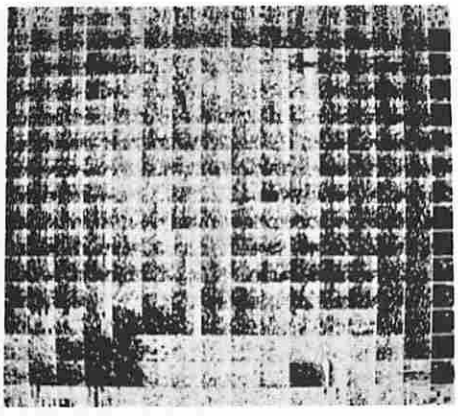
d. Inverse transform of transform

Figure 4-3 Fourier Transforms in 16 x 16 Element Blocks



a. Threshold display of Surveyor box transform

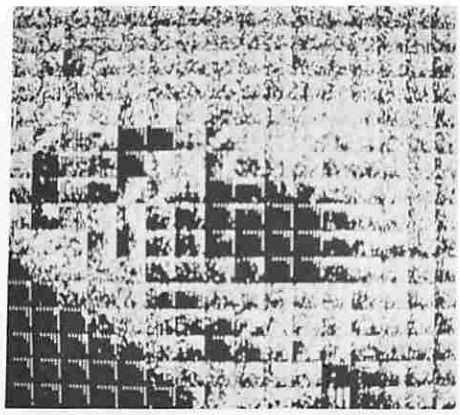
b. Inverse transform of transform



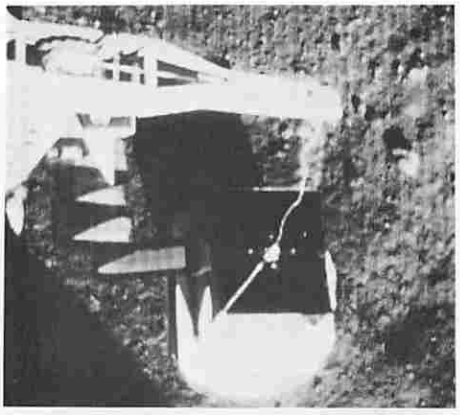
c. Threshold display of Girl transform

d. Inverse transform of transform

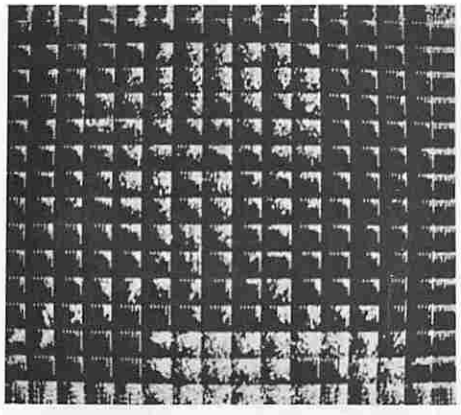
Figure 4-4 Hadamard Transforms in 16 x 16 Element Blocks



a. Threshold display of Surveyor box transform



b. Inverse transform of transform



c. Threshold display of Girl transform

d. Inverse transform of transform

Figure 4-5 Karhunen-Loeve Transforms in 16 x 16 Element Blocks

a. 2:1 sample reduction

b. 4:1 sample reduction

c. 8:1 sample reduction

d. 11:1 sample reduction

Figure 4-6 Fourier Transform Zonal Sampling in 16 x 16 Element Blocks -- Circular Zone

a. 2:1 sample reduction

b. 4:1 sample reduction

c. 8:1 sample reduction

d. 13:1 sample reduction

Figure 4-7 Fourier Transform Zonal Sampling in 16 x 16 Element Blocks -- Hyperbolic Zone

a. 2:1 sample reduction

b. 4:1 sample reduction

c. 8:1 sample reduction

d. 12:1 sample reduction

Figure 4-8 Hadamard Transform Zonal Sampling in 16 x 16 Element Blocks -- Circular Zone

a. 2:1 sample reduction

b. 4:1 sample reduction

c. 8:1 sample reduction

d. 12:1 sample reduction

Figure 4-9 Hadamard Transform Zonal Sampling in 16 x 16 Element Blocks -- Hyperbolic Zone

a. 2:1 sample reduction

b. 4:1 sample reduction

c. 8:1 sample reduction

d. 12:1 sample reduction

Figure 4-10 Karhunen-Loeve Transform Zonal Sampling in 16 x 16 Element Blocks

brightness changes in periods of 16 elements are excluded from the circular pass band. To prevent this grid effect a hyperbolic shaped zone similar to the sampling mask of Figure 4-1a was employed. The results shown in Figures 4-7 and 4-9 show a definite improvement in the elimination of the grid structure as compared to Figures 4-6 and 4-8. In all four images there is an expected loss of resolution. Figure 4-10 illustrates zonal low pass filtering with the Karhunen-Loeve transform. The filter is a mask passing those transform domain samples corresponding to the largest eigenvalues of the image covariance function. The reconstructed images do not show the 16 by 16 element grid structure, but there is some loss in resolution. Summarizing these results: for a given sample reduction factor the Karhunen-Loeve transform results in the smallest mean square error and the least image degradation from a subjective viewpoint. With a hyperbolic shaped pass band the Fourier transform is somewhat better than the Hadamard transform for both measures of image quality.

Figures 4-11, 4-12, and 4-13 show the effects of threshold coding in the transform domain for the three types of image transforms. The quality rating of the three transforms from the standpoint of subjective quality is: Karhunen-Loeve, first; Hadamard, second; and Fourier, third. It should be noted that the sample reduction factors obtained for equivalent image quality

a. 5:1 sample reduction

b. 10:1 sample reduction

c. 20:1 sample reduction

Figure 4-11 Fourier Transform Threshold Sampling in 16 x 16 Element Blocks

a. 5:1 sample reduction

b. 10:1 sample reduction

c. 20:1 sample reduction

Figure 4-12 Hadamard Transform Threshold Sampling in 16 x 16 Element Blocks

a. 5:1 sample reduction

b. 10:1 sample reduction

c. 20:1 sample reduction

Figure 4-13 Karhunen-Loeve Transform Threshold Sampling in 16 x 16 Element Blocks

are much higher for threshold sampling than zonal sampling. For a fair comparison, however, to account for position code bits, the sample reduction factors for threshold coding should be multiplied by a factor of about 0.6 to 0.8 to obtain equivalent bandwidth reduction factors.

In summary the following conclusions can be drawn from these experiments:

a. For both zonal and threshold sampling in the transform domain, the best transform is the Karhunen-Loeve transform, followed by the Hadamard transform, followed by the Fourier transform.

b. Threshold sampling provides higher sample and bandwidth reduction factors than zonal sampling for all three transforms.

c. The effect of a limited block size does not appear to be a serious problem either from the standpoint of image quality or performance.

While the Karhunen-Loeve transform does appear to provide better performance than the Fourier and Hadamard transforms, the margin of performance is not too large. In view of the considerably greater amount of computation involved with the Karhunen-Loeve transform as compared to the Fourier and Hadamard transforms, its utilization will probably be limited. The following sections are therefore restricted in scope to the Fourier and Hadamard transforms. These sections present an analysis of transform domain quantization, a further discussion of image coding for bandwidth reduction, and a study of the error tolerance properties of image transforms.

5. Fourier and Hadamard Image Transform Quantization

The dynamic range of Fourier and Hadamard transform domain samples in integer arithmetic is 1 to N^2A where N denotes the number of elements per line of the image and A is the maximum integer value of the amplitude of an image sample. If each transform domain sample were simply coded in a binary code, $\log_2(N^2A)$ bits would be required for each code word. For a 256 x 256 element image of 64 grey levels, each code word would be 22 bits in length. Even with threshold coding it would be unlikely that a significant bandwidth compression could be achieved for such large length code words. In order to achieve a bandwidth compression with transform coding it is necessary to recode, or quantize, each transform domain sample so that it may be represented by relatively short length code words. There are two basic approaches to this process: each sample could be quantized to the same number of levels, with the quantization levels possibly chosen according to a nonlinear scale; or the number of levels could be permitted to vary from sample to sample with a linear spacing of quantization levels. The latter approach will result in the most efficient coding, but the code words will be of variable length. This creates problems in data synchronization and channel coding for error correction. The former method can be adapted for relatively efficient constant word length coding.

In the subsequent discussion only the first method of quantization is considered.

5.1 Quantization Scales

In the quantization process let the transform sample component (amplitude, real part, imaginary part, magnitude, or phase) to be quantized be represented by the function $F_c(u,v)$. The range of the sample component is assumed to be broken up into K positive and K negative bands separated by quantization levels $Q_j (j = 0, \pm 1, \pm 2, \cdots, \pm K)$. The zeroth quantization level and the upper and lower quantization levels are assigned the values

$$Q_0 = 0 \tag{5-1a}$$

$$Q_K = \frac{NA}{2} \tag{5-1b}$$

$$Q_{-K} = \frac{-NA}{2} \tag{5-1c}$$

where A represents the maximum value of a sample of the original image of N by N elements. If a transform component falls in a band bounded by quantization levels Q_{j-1} and Q_j, the component is quantized, and subsequently reconstructed, to the value $F_j(u,v)$ which lies within the band. The relationship between quantization and reconstruction levels is given in Figure 5-1.

Quantization and reconstruction levels are logically chosen

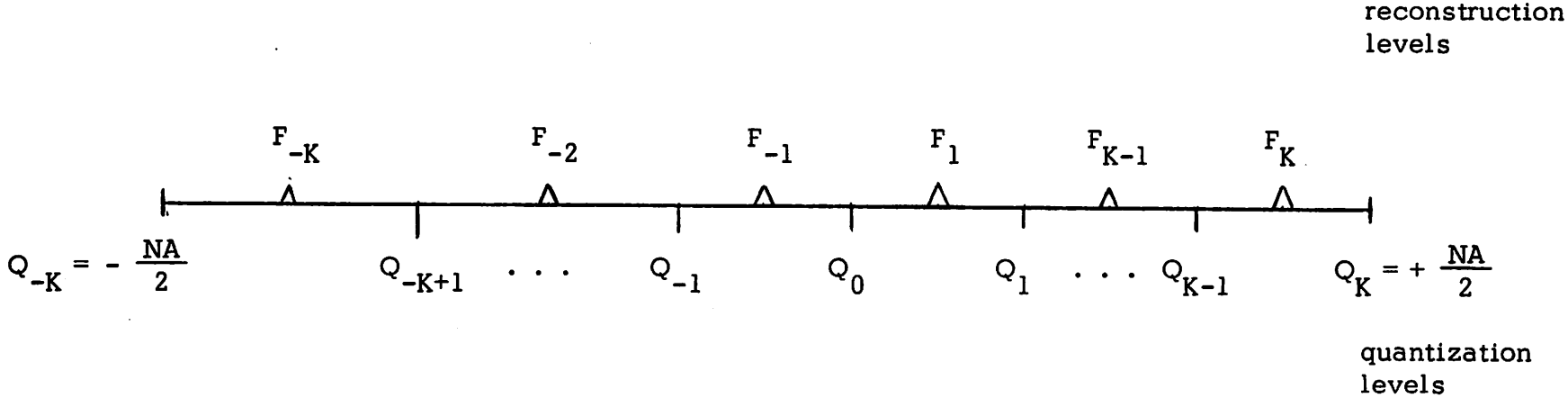


Figure 5-1 Quantization and Reconstruction Levels

to minimize the effects of the quantization error introduced by the amplitude truncation of samples. Table 5-1 lists some error criteria for the selection of quantization and reconstruction levels. The quantization error criterion depends upon the application of a reconstructed image. The principal consideration is whether the image is to be used for subjective viewing or photometric measurements.

For subjective viewing the relative spatial error criterion listed in Table 5-1 provides an indication of image quality. This relative spatial error criterion is predicated upon the fact that incremental brightness changes in the reconstructed image are much more noticeable if the brightness level is low than if it is high. Thus, to minimize the relative spatial error, the density of quantization levels in the spatial domain should be greater at the lower amplitude levels. But, since the brightness of every point of a reconstructed image is a function of the amplitude of a single transform sample, then by the same reasoning, the density of quantization levels should be greater for low level transform samples. From psychophysical tests, it is known that the human viewer is very sensitive to the location of high frequency brightness transitions, but relatively insensitive to their actual magnitude. In fact images which have been "crispened" by high pass filtering often appear preferable to the original image. From this characteristic

TABLE 5-1

Quantization Error Criteria

Cumulative mean square spatial error	$\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\left[f(x,y)-\tilde{f}(x,y)\right]^2$
Cumulative mean square transform error	$\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\left[F(u,v)-\tilde{F}(u,v)\right]^2$
Cumulative spatial error	$\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\left\|f(x,y)-\tilde{f}(x,y)\right\|$
Cumulative transform error	$\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\left\|F(u,v)-\tilde{F}(u,v)\right\|$
Relative spatial error	$\frac{\left\|f(x,y)-\tilde{f}(x,y)\right\|}{\left\|f(x,y)\right\|}$
Relative transform error	$\frac{\left\|F(u,v)-\tilde{F}(u,v)\right\|}{\left\|F(u,v)\right\|}$

$\tilde{F}(u,v)$ = quantized value of $F(u,v)$

$\tilde{f}(x,y)$ = inverse transform of quantized value of $F(u,v)$

of subjective viewing it would seem that the density of quantization levels, at low transform sample amplitudes, should be greater at the higher spatial frequencies than at the lower spatial frequencies.

If photometric measurements are to be made on an image the cumulative mean square spatial error is a common fidelity criterion. For a mean square error criterion the quantization levels in the transform domain must be selected to minimize the cumulative mean square error in the spatial domain. Let

$$\delta_s \equiv \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} E\left\{[f(x,y)-\tilde{f}(x,y)]^2\right\} \tag{5-2}$$

represent the cumulative mean square spatial domain error where $\tilde{f}(x,y)$ is the image reconstruction from the quantized transform samples, $\hat{F}(u,v)$. For a Fourier or Hadamard transform

$$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v)\, b(x,y,u,v) \tag{5-3}$$

and

$$\tilde{f}(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \tilde{F}(u,v)\, b(x,y,u,v) \tag{5-4}$$

Hence, the spatial domain mean square error can be written as

$$\delta_s = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} E\left\{\left[\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} [F(u,v)-\tilde{F}(u,v)] \cdot b(x,y,u,v)\right]^2\right\} \tag{5-5}$$

Expanding the integrand yields

$$\delta_s = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} E\left\{\left[\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{u'=0}^{N-1} \sum_{v'=0}^{N-1} [F(u,v)-\tilde{F}(u,v)] \cdot [F(u',v')-\tilde{F}(u',v')]\, b(x,y,u,v)\, b(x,y,u',v')\right]\right\} \tag{5-6}$$

Rearranging the order of summation gives

$$\delta_s = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{u'=0}^{N-1} \sum_{v'=0}^{N-1} E\left\{\left([F(u,v)-\tilde{F}(u,v)] \cdot [F(u',v')-\tilde{F}(u',v')] \cdot \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} b(x,y,u,v)\, b(x,y,u',v')\right)\right\} \tag{5-7}$$

As a result of the orthogonality of the transform kernel,

$$\delta_s = \frac{1}{N^2} \sum_{u} \sum_{v} \sum_{u'} \sum_{v'} E\left\{[F(u,v)-\tilde{F}(u,v)]\,[F(u',v')-F(u',v')] \cdot \delta(u-u', v-v')\right\} \tag{5-8a}$$

or

$$\delta_s = \frac{1}{N^2} \sum_{u} \sum_{v} E\left\{[F(u,v)-\tilde{F}(u,v)]^2\right\} \tag{5-8b}$$

The cumulative mean square error in the spatial domain of the reconstructed image is therefore equal to the cumulative mean square error in the transform domain. Minimization of δ_s can then be accomplished by minimization of the mean square error

$$\mathcal{E}(u,v) \equiv E\left\{[F(u,v)-\tilde{F}(u,v)]^2\right\} \tag{5-9}$$

in the transform domain for all spatial frequencies. For the Hadamard transform, quantization and reconstruction levels for the transform sample <u>amplitude</u> must be found to minimize $\mathcal{E}(u,v)$. In the case of the Fourier transform the mean square error of the <u>real and imaginary</u>, or <u>phase and magnitude</u>, components of a transform must <u>each</u> be minimized. The mean square error of a transform component may be written in explicit form as

$$\mathcal{E}(u,v) = \mathcal{E}_+(u,v) + \mathcal{E}_-(u,v) \tag{5-10}$$

in which

$$\mathcal{E}_+(u,v) = \sum_{j=1}^{K} \int_{Q_{j-1}(u,v)}^{Q_j(u,v)} [F_c(u,v)-F_j(u,v)]^2 \, p(F_c)\, dF_c \tag{5-11a}$$

and

$$\delta_{-}(u,v) = \sum_{j=-1}^{-K} \int_{Q_{j+1}(u,v)}^{Q_j(u,v)} [F_c(u,v) - F_j(u,v)]^2 \, p(F_c) \, dF_c \qquad (5\text{-}11b)$$

where $p(F_c)$ is the probability density of the transform sample component to be quantized. If $p(F_c)$ is a symmetrical probability density about $Q_0 = 0$, then $\delta_{+}(u,v)$ equals $\delta_{-}(u,v)$, and the quantization rule determined by the minimization of $\delta_{+}(u,v)$ is the same as that determined from $\delta_{-}(u,v)$.

The optimum placement of the quantization and reconstruction levels to minimize the mean square error of a quantized signal has been determined by Panter and Dite [40]. The reconstruction levels should be located midway between each pair of quantization levels. Thus

$$F_j(u,v) = \frac{Q_j(u,v) + Q_{j+1}(u,v)}{2} \qquad (5\text{-}12)$$

The quantization levels can be determined to a good approximation [40] from

$$Q_j(u,v) = \frac{\frac{NA}{2} \int_0^{j \frac{NA}{2}} \left(p\{F_c\} \right)^{-\frac{1}{3}} dF_c}{\int_0^{\frac{NA}{2}} \left(p\{F_c\} \right)^{-\frac{1}{3}} dF_c} \qquad (5\text{-}13)$$

Three cases of interest for quantization of the Fourier and Hadamard transforms are listed below.

Uniform distribution:

$$p\{F_c\} = \frac{1}{NA} \tag{5-14a}$$

$$Q_j(u,v) = j\,\frac{NA}{2K} \tag{5-14b}$$

Rayleigh distribution:

$$p\{F_c\} = \frac{F_c}{\sigma^2(u,v)}\ \exp\left\{-\frac{F_c^2(u,v)}{2\sigma^2(u,v)}\right\} \tag{5-15a}$$

$$Q_j(u,v) = \frac{\frac{NA}{2}\int_0^{j\frac{NA}{2K}} [F_c(u,v)]^{-\frac{1}{3}}\ \exp\left\{\frac{F_c^2(u,v)}{6\sigma^2(u,v)}\right\}\ dF_c}{\int_0^{\frac{NA}{2}} [F_c(u,v)]^{-\frac{1}{3}}\ \exp\left\{\frac{F_c^2(u,v)}{6\sigma^2(u,v)}\right\}\ dF_c} \tag{5-15b}$$

Gaussian distribution:

$$p\{F_c\} = [2\pi\sigma^2(u,v)]^{-\frac{1}{2}}\ \exp\left\{-\frac{F_c^2(u,v)}{2\sigma^2(u,v)}\right\} \tag{5-16a}$$

$$Q_j(u,v) = \frac{\frac{NA}{2}\int_0^{j\frac{NA}{2K}} \exp\left\{\frac{F_c^2(u,v)}{6\sigma^2(u,v)}\right\}\ dF_c}{\int_0^{\frac{NA}{2}} \exp\left\{\frac{F_c^2(u,v)}{6\sigma^2(u,v)}\right\}\ dF_c} \tag{5-16b}$$

The quantization scales determined by Equations (5-15b) and (5-16b) for the Rayleigh and Gaussian distributions have the desired subjective property that the quantization levels are more closely spaced at the lower quantization levels, and also are more closely spaced at the higher spatial frequencies for which the variance $\sigma^2(u, v)$ is smaller. Unfortunately, the quantization levels are nonlinearly related to the sample variance. Hence, it becomes necessary to compute a separate quantization scale for each transform sample.

There are two other scaling laws--the Gaussian error function and the logarithmic--that have the same general characteristics as the optimum mean square error quantizer, but for which the quantization levels are linearly related to the sample variance. In the Gaussian error function quantizer the quantization levels are selected so that when the probability density of transform samples is Gaussian with variance $\sigma^2(u, v)$, the probability that a transform sample is quantized to a given reconstruction level is the same for all levels. This results in a uniform entropy for all reconstruction levels, and therefore a constant word length code may be used for each quantized sample. The quantization levels are given by the solution of the equation

$$\frac{1}{K} = \frac{\int_{Q_{j-1}}^{Q_j} \frac{1}{\sqrt{2\pi\,\sigma^2(u,v)}} \exp\left\{-\frac{F_c^2(u,v)}{2\sigma^2(u,v)}\right\} dF_c}{\int_0^{\frac{NA}{2}} \frac{1}{\sqrt{2\pi\,\sigma^2(u,v)}} \exp\left\{-\frac{F_c^2(u,v)}{2\sigma^2(u,v)}\right\} dF_c} \quad \text{for } j=1,2,\cdots,K-1$$

(5-17)

For $\frac{NA}{2}$ large the denominator approaches one-half and the scaling law can be expressed in terms of the Gaussian error function as

$$\frac{1}{2K} = 1 + \text{erf}\left\{\frac{Q_j}{\sqrt{2}\,\sigma(u,v)}\right\} - \text{erf}\left\{\frac{Q_{j-1}}{\sqrt{2}\,\sigma(u,v)}\right\} \tag{5-18}$$

where

$$\text{erf}\{x\} \equiv \frac{2}{\sqrt{\pi}} \int_0^x \exp\{-Z^2\}\, dz$$

The logarithmic quantizer obeys the function

$$\ell n\left\{1 + \frac{Q_j}{W(u,v)}\right\} = \frac{j}{K}\,\ell n\left\{1 + \frac{NA}{2}\right\} \quad \text{for } j = 0, 1, 2, \cdots, K-1$$

(5-19)

in the positive quadrant and the inverted and reversed function in the negative quadrant where W(u, v) is a spatial frequency weighting function. The quantization levels are approximately

$$Q_j = W(u,v)\left[\left(\frac{NA}{2}\right)^{\frac{j}{K}} - 1\right] \qquad \text{for } j = 0, 1, 2, \cdots, K-1 \qquad (5\text{-}20a)$$

$$Q_K = \frac{NA}{2} \qquad (5\text{-}20b)$$

A convenient implementation of the logarithmic quantizer can be realized by adding plus one to each sample component and then taking the logarithm. The resulting continuous function can then be quantized linearly.

Figure 5-2 shows the relationship between the quantization levels set by the optimum, Gaussian error function, and logarithmic quantizers when the probability density of the transform domain samples is Gaussian with variance $\sigma^2(u,v)$. This figure indicates that Gaussian error function scale is a reasonably good approximation to the optimum scale for a transform sample maximum standard deviation in the range of about 1,500 to 4,000.

5.2 Quantization Experiments

A series of experiments has been conducted to assess the effects of quantization of Fourier and Hadamard transform domain samples. In these experiments the transform domains were quantized and reconstructions were obtained of the quantized samples. The cumulative root mean square quantization error

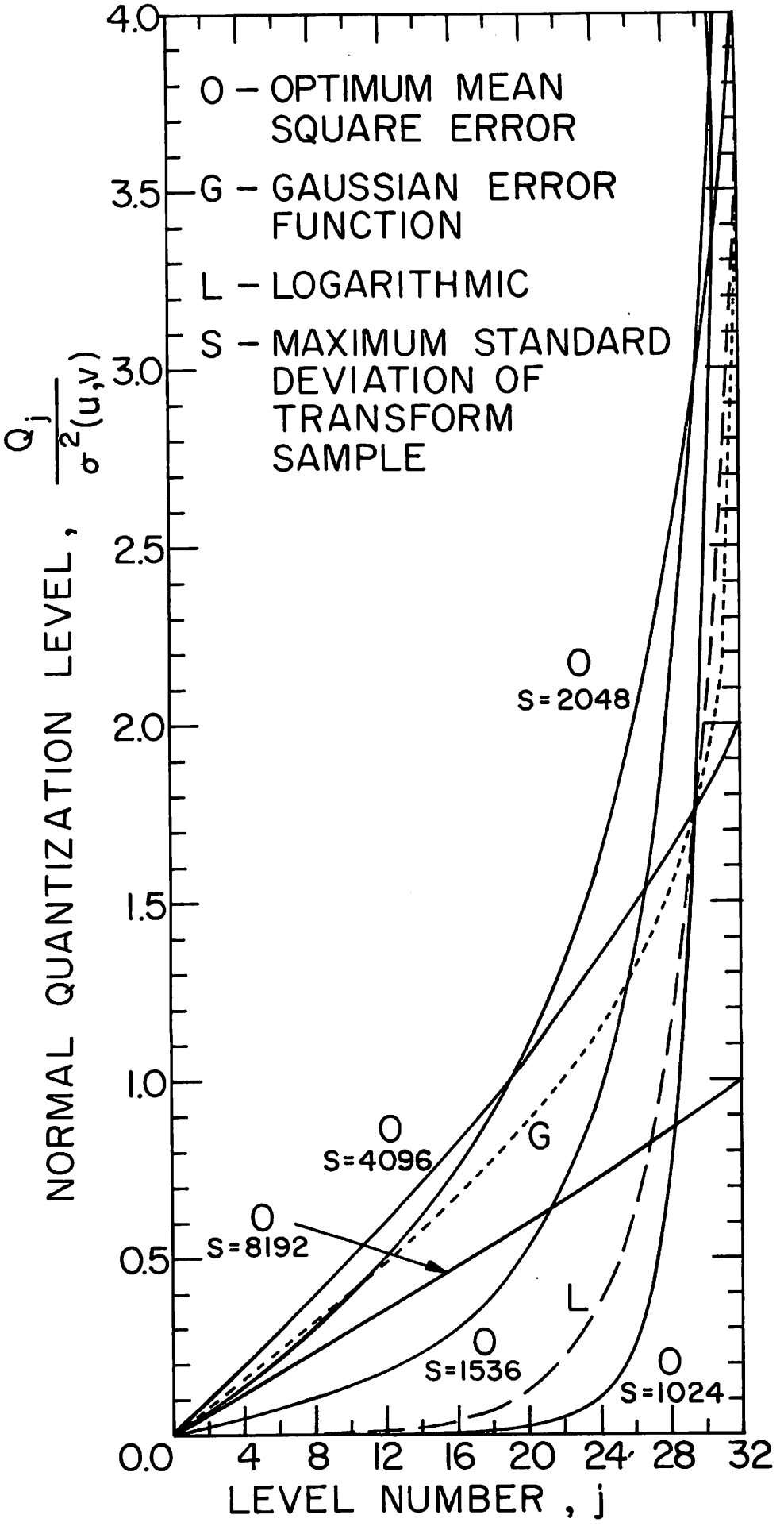


Figure 5-2 Comparison of Optimum and Gaussian Error Function Quantization Scales

$$\delta_Q = \left[\frac{1}{N^2} \sum_x \sum_y [f(x,y) - \hat{f}(x,y)]^2 \right]^{\frac{1}{2}} \quad (5\text{-}21)$$

was measured for each quantized image. In addition the difference function

$$d(x,y) = |f(x,y) - \hat{f}(x,y)| \quad (5\text{-}22)$$

was formed to indicate the spatial correlation of errors.

Figures 5-3 and 5-6 show the effects of quantization on the Fourier and Hadamard transforms, respectively, for a Gaussian error function quantizer with 64 quantization levels. In these experiments the transform domain variance function was modeled as

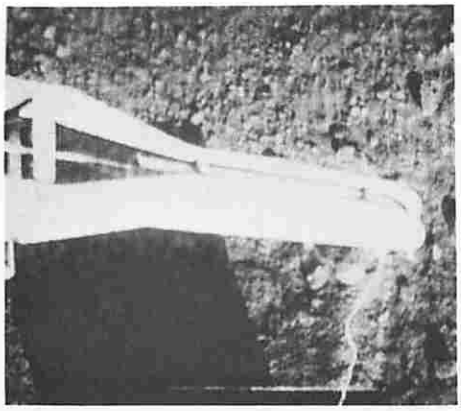
$$\sigma^2(u,v) = S^2 \exp\left\{ - \frac{u^2 + v^2}{p/2} \right\}$$

where S and p are the maximum and spread variance parameters. A computer search procedure was developed to determine the best values of S and p to minimize the quantization error δ_Q. The reconstructions in Figures 5-3 and 5-6 were made on images quantized with the values of S and p giving a minimum value of δ_Q. Figures 5-4, 5-5, 5-7, and 5-8 illustrate the effect of an incorrect choice of the variance parameters S and p. There is a broad range in the values of S and p which provide good quality

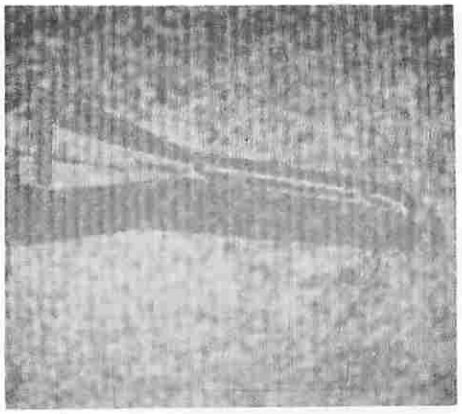
a. amp. par. = 91,000
spread par. = 4,000

b. difference
R.M.S. error = 2.2

Surveyor box



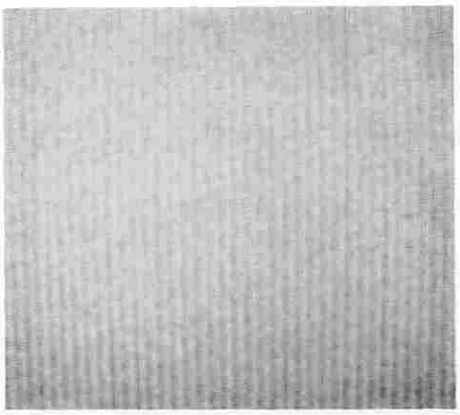
c. amp. par. = 174,000
spread par. = 1,500



d. difference
R.M.S. error = 2.9

Surveyor boom

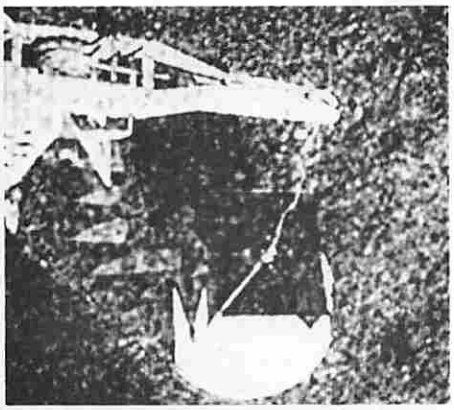
e. amp. par. = 43,000
spread par. = 4,300



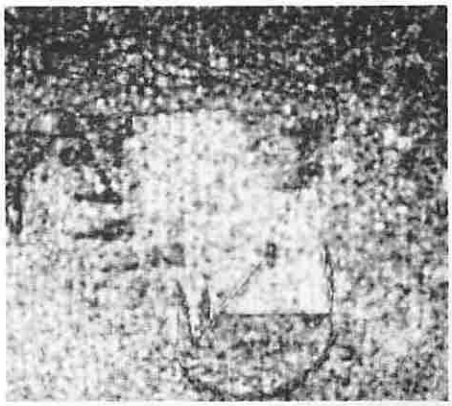
f. difference
R.M.S. error = 1.1

Surveyor footpad

Figure 5-3 Fourier transform quantization: examples of correct parameter scaling -- Gaussian error function quantizer, 64 levels

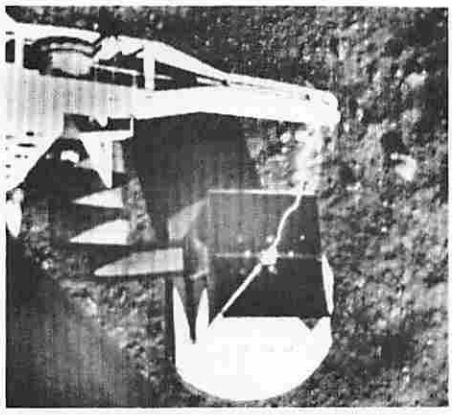


a. amp. par. = 350,000
spread par. = 4,000

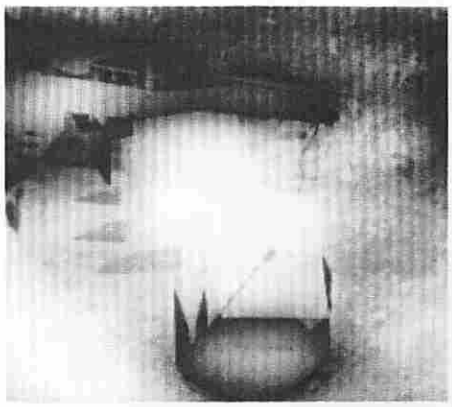


b. difference
R.M.S. error = 11.9

too large amplitude parameter



c. amp. par. = 20,000
spread par. = 4,000

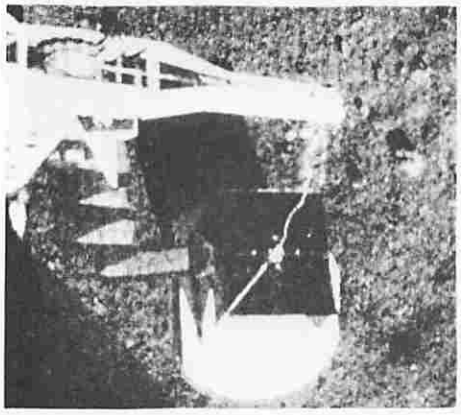


d. difference
R.M.S. error = 14.5

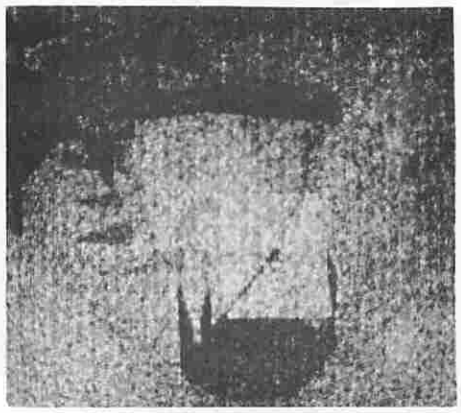
too small amplitude parameter

Proper spread parameter

Figure 5-4 Fourier transform quantization: examples of incorrect amplitude parameter scaling -- Gaussian error function quantizer, 64 levels



a. amp. par. = 91000
spread par. = 16000



b. difference
R.M.S. error = 8.4

too large spread parameter

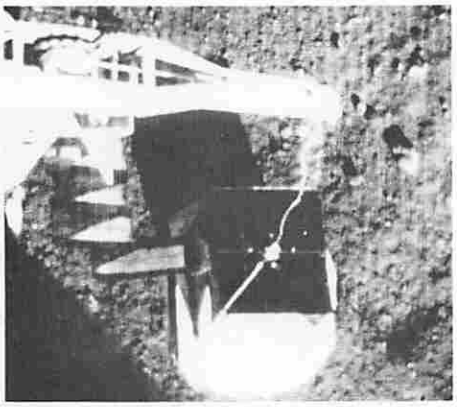
c. amp. par. = 91000
spread par. = 1000

d. difference
R.M.S. error = 5.3

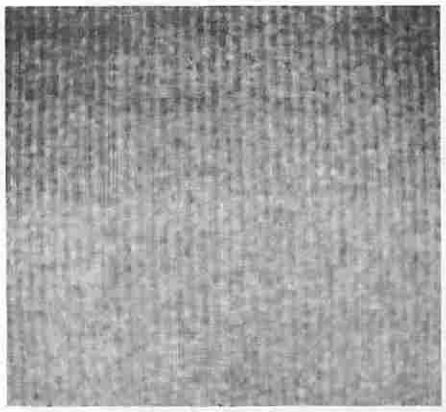
too small spread parameter

Proper amplitude parameter

Figure 5-5 Fourier transform quantization: examples of incorrect spread parameter -- Gaussian error function quantizer, 64 levels

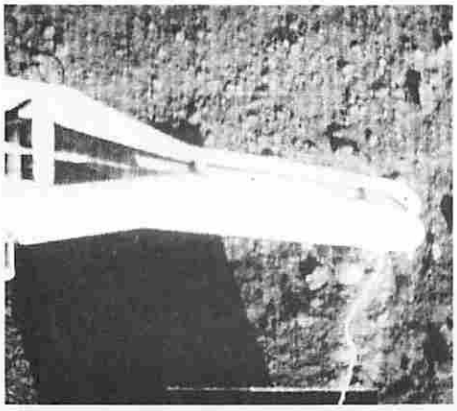


a. amp. par. = 475
spread par. = 11,100

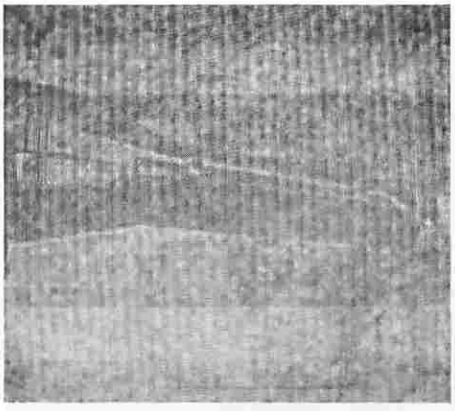


b. difference
R.M.S. error = 1.9

Surveyor box



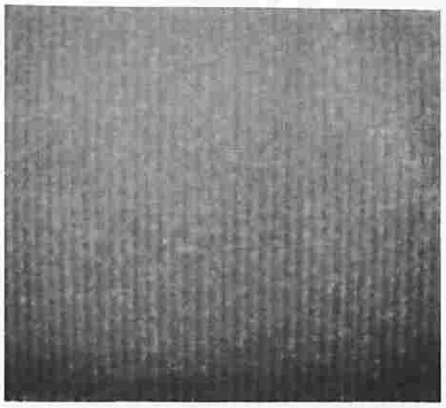
c. amp. par. = 575
spread par. = 10,000



d. difference
R.M.S. error = 2.3

Surveyor boom

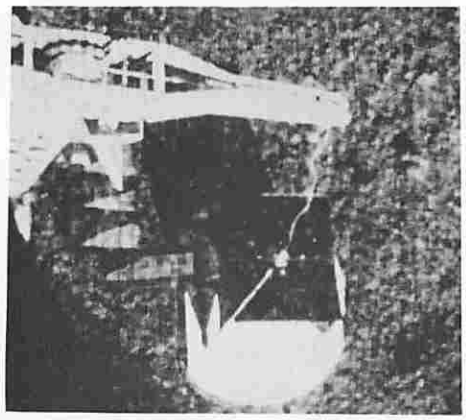
e. amp. par. = 400
spread par. = 10,000



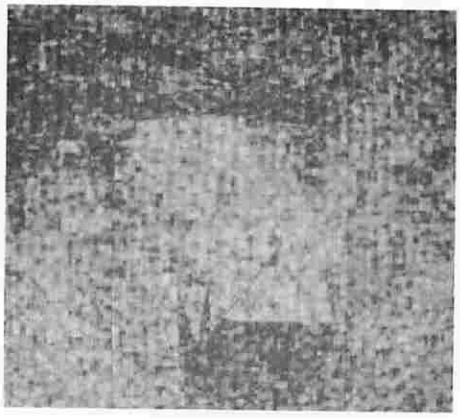
f. difference
R.M.S. error = 1.3

Surveyor footpad

Figure 5-6 Hadamard transform quantization: examples of correct parameter scaling -- Gaussian error function

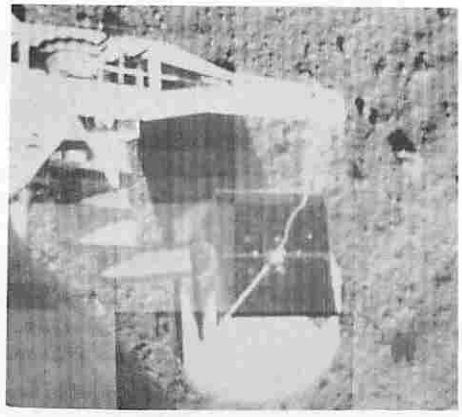


a. amp. par. = 1400
spread par. = 11,100

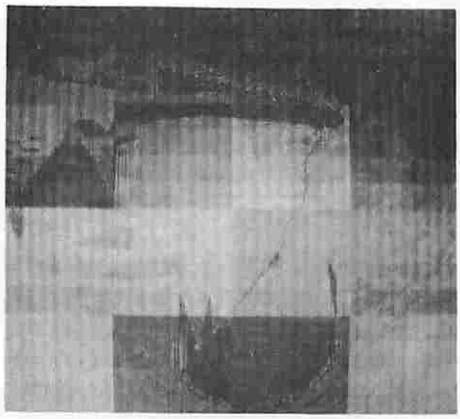


b. difference
R.M.S. error = 7.8

too large amplitude parameter



c. amp. par. = 200
spread par. = 11,100

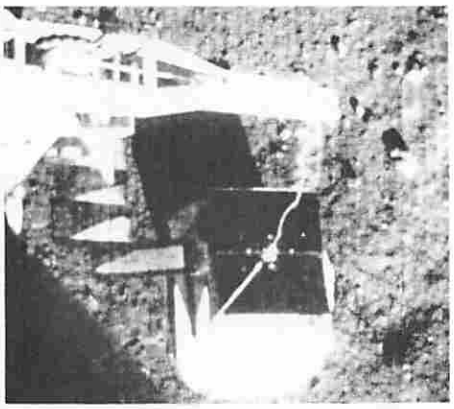


d. difference
R.M.S. error = 8.2

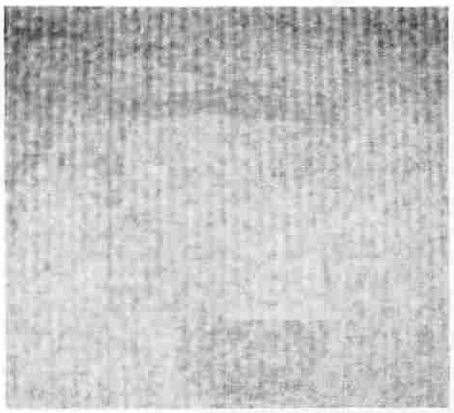
too small amplitude parameter

Proper spread amplitude

Figure 5-7 Hadamard transform quantization: examples of incorrect amplitude parameter scaling - Gaussian error function quantizer, 64 levels

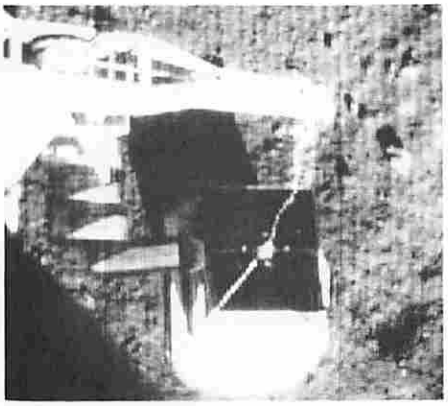


a. amp. par. = 475
spread par. = 15,000

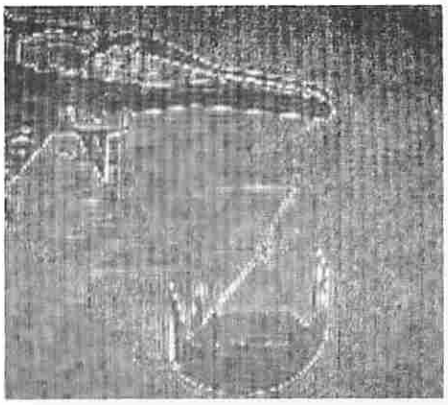


b. difference
R.M.S. error = 5.2

too large spread parameter



c. amp. par. = 475
spread par. = 4000



d. difference
R.M.S. error = 4.2

too small spread parameter

Proper amplitude parameter

Figure 5-8 Hadamard transform quantization: examples of incorrect spread parameter -- Gaussian error function quantizer, 64 levels

reconstructions. Experiments previously reported [27] show the effect of using 32 and 16 quantization levels picked according to the Gaussian error function quantization scale. The quantization error is noticeable for 32 levels and quite bad for 16 levels.

In summary of the quantization experiments, it has been found that good quality Fourier and Hadamard transform reconstructions are possible when the transform samples have been quantized to as few as 64 levels using the Gaussian error function quantization scale.

6. Fourier and Hadamard Image Transform Bandwidth Reduction

A sample reduction, and a subsequent bandwidth reduction by proper coding, are possible by coding the Fourier or Hadamard transform of an image rather than the image itself. This sample reduction is obtainable because, as a result of the element-to-element correlation in the image, many of the transform domain samples are of extremely low magnitude and may be deleted from the image reconstruction without seriously degrading the quality of the reconstructed image.

The process of selecting samples for inclusion in the image reconstruction can be conveniently analyzed from the viewpoint of two dimensional sampling. Figure 6-1 illustrates a generalized block diagram of a transform sampling system. The forward transform of an image, $F(u,v)$, is multiplied by a two dimensional sampling function, $S(u,v)$, which takes on the values zero or one according to some a priori or adaptive rule. The sampled image transform, $F_s(u,v)$, is

$$F_s(u,v) = F(u,v)\ S(u,v) \tag{6-1}$$

The reconstructed image, $f_s(x,y)$, is then the reverse transform of $F_s(u,v)$, Thus,

$$f_s(x,y) = \sum_{n=0}^{N-1} \sum_{n=0}^{N-1} F(u,v)\ S(u,v)\ b(x,y,u,v) \tag{6-2}$$

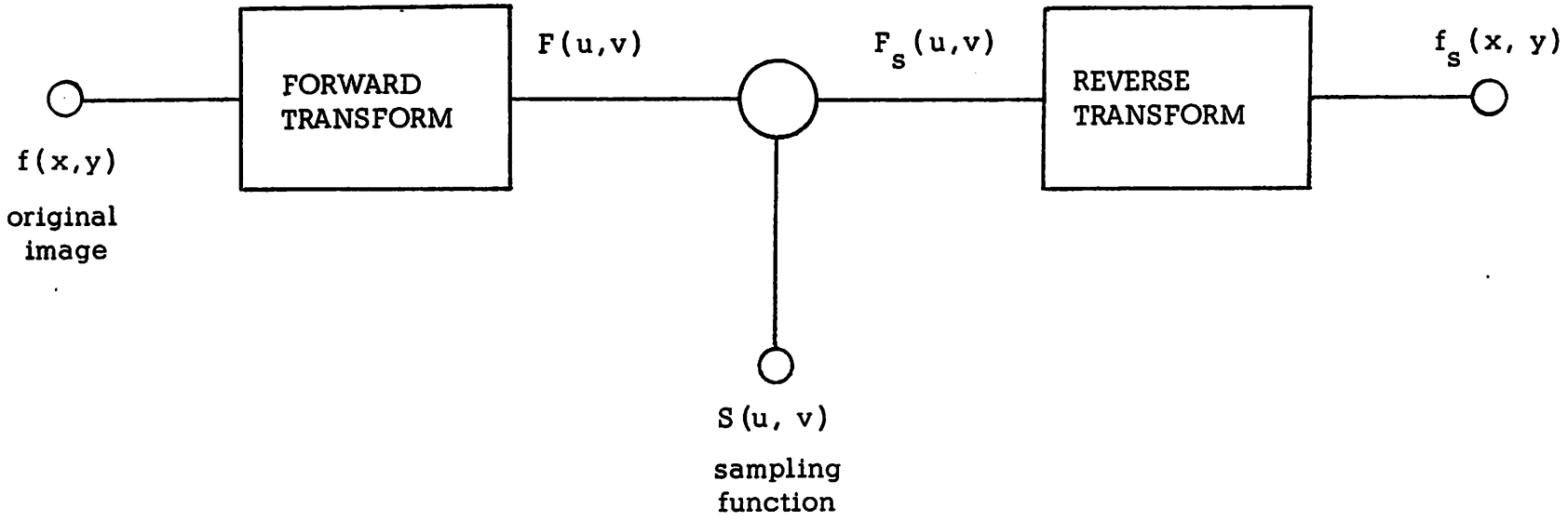


Figure 6-1 Transform Domain Sampling

In the case of the Fourier transform, as a consequence of the frequency translation theorem, the reconstructed image can be expressed as a spatial convolution, denoted by the symbol $\circledast$, of the original image and the inverse Fourier transform, $s(x, y)$, of $S(u, v)$. Thus,

$$f(x, y) = f_s(x, y) \circledast s(x, y) \tag{6-3}$$

It should be noted that this result does not hold for the Hadamard transform since the Hadamard transform does not possess a sequency translation property.

Table 6-1 lists three basic transform sampling methods. With the random sampling method the sampling function, $S(u, v)$, assumes the value 0 or 1 according to some probability distribution $p(u, v)$ over the transform domain. Experiments have been performed in which one-half of the Fourier transform samples have been randomly discarded independent of their location in the transform domain. The resultant reconstructions were of poor quality due to errors in deleting large magnitude low frequency samples. Several variations were attempted in which more of the lower spatial frequencies were included, but the results were not particularly encouraging. It appears that at most a 2:1 sample reduction can be obtained by random sampling at the cost of moderate degradation.

TABLE 6-1

Classification of Transform Sampling Methods

Description	Sampling Function S(u,v)	Conditions
Random sampling	1	with probability $p(u,v)$
	0	with probability $1 - p(u,v)$
Zonal sampling	1	u,v in sampling region
	0	u,v not in sampling region
Threshold sampling	1	$\lvert F(u,v) \rvert > M_T(u,v)$
	0	$\lvert F(u,v) \rvert \leq M_T(u,v)$

The zonal and threshold sampling techniques are discussed in the following sections.

6.1 Zonal Transform Sampling

In most scenes of interest the energy in the Fourier transform domain tends to be clustered toward the lowest spatial frequencies. Similarly, the Hadamard transform domain energy is greatest at the low sequencies. For example, in the three Surveyor spacecraft scenes, 95% of the image energy in the Fourier transform is contained in 1% or less of the transform samples [27].

With an image energy distribution clustered at the low spatial frequencies or sequencies, the most obvious means of conserving bandwidth is simply to not transmit the high spatial frequency or sequency samples. Discarding the high spatial frequencies or sequencies is equivalent to passing the image through a circular, zonal, low pass filter; the result is a loss of focus. If some degree of resolution loss is acceptable, zonal low pass filtering of the transform domain yields relatively large bandwidth reductions.

Zonal low pass filtering of a sequency ordered Hadamard transform is equivalent to multiplying the transform samples by the sampling function

Hadamard transform:

$$\left.\begin{array}{l} S(u,v) = 0 \quad \text{if} \quad u^2 + v^2 > R^2 \\ S(u,v) = 1 \quad \text{otherwise} \end{array}\right\} \quad u,v = 0,1,\cdots,N-1 \qquad (6\text{-}4)$$

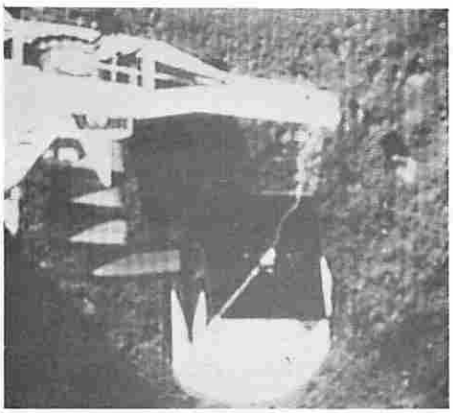
For the Fourier transform the sampling function is

Fourier transform:

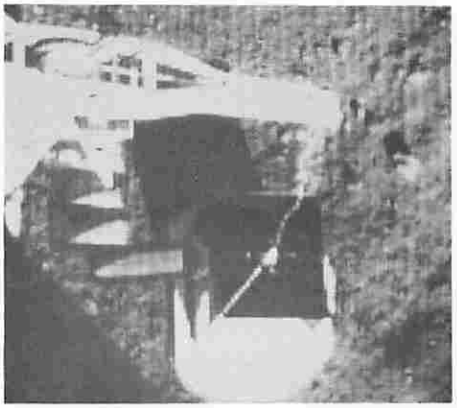
$$\left.\begin{array}{l} S(u,v) = 1 \\ S(N-1-u,\ v) = 1 \\ S(u,\ N-1-v) = 1 \\ S(N-1-u,\ N-1-v) = 1 \\ S(u,v) = 0 \qquad \text{otherwise} \end{array}\right\} \quad \text{if } u^2 + v^2 > R^2 \quad \left.\begin{array}{l} \\ \\ \\ \\ \\ \end{array}\right\} u,v = 0,1,\cdots,\frac{N}{2}-1$$

(6-5)

Figure 6-2 shows the effects of Fourier and Hadamard transform zonal low pass sampling of the Surveyor box scene over the full frame of 256 by 256 elements. These experiments support the widely known fact that the high frequency and sequency brightness transitions are important even though they are relatively few in number and contain a low proportion of the image energy. The image degradation tends to be more noticeable for zonal filtering of the Hadamard transform than the Fourier transform for the same sample reduction factor because of the rectangular shape of

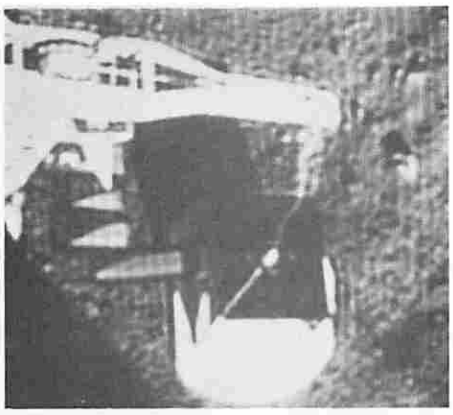


a. 4:1 Fourier

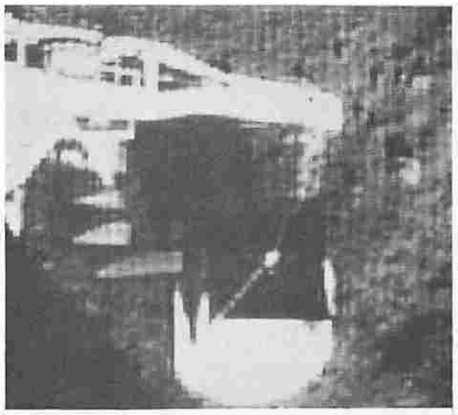


b. 4:1 Hadamard

R = 143

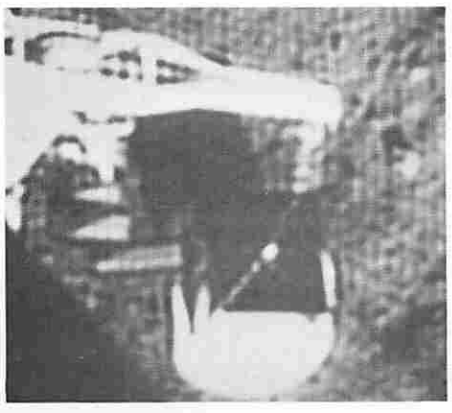


c. 8:1 Fourier

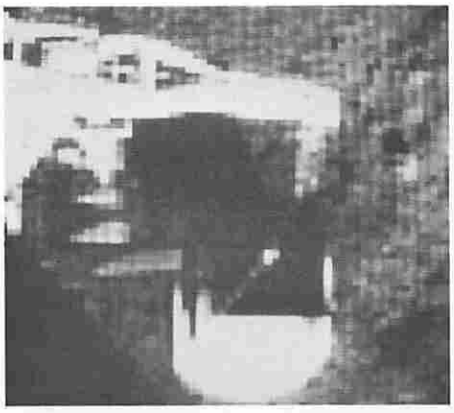


d. 8:1 Hadamard

R = 101



e. 16:1 Fourier



f. 16:1 Hadamard

R = 71

Figure 6-2. Circular Low Pass Zonal Fourier and Hadamard Transform Sampling of Surveyor Box over Full Frame of 256 x 256 Elements, unquantized transform.

the two dimensional Hadamard reconstruction waveforms. The eye is very sensitive to the presence of sharp brightness transitions within an image. With the Hadamard transform all transitions occur within one element, whereas in the Fourier transform the brightness transitions are spread over many elements since the reconstruction waveforms are two dimensional sinusoids.

If the zonal low pass filter has square rather than circular boundaries in the transform domain, it is possible to produce a low pass version of the original by the simple expedient of spatial averaging of elements in the original image. In this case the complexities of the transform operation would probably not be warranted if a low pass reconstruction is acceptable.

It has been conjectured that to produce a subjectively pleasing image, the eye only requires the low spatial frequencies of an image signal to provide the overall grey scale and the high spatial frequencies to provide the edge transitions. The mid-spatial frequencies are assumed to play a minor part in the reconstruction of an image. This conjecture has been tested by sampling the Fourier and Hadamard transforms of an image with a circular zonal rejection filter with the characteristic functions

Fourier transform:

$$\left.\begin{array}{l} S(u,v) = 0 \\ S(N-1-u,v) = 0 \\ S(u,N-1-v) = 0 \\ S(N-1-u,\ N-1-v) = 0 \end{array}\right\} \text{if } R_1^2 < u^2+v^2 < R_2^2 \text{ or } u^2+v^2 > R_3^2 \quad \Big\} \; u,v = 0,1,\cdots,\frac{N}{2}-1$$

$$S(u,v) = 1 \quad \text{otherwise} \tag{6-6}$$

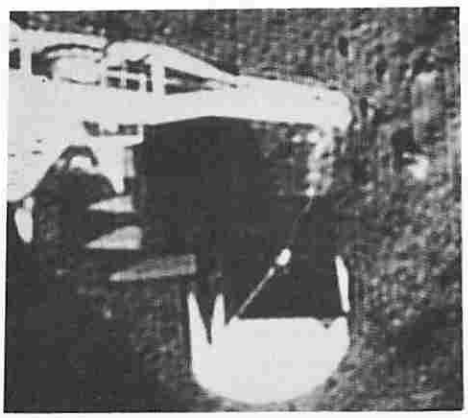
Hadamard transform:

$$S(u,v) = 0 \quad \text{for } R_1^2 < u^2+v^2 < R_2^2 \text{ or } u^2+v^2 > R_3^2$$

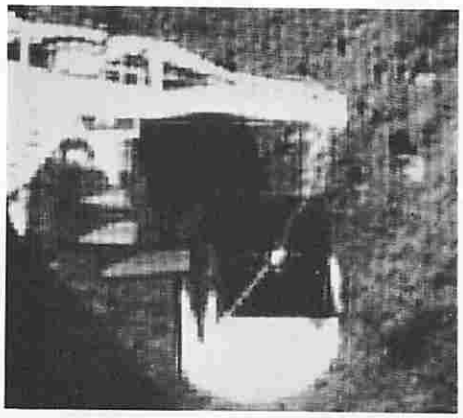
$$S(u,v) = 1 \quad \text{otherwise} \tag{6-7}$$

Figure 6-3 shows the effect of band rejection filtering on the Surveyor box scene. The image quality appears to be somewhat degraded as compared to the results of Figure 6-2 with a simple circular zonal low pass selection of transform samples.

In the development of zonal transform sampling techniques presented in Section 4 for transforms in 16 by 16 element blocks, it was found that a hyperbolic zone suppressed the grid effect for high sample reduction factors better than a circular zone. The effect of the use of a hyperbolic zone of full size, 256 by 256 point, images is shown in Figure 6-4. These reconstructions show a perceptible improvement to their counterparts in Figure 6-2 for a circular zone. For natural images it appears that a hyperbolic zone matches the



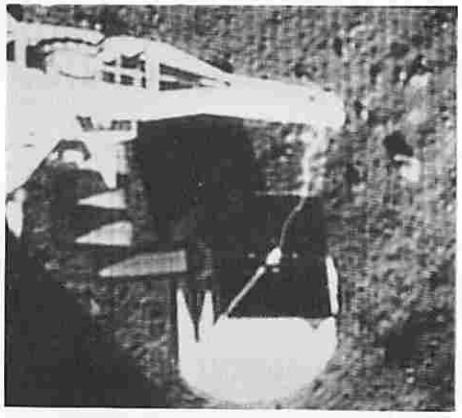
a. Fourier



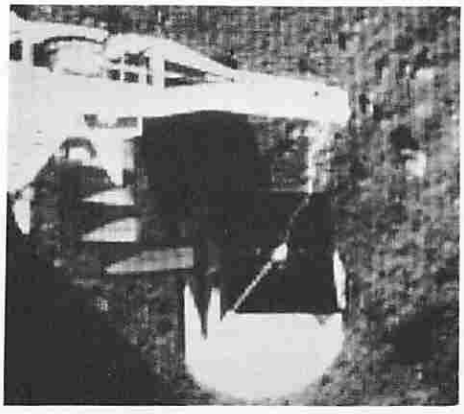
b. Hadamard

$R_1 = 100$; $R_2 = 233.8$

$R_3 = 256$



c. Fourier



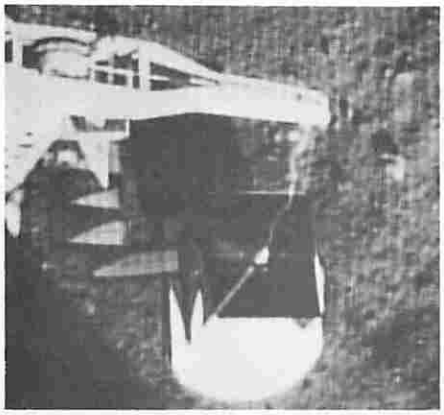
d. Hadamard

$R_1 = 125$; $R_2 = 245.5$

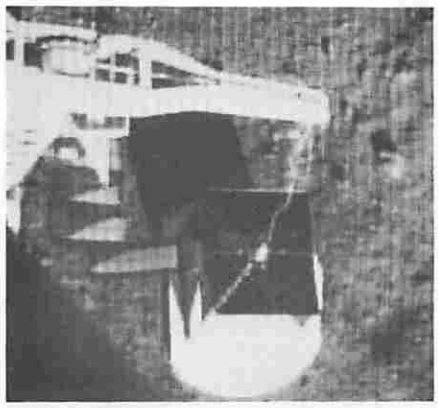
$R_3 = 256$

4:1 sample reduction

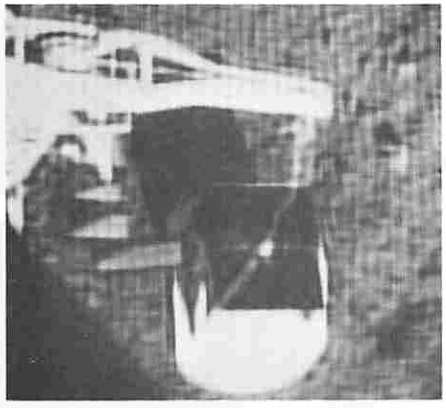
Figure 6-3 Circular Band Rejection Zonal Fourier and Hadamard Transform Sampling of Surveyor Box over Full Frame of 256 x 256 Elements, unquantized transform.



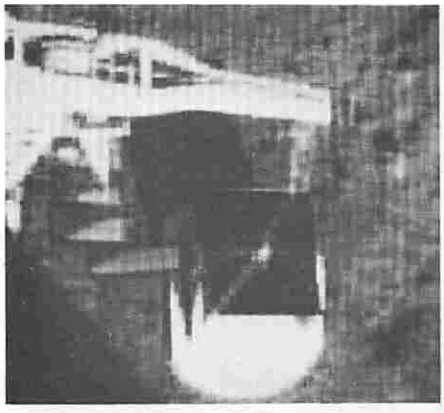
a. 4:1 Fourier



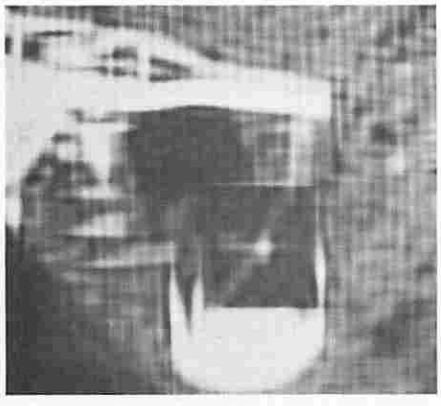
b. 4:1 Hadamard



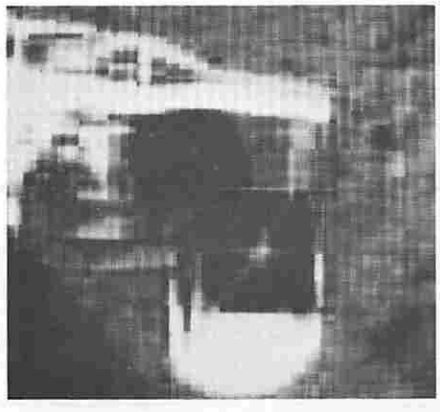
c. 8:1 Fourier



d. 8:1 Hadamard



e. 16:1 Fourier



f. 16:1 Hadamard

Figure 6-4 Hyperbolic Low Pass Zonal Fourier and Hadamard Transform Sampling of Surveyor Box over Full Frame of 256 x 256 Elements, unquantized transform.

energy distribution better in a spatial frequency or sequency ordered transform plane than does a circular zone.

6.2 Threshold Transform Sampling

The difficulty with zonal transform sampling is that large magnitude transform samples often are included in the rejection zone and therefore deleted from the reconstruction. This problem can be overcome by the threshold sampling technique in which samples whose magnitudes are greater than a pre-specified threshold level are included in the image reconstruction independent of their position in the transform domain.

Figure 6-5 and 6-6 are plots of the percentage of transform domain samples lying below a magnitude threshold level for the Fourier and Hadamard transforms. Maps showing the location of transform samples exceeding the threshold level for the Fourier and Hadamard transforms are shown in Figure 6-7. It should be noted that the large magnitude samples tend to be located at the lower spatial frequencies or sequencies. But many high spatial frequency and sequency samples exceed the threshold. In low pass zonal filtering these transform domain samples would not have been included in the image reconstruction.

Figure 6-8 to 6-11 show the effects of threshold coding in the transform domain for the Fourier and Hadamard transforms. Each transform domain has been quantized to 64 levels per transform sample

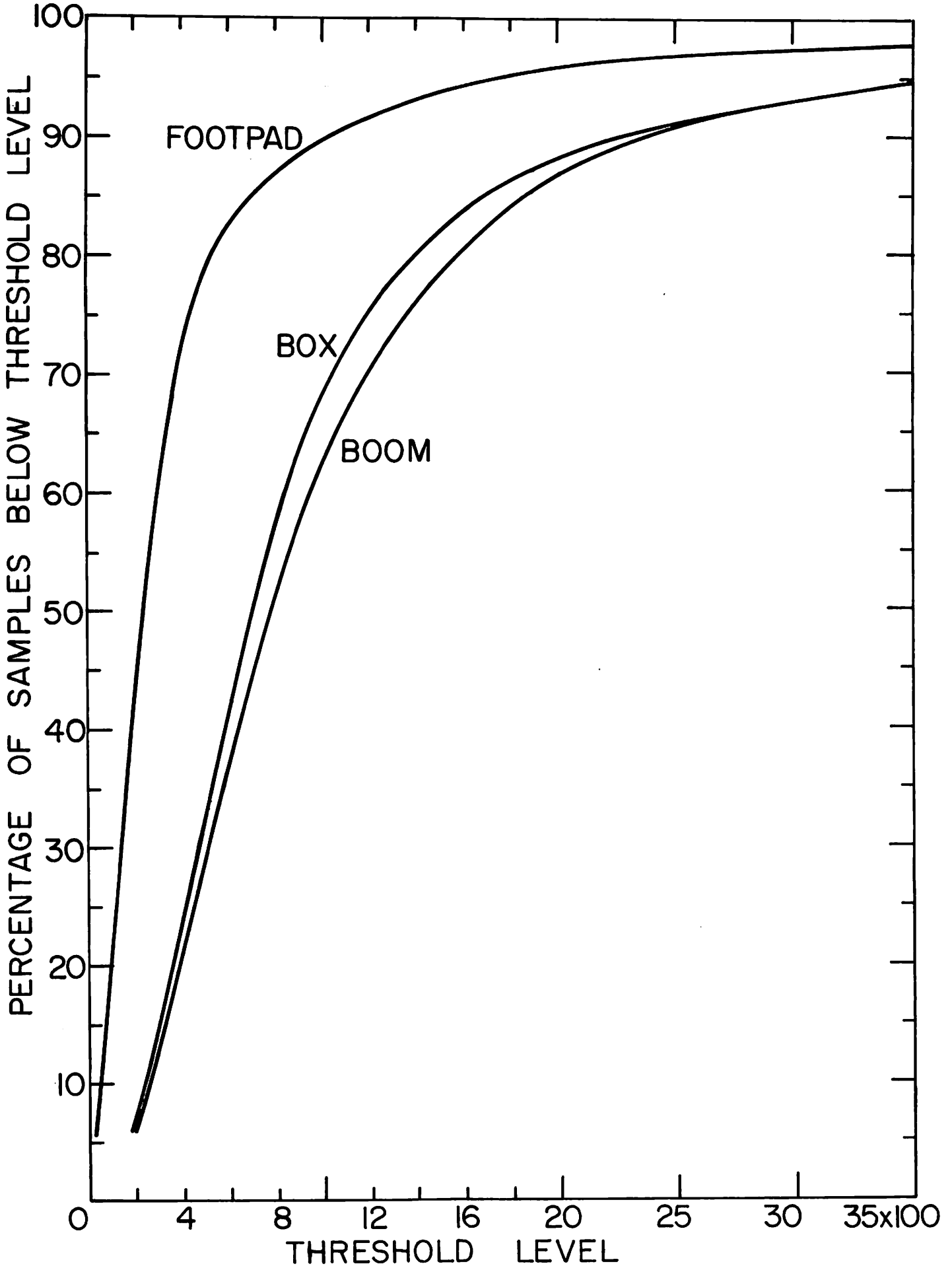


Figure 6-5 Number of Fourier Transform Samples below Threshold versus Threshold Level

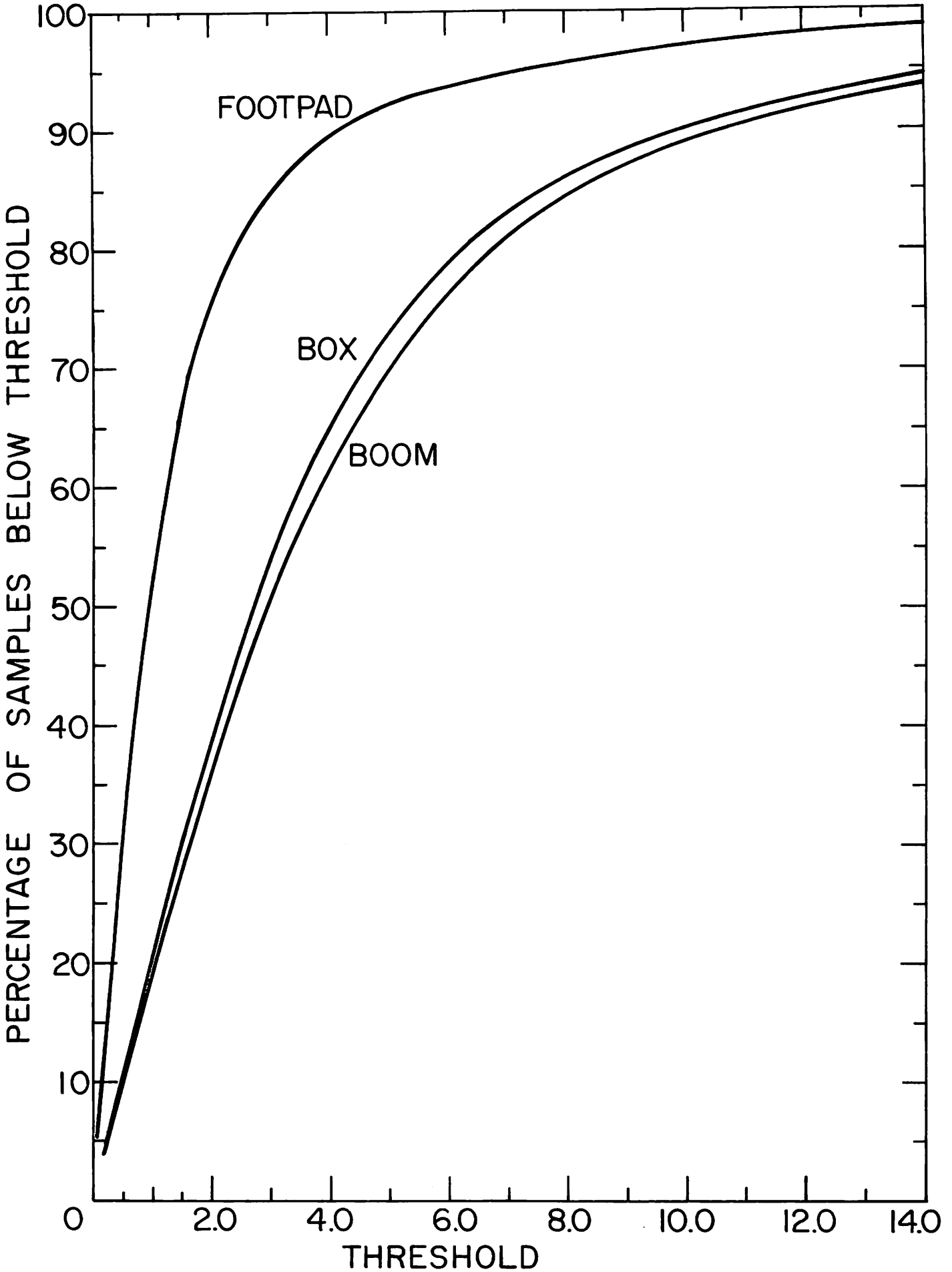
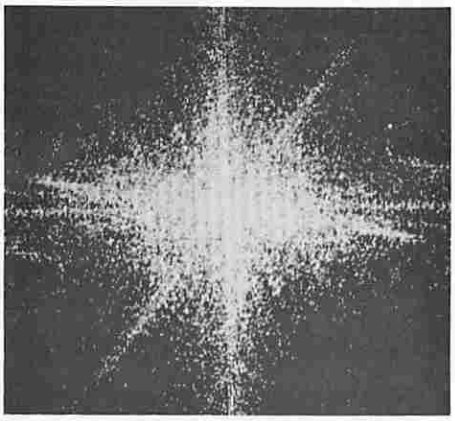
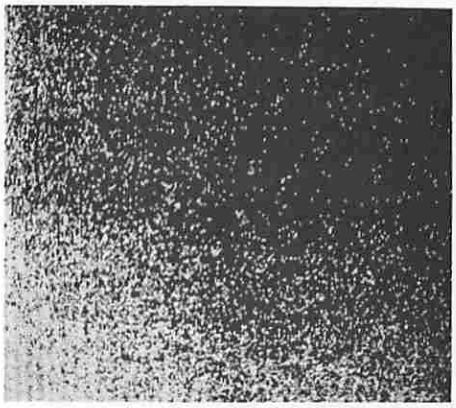


Figure 6-6 Number of Hadamard Transform Samples below Threshold versus Threshold Level

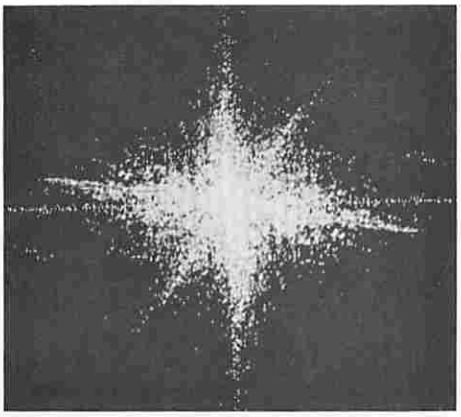


a. Fourier

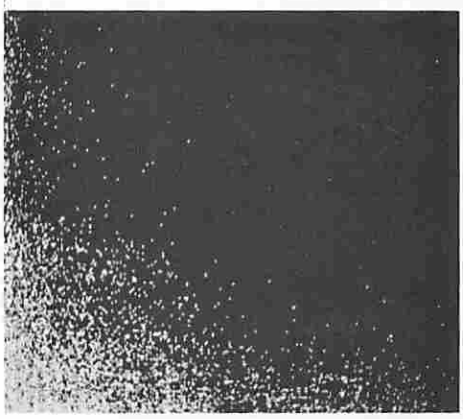


b. Hadamard

5:1 sample reduction

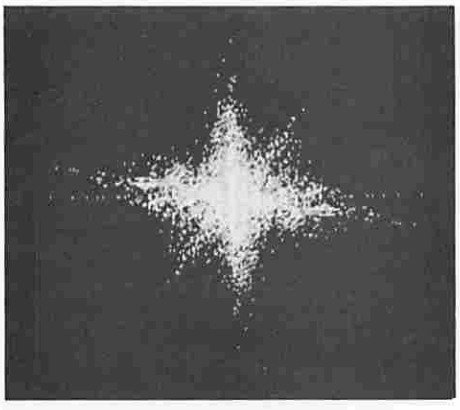


c. Fourier

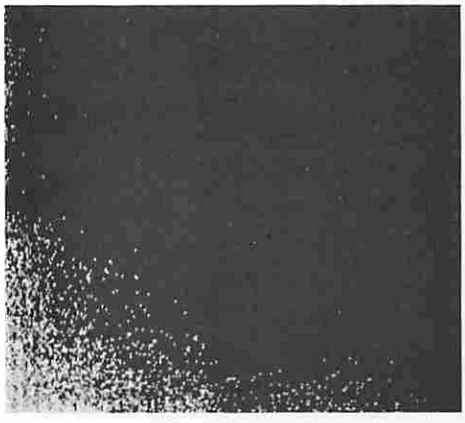


d. Hadamard

10:1 sample reduction



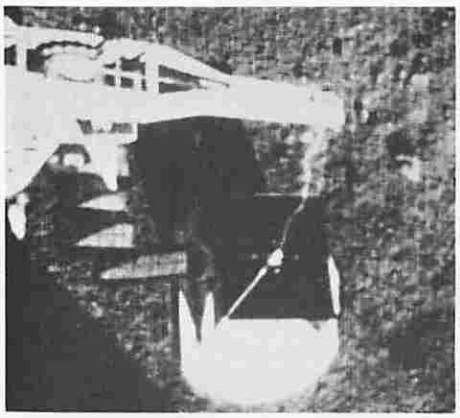
e. Fourier



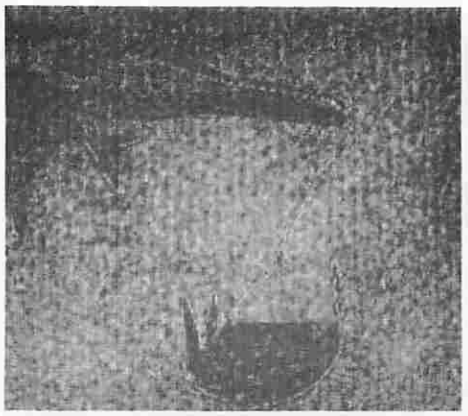
f. Hadamard

20:1 sample reduction

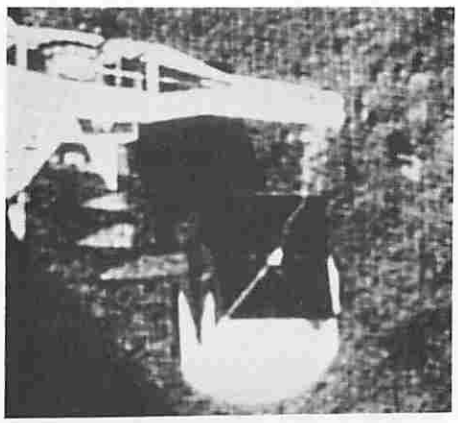
Figure 6-7 Maps of Fourier and Hadamard Transform Samples above Threshold for Surveyor Box over a Full Frame of 256 x 256 Elements.



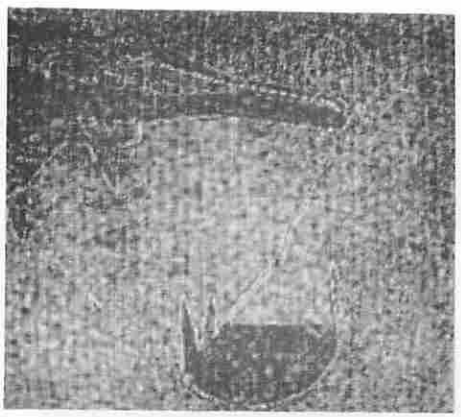
a. 5:1 sample reduction



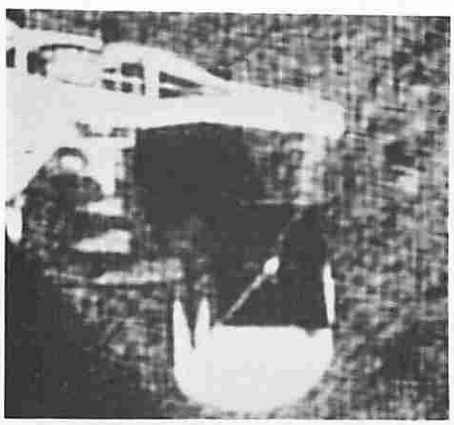
b. Difference
R.M.S. error = 3.6



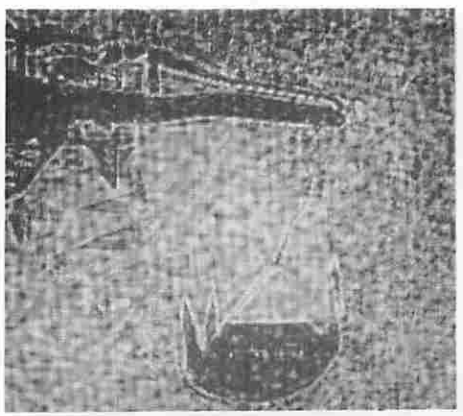
c. 10:1 sample reduction



d. Difference
R.M.S. error = 3.8

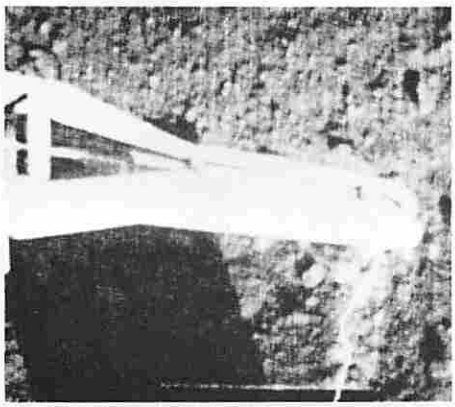


e. 20:1 sample reduction

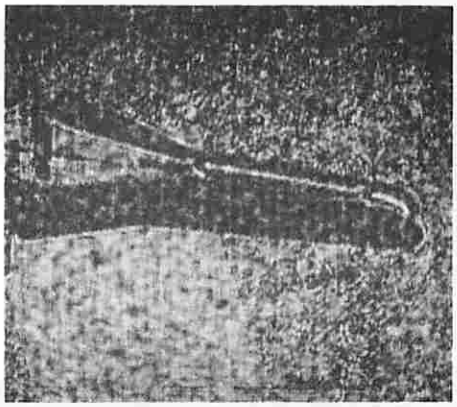


f. Difference
R.M.S. error = 4.7

Figure 6-8 Fourier Transform Threshold Coding: Effects of Thresholding for Surveyor Box over a Full Frame of 256 x 256 Elements, quantized transform.



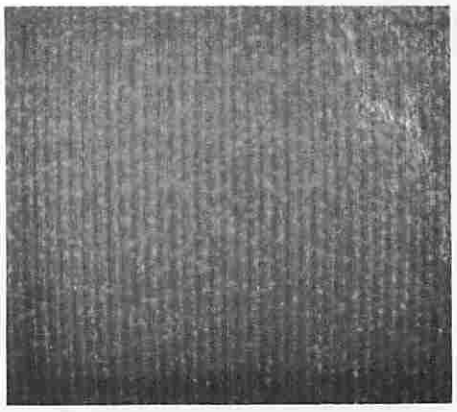
a. 10:1 sample reduction



b. Difference
R.M.S. error = 3.7

Surveyor Boom

c. 10:1 sample reduction

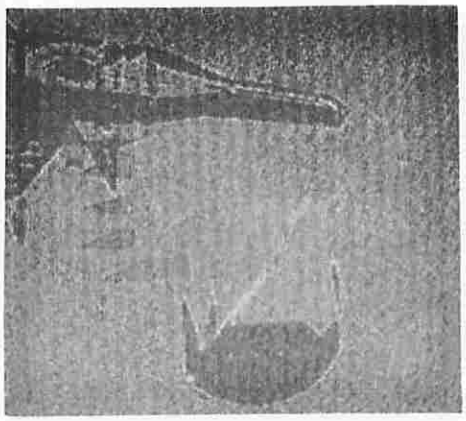


d. Difference
R.M.S. error = 1.6

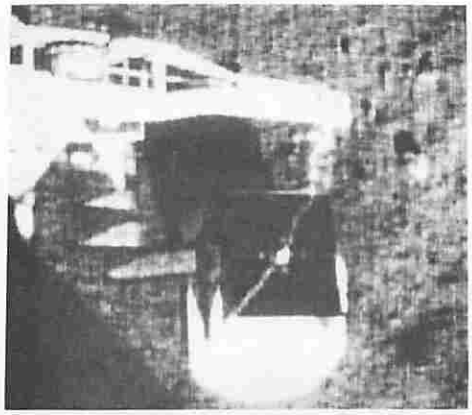
Surveyor Footpad

Figure 6-9 Fourier Transform Threshold Coding: Effects of Thresholding for Surveyor Footpad and Boom over a Full Frame of 256 x 256 Elements, quantized transform.

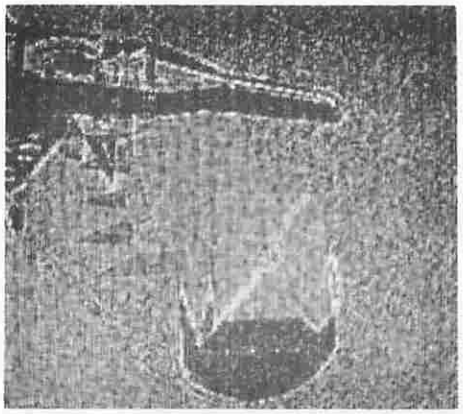
a. 5:1 sample reduction



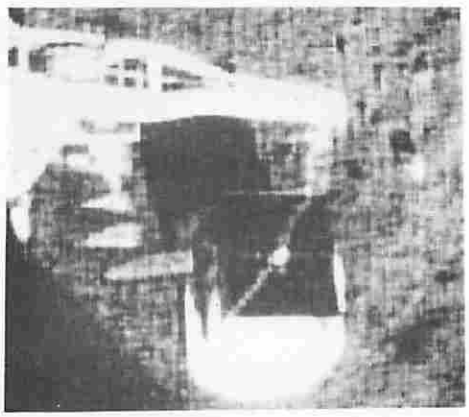
b. Difference
R.M.S. error = 2.8



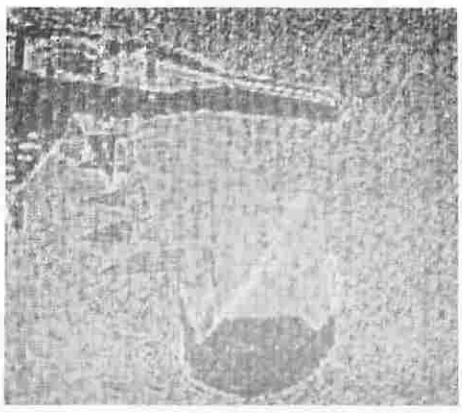
c. 10:1 sample reduction



d. Difference
R.M.S. error = 3.9

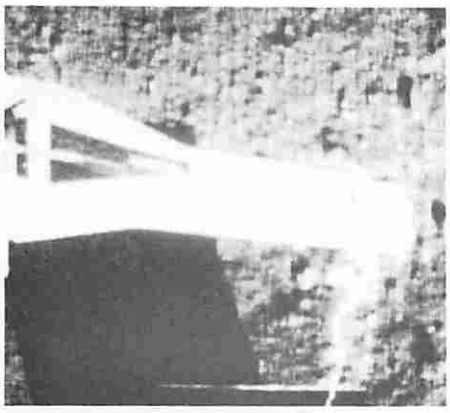


e. 20:1 sample reduction

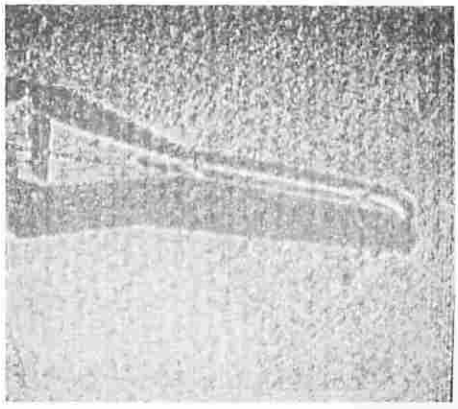


f. Difference
R.M.S. error = 4.5

Figure 6-10 Hadamard Transform Threshold Coding: Effects of Thresholding for Surveyor Box over a Full Frame of 256 x 256 Elements, quantized transform.

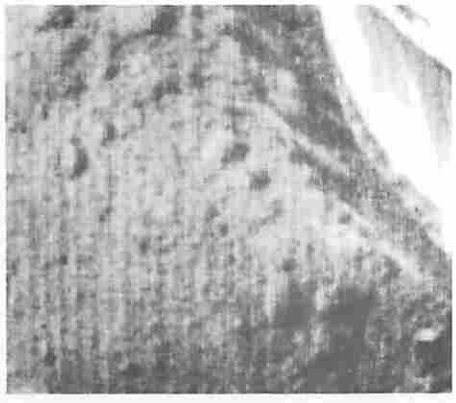


a. 10:1 sample reduction

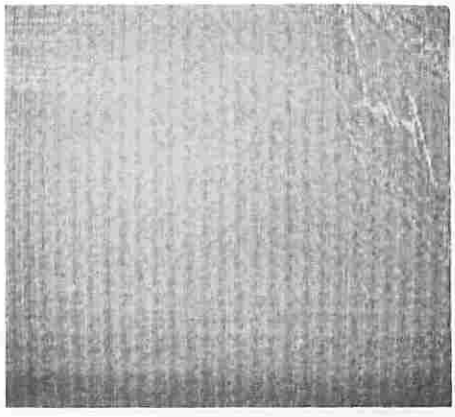


b. Difference
R.M.S. error = 4.8

Surveyor Boom



c. 10:1 sample reduction



d. Difference
R.M.S. error = 1.4

Surveyor Footpad

Figure 6-11 Hadamard Transform Threshold Coding: Effects of Thresholding for Surveyor Footpad and Boom over a Full Frame of 256 x 256 Elements, quantized transform.

component according to the Gaussian error function quantizer scale. Thus, the image reconstructions exhibit the joint effects of sample deletion and quantization. Difference pictures are displayed to illustrate the spatial distribution of errors. Also the cumulative average mean square error has been measured for each reconstruction. From these experiments it can be concluded that the Fourier and Hadamard transforms both provide good quality reconstructions for sample reduction factors of 5:1. Some image degradation is noticeable for a sample reduction factor of 10:1.

In order to achieve a bandwidth reduction for digital image transmission with transform domain threshold coding it is necessary to code the position of the samples exceeding the threshold level. There are a variety of ways of position coding that could be employed. The simplest conceptually would be to code the coordinates of each significant transform sample. Higher coding efficiency can be obtained, however, by coding the number of non-significant samples between significant samples. This scheme, called run length coding, has been used quite successfully in the spatial domain for black or white pictures. To achieve a short position code length, runs are usually restricted in length to some maximum value, normally a power of two. By including a line synchronization code group it becomes unnecessary to code the line number. Another advantage of the employment of a line synchronization code is that it prevents the propagation of channel errors over more than one line.

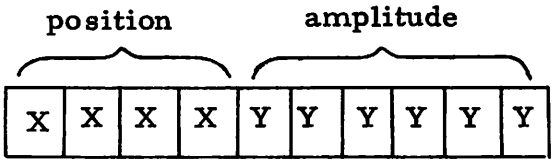
A run length coding procedure for Fourier and Hadamard transform threshold coding has been implemented on a general purpose digital computer. The coding procedure is "fail safe" in that every transform domain sample is coded for a zero level threshold; there is no truncation of transform samples. The basic properties of the run length coding procedure are outlined below:

a. The first sample along each line is coded regardless of its magnitude. A position code of all zero bits is affixed to the amplitude code to compromise the line synchronization code group.

b. 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.

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

A simple code to implement this run length coding procedure is given below.



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	Y	Y	first sample of a line/above threshold
0	0	0	1	Y	Y	Y	Y	Y	Y	run length = 1 (adjacent significant samples)
0	0	1	0	Y	Y	Y	Y	Y	Y	run length = 2
		⋮								
1	1	1	0	Y	Y	Y	Y	Y	Y	run length = 14
1	1	1	1	1	1	1	1	1	1	pseudo-run of length 14

This run length coding procedure for transform threshold coding has been tested for the Surveyor box and boom scenes. As expected, the run length coding does not introduce any reconstruction errors. The effect of channel errors on position bits is considered in the next section. Table 6-2 shows the bandwidth reduction factors obtained for these test scenes as a function of the sample reduction factor. In all cases the run length code employed four position bits and runs were truncated in length to 14 samples. Better performance could, no doubt, be obtained if the number of position code bits were tailored to match the run statistics.

TABLE 6-2

Experimental Bandwidth Reduction Achieved by Fourier and Hadamard Transform Threshold Coding for Surveyor Box Scene

Sample Reduction	Fourier Transform				Hadamard Transform			
	Number of position bits				Number of position bits			
	3	4	5	6	3	4	5	6
5 : 1	2.6	3.1	3.2	3.2	2.2	2.6	2.6	2.5
10 : 1	3.3	4.8	5.7	6.0	3.0	4.3	4.9	5.1
20 : 1	3.9	6.6	8.7	10.0	3.3	5.4	6.9	7.7

7. Fourier and Hadamard Image Transform Channel Error Tolerance

A major concern of communication system designers is the susceptibility of data to noise interference. It is important, then, to study the effects of noise on the image transform coding communication system. The inherent "error averaging" property of transform coding combined with error correction coding of specific transform samples provides a means of image coding for which channel errors are less deleterious than for conventional spatial coding of an image. This property, of course, is predicated on the assumption that the particular transform used tends to compact image energy in a few number of coefficients in the transform domain.

In most digital communication systems the code alphabet consists of two symbols which are subject to perturbations in the channel, and these perturbations introduce random noise at the receiver. The binary symmetric channel is used as the noise model in the study of channel effects on image transform coding. The classical representation of such a communication channel is given in Figure 7-1, where the probability of receiving an incorrect symbol is p regardless of which symbol is transmitted.

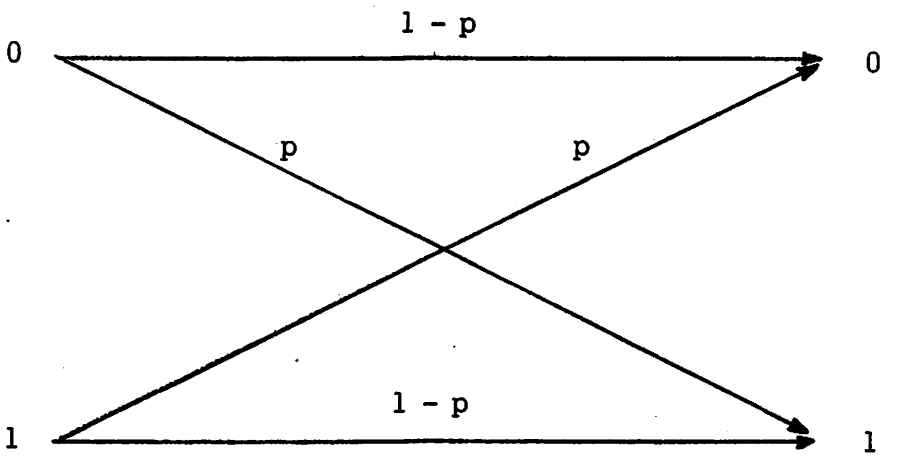
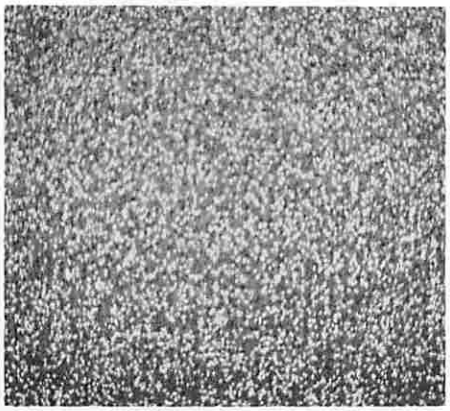


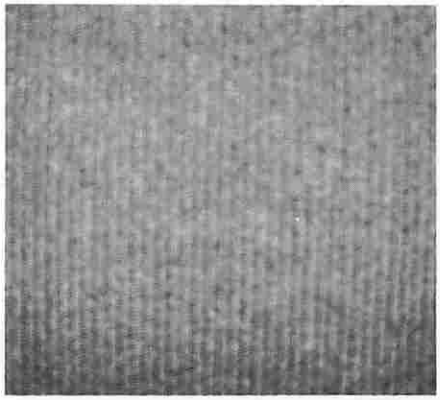
Figure 7-1 Model of a Binary Symmetric Channel

7.1 Channel Noise Effects

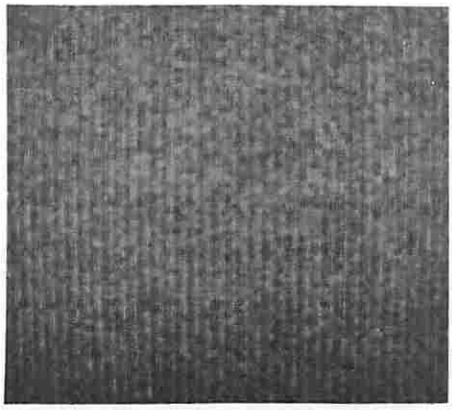
An intuitive justification for transmitting the transform rather than the spatial domain of an image is the fact that channel noise introduced in the transform of an image tends to be distributed over the entire reconstructed image. Consequently, the noise manifests itself as a combination of low order orthogonal functions in the image due to noise introduced in the large amplitude coefficients of the transform domain. If the Fourier transform is used, the noise presents itself as a low frequency effect and if the Hadamard transform is used, the effect is low sequency corresponding to non-periodic checkerboards of a low number of zero crossings. Finally, if the Karhunen-Loeve transform could be used, the noise introduced in the large valued coefficients would correspond to those orthogonal functions representing the largest eigenvalues and matching the original image closest in a mean square error sense. In all cases, since the eye is more sensitive to the high frequency "salt and pepper" effect of channel noise in the spatial domain, the same channel error rate in the transform domain is somewhat less offensive. Figure 7-2a shows a mid-grey scene after having passed through a channel with probability of error of $Pe = 10^{-1}$. Figure 7-2b is the Fourier transform of the output of the same channel whose input was the Fourier transform of the mid-grey scene. Figure 7-2c is the same experiment replacing the Fourier transform with the Hadamard transform. All



a. BSC noise in spatial domain



b. Fourier transform of BSC noise in Fourier domain



c. Hadamard transform of BSC noise in Hadamard domain

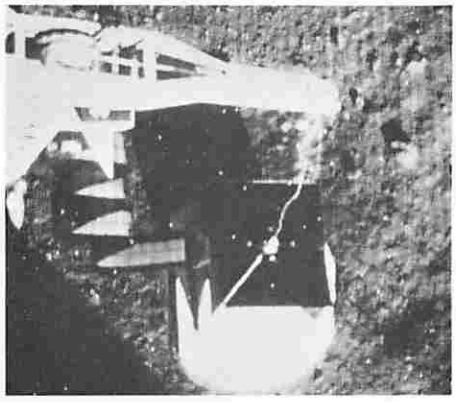
Figure 7-2 Binary Symmetric Channel with Error Rate Pe = 10^{-1}

three scenes have the same error rate but the induced noise energy is distributed quite differently. A quantizing and coding method can be developed to take advantage of the inherent high frequency or "salt and pepper" noise immunity that transform domain coding offers. As a first step in this direction a requirement will be made that each quantum level occur equally likely as any other quantum level. This quantization criterion will guarantee that each code word is equally likely to occur and will avoid any unexpected noise biasing, since the binary symmetric channel affects each code bit, and therefore each code word, independently of all others. Such a quantization requirement results in the quantization rule employed in the earlier sections of this report. As was mentioned earlier, such a scheme is sub-optimum with respect to quantization noise error, but is better suited for channel noise immunity.

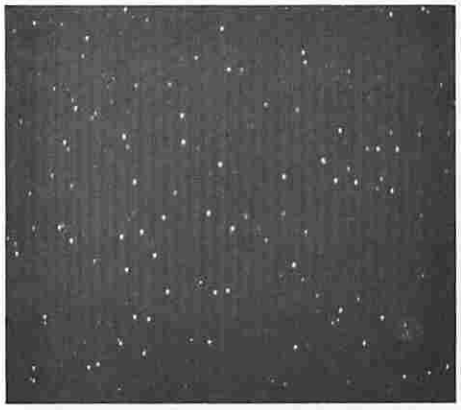
A sequence of computer noise simulation experiments have been conducted in order to verify the concepts presented earlier. Figures 7-3, 7-4, and 7-5 present the results of the simulation where three different noise rates: $Pe = 10^{-4}$, $Pe = 10^{-3}$, $Pe = 10^{-2}$, were introduced into the spatial, Fourier, and Hadamard domains respectively. In addition the difference pictures are included for visual purposes. The "salt and pepper" effect is quite evident in Figure 7-3 for spatial domain errors. For errors less than $Pe = 10^{-4}$ the transform domains indicate little or no degradation while a few errors are still

a. Pe = 10^{-4}

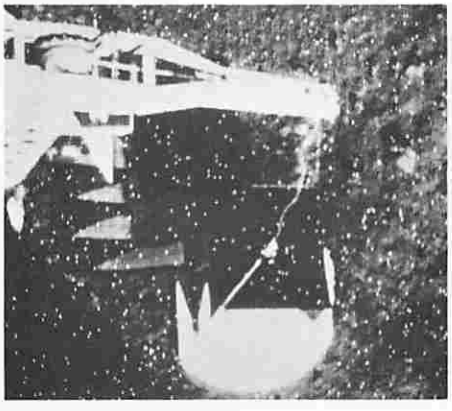
b. Difference



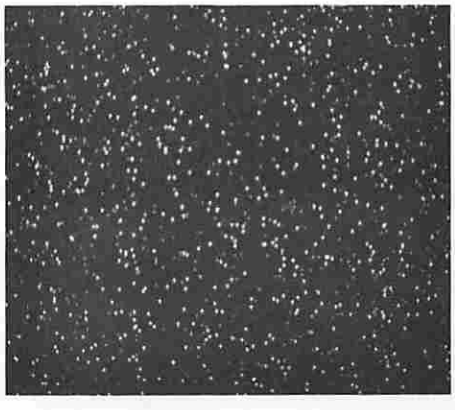
c. Pe = 10^{-3}



d. Difference



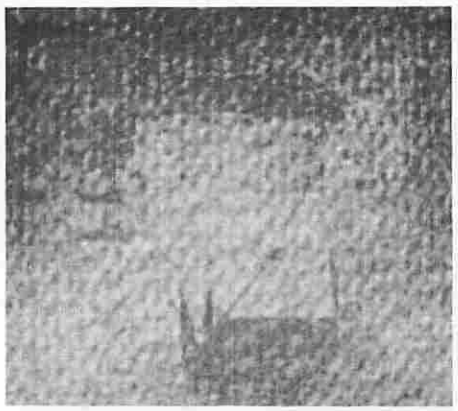
e. Pe = 10^{-2}



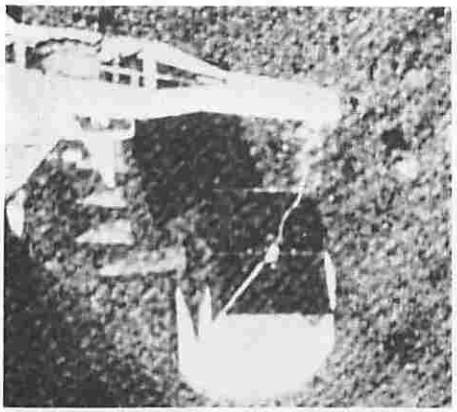
f. Difference

Figure 7-3 Spatial Domain Coding Effects of Channel Errors

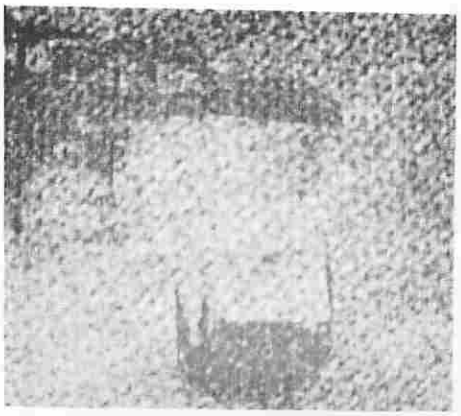
a. Pe = 10^{-4}



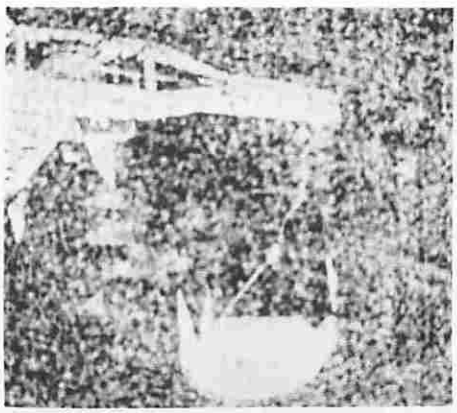
b. Difference



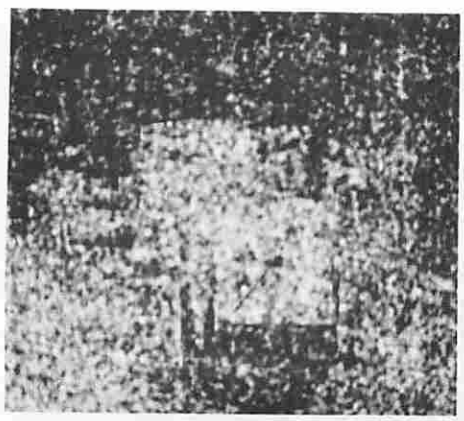
c. Pe = 10^{-3}



d. Difference



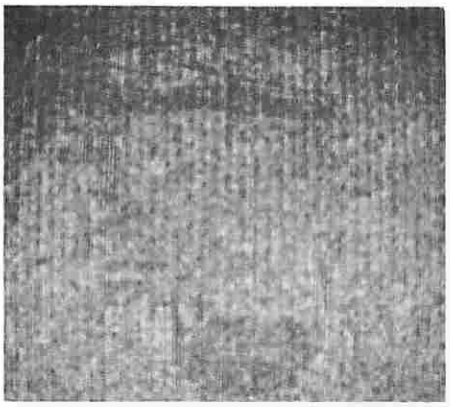
e. Pe = 10^{-2}



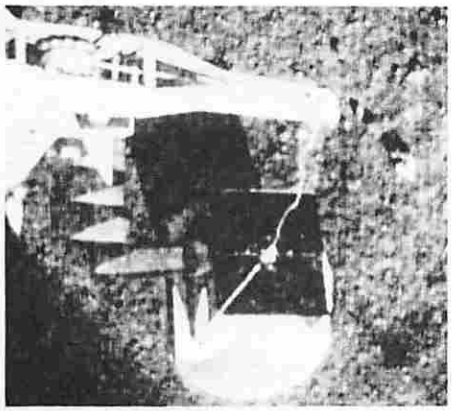
f. Difference

Figure 7-4 Fourier Transform Coding Effects of Channel Errors

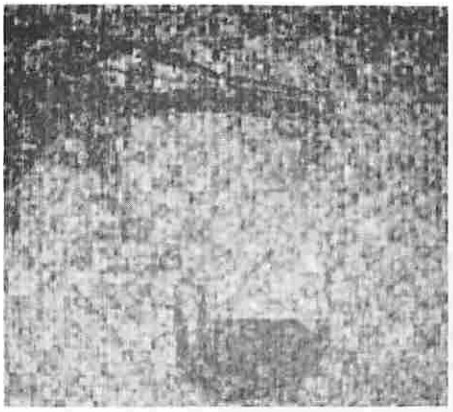
a. Pe = 10^{-4}



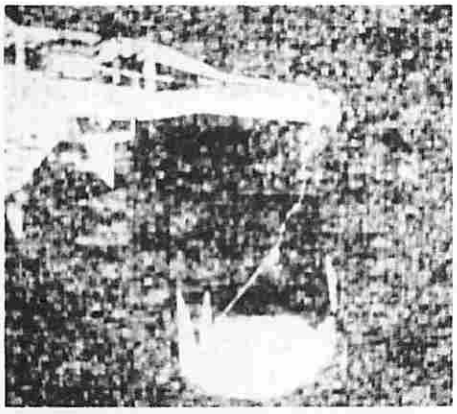
b. Difference



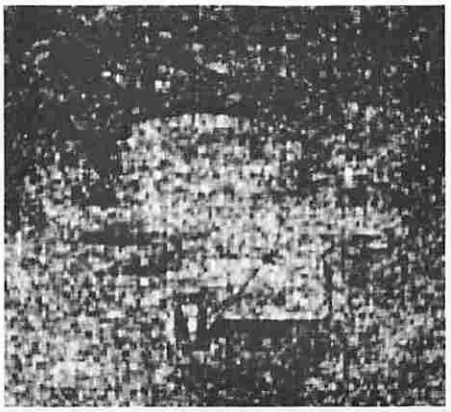
c. Pe = 10^{-3}



d. Difference



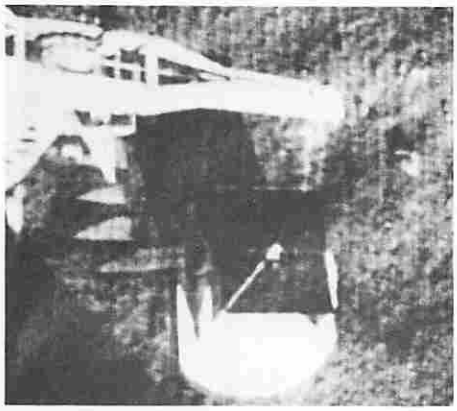
e. Pe = 10^{-2}



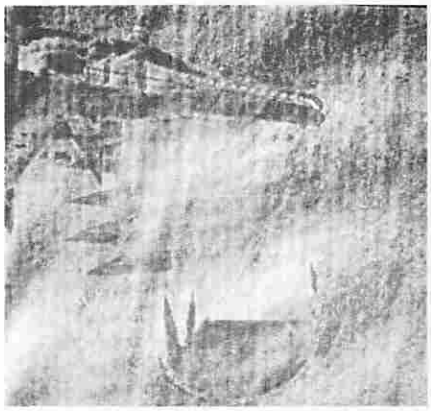
f. Difference

Figure 7-5 Hadamard Transform Coding Effects of Channel Errors

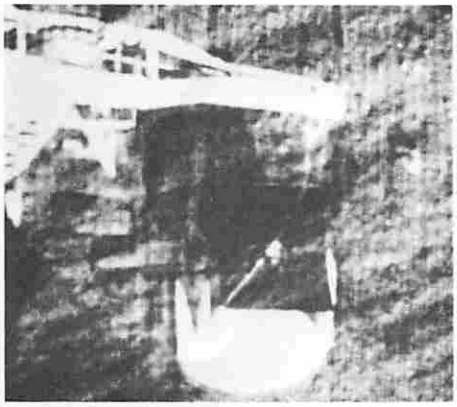
evident in the spatial domain noise. However, for larger noise rates the low order orthogonal functions that make up the respective transformations tend to swamp out the reconstructed image. This can be explained by the fact that the absolute, as opposed to relative value of a bit error is much larger in the regions where the transform coefficients (eigenvalues for the Karhunen-Loeve transform) are large in the transform domain. This explains the effect in Figures 7-4e and 7-5e. Further demonstration of this effect was presented in reference [27] where it was shown that by protecting certain areas of the transform domain from noise effects, large improvements in noise immunity could be obtained. This suggests an error correction procedure, a simulation of which is presented in the following section. However, before developing some error correction techniques, it is instructive to investigate the effects of a noisy channel on thresholded transform domains in order that both bandwidth reduction and noise immunity be combined. Figures 7-6 and 7-7 present results of such a simulation in which a threshold has been selected to provide a 5:1 sample reduction ratio. Again the difference pictures are presented for visual evaluation purposes. The noise effects now include run length errors in the transform domain which manifest themselves as a unique type of one dimensional blurring in the reconstructed images. Again, noise with errors less than $Pe = 10^{-4}$ tend to be averaged out due to the reconstruction process. The threshold coding technique



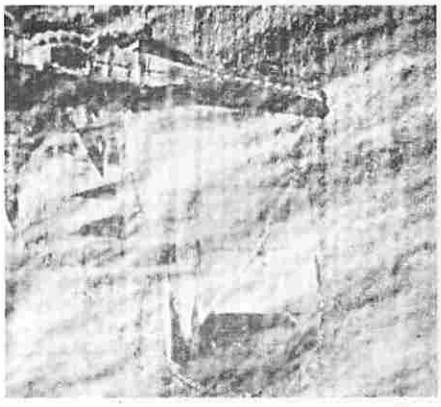
a. Pe = 10^{-4}



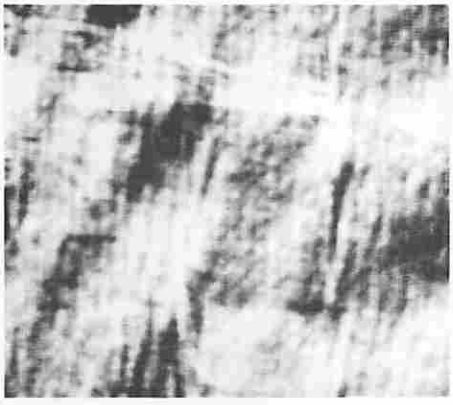
b. Difference



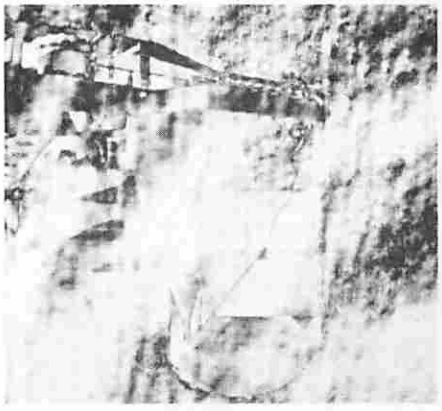
c. Pe = 10^{-3}



d. Difference

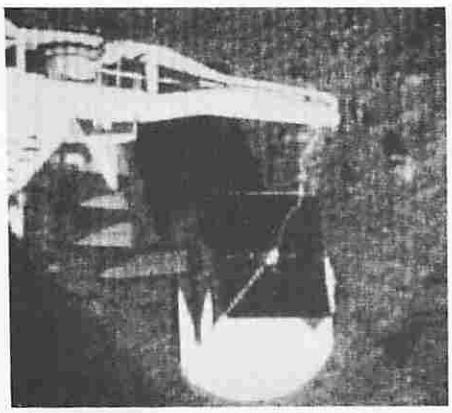


e. Pe = 10^{-2}

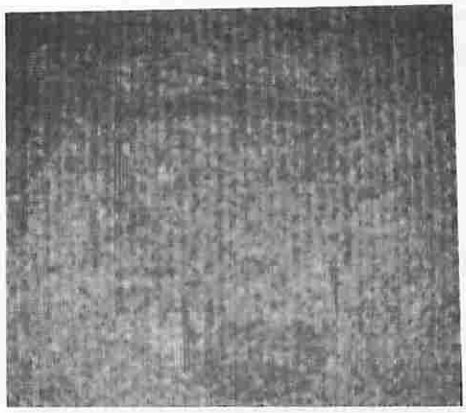


f. Difference

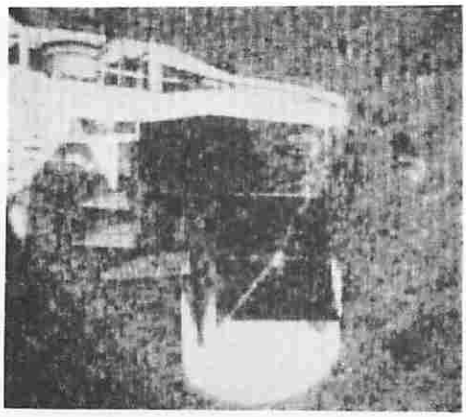
Figure 7-6 Fourier Transform Threshold Coding
Effects of Channel Errors S.R. = 5:1



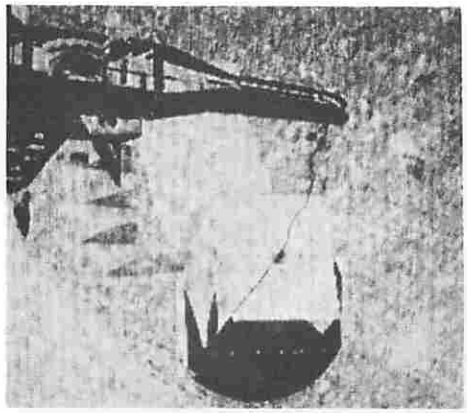
a. Pe = 10^{-4}



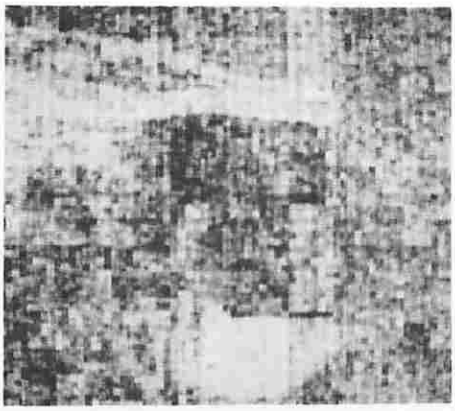
b. Difference



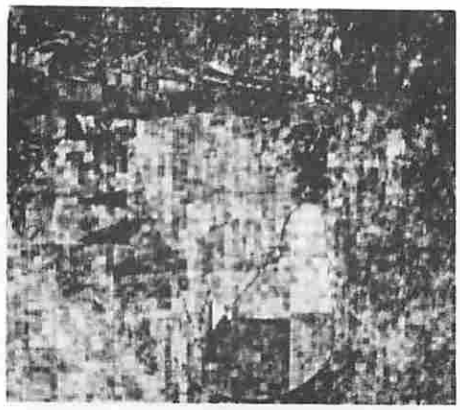
c. Pe = 10^{-3}



d. Difference



e. Pe = 10^{-2}



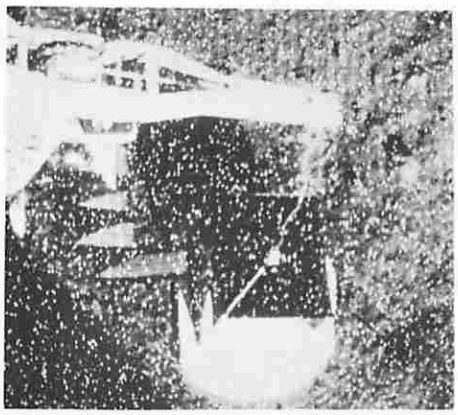
f. Difference

Figure 7-7 Hadamard Transform Threshold Coding
Effects of Channel Errors, S.R. = 5:1

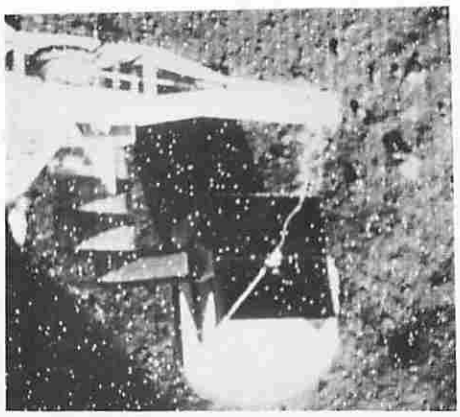
requires coding for position information. These code words should be uniformly distributed so that unexpected noise biasing does not occur in the position code as well as data code. More sophisticated coding techniques might be pursued in this area.

7.2 Error Correction Transform Coding

As a result of the statistical regularity of samples in the transform domain, a smaller amount of error correction in this domain will yield a better noise immunity than the same amount of error correction in the spatial domain. The nature of the quantization law is such that errors in certain positions of the transform domain are much more bothersome than in other positions due to the large statistical variance of samples at these coefficients. Therefore, it is natural to develop an error correction rule to correct for errors only in these large variance regions. One such rule would be to error correct code those transform samples which correspond to positions in the transform domain where the transform spectrum of the covariance function indicates a high probability of large sample value. This technique alone requires an increase in bandwidth to facilitate the error correction. However, it has been found that the small increase in bandwidth in the transform domain will result in better reconstructions than the same increase in the spatial domain. Figure 7-8 demonstrates this situation where a 3.5:1 increase in



a. Spatial domain errors



b. Spatial domain error correction

c. Hadamard domain error correction

d. Fourier domain error correction

Figure 7-8 Channel Error Rate Pe=4×10^{-2} 3.5:1 Increase in Bandwidth

bandwidth has greatly improved the transform coded image over the spatial coded image. It is important to emphasize that the coding technique used for the transform domain should be tailored to a particular channel capacity. If the channel noise has an error rate less than about 10^{-4}, then it appears that no error correction is necessary as in Figures 7-4a, 7-5a, 7-6a, and 7-7a. However, under the circumstances of a high error rate, it often becomes desirable to transmit as many error corrected samples as possible at the expense of either increased bandwidth or of not transmitting the entire transform plane. Using such a system, corrected, but not necessarily errorless, data could be received until either all data (and parity bits) are received for a complete picture or until normal picture bandwidth has been reached, at which time transmission is terminated. In order to implement such a scheme, an error correcting code must be selected.

A specific example of the potential of the transform error correction coding technique is presented below. A high error rate channel is assumed with rate $Pe = 4 \times 10^{-2}$. Three experiments are implemented, one of which uses an increased bandwidth and the other two utilize an equal bandwidth criterion such that the exact same number of bits is necessary to transmit the spatial domain as the transform coded domain, (256) (256) (6). A Bose Chandhuri-Hocquenghem (BCH) code [41, p. 163] which is capable of correcting a total of

seven errors is 31 bits long with 6 information bits (31, 6). Utilizing an error correcting code capable of seven error corrections does not mean that the six information bits will be received over the noisy channel error free. Since each code word length has been increased to thirty-one bits, eight or more errors per code word cannot be guaranteed to be corrected. The probability of having eight or more errors in the BCH code (31, 6) is given by the partial sum of the binomial distribution

$$p\,(8 \text{ or more errors}) = \sum_{i=8}^{31} \binom{31}{i} p^{i}(1-p)^{31-i} \qquad (7\text{-}1)$$

where p is the binary symmetric channel error rate. This probability is an upper bound for the incorrect reception of a code word since the possibility of correct reception for greater than seven errors still exists but is unknown. For the specific channel error rate of 4×10^{-2} the error corrected data samples will be received with probability of error no greater than 2.26×10^{-5} [42].

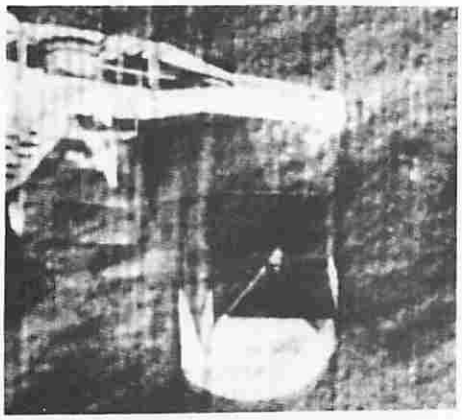
Figure 7-8 presents the results of an experiment in which an increased bandwidth has been allowed to compensate for the parity bits necessary in the error correction code. However, a (31/6):1 increase in bandwidth would be necessary to completely transmit the full data of either the space or transform domain. Allowing only a 3.5:1 bandwidth increase means that not all the data in the transform

domain is transmitted and thus Figures 7-8c and 7-8d are 1.44:1 low pass sequency and frequency filters with error rate 2.26×10^{-5}. The spatial domain error correction is the average of 70% BCH code and 30% no error coding.

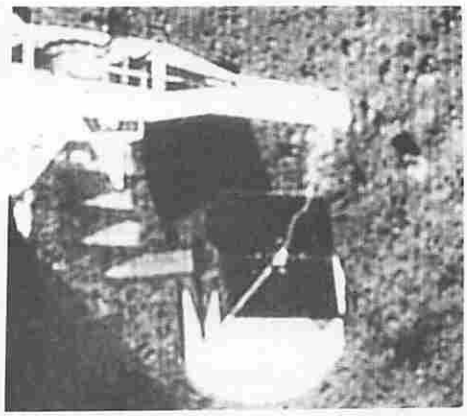
For spacecraft implementations it is desirable to transmit the error corrected image with no increased bandwidth requirement over conventional spatial domain transmission. Thus a (31/6):1 low pass frequency or sequency filter with error rate 2.26×10^{-5} will result in an equal bandwidth requirement. The results of this experiment are displayed in Figures 7-9b and 7-9d.

Because zonal low pass transform filtering is a non-adaptive technique for bandwidth reduction, it is desirable to utilize the adaptive feature of threshold coding as a means of more optimally compensating for the parity bits necessary for error coding. Thus a run length coding technique utilizing 4 position bits will be used in the transform domain.* Consequently, 4 position and 6 data bits will be used as information bits in the transform domain for run length coded thresholded transform samples. Thus a new error correcting code is necessary and a convenient candidate is a BCH (31, 11) code. This

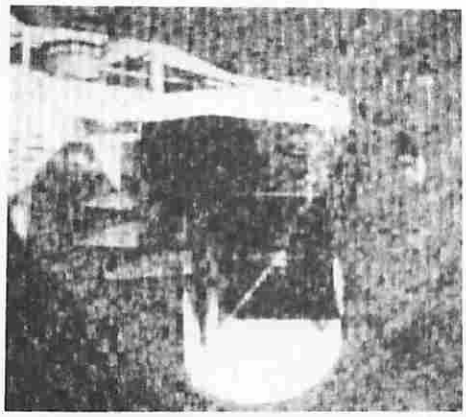
*A pseudo-run length coding technique is alluded to here enabling 4 rather than 8 bits necessary for position coding [43].



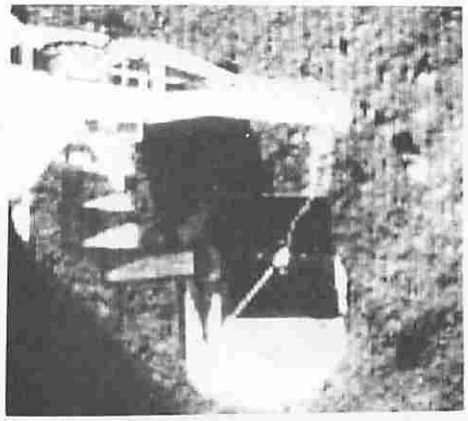
a. Fourier run length
error corrected retransformation



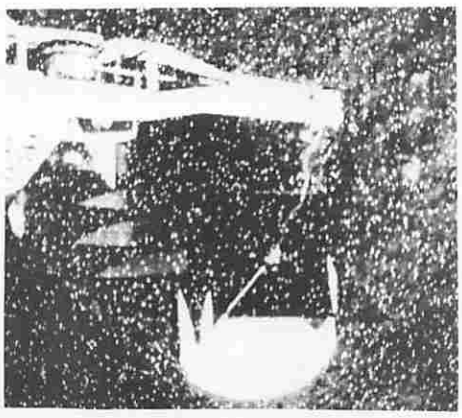
b. Fourier zonal error
corrected retransformation



c. Hadamard run length
error corrected retransformation



d. Hadamard zonal error
corrected retransformation



e. Spatial domain errors

Figure 7-9 Surveyor Box Equal Bandwidth
Error Correction Technique, Pe = 4×10^{-2}

code, again of length 31 bits, has 11 information bits and is capable of correcting 5 or less bit errors. Consequently, the probability of having 6 or more errors in the BCH code for each sample is given by the partial sum of the binomial distribution

$$p\,(6 \text{ or more errors}) = \sum_{i=6}^{31} \binom{31}{i} p^{i}(1-p)^{31-i} \qquad (7\text{-}2)$$

and is equal to 1.27×10^{-3} for a channel with error rate 4×10^{-2} [43]. Thus the cost of run length coding has changed the effective error rate from 2.26×10^{-5} to 1.27×10^{-3} for this example. Figure 7-9a and 7-9c are the run length error corrected retransformations with a 5:1 bandwidth reduction to compensate for the (31/6):1 parity information bandwidth increase. Consequently, again, an equal bandwidth criterion has been maintained.

It is suggested that other coding techniques could be developed which would improve upon these results. In fact, for potential hardware systems, research ought to be undertaken to develop the best code for the channel error rate, bandwidth, and computational complexity allowable.

8. Summary

This report has presented a theoretical development of several two dimensional transforms that are potentially useful for image coding. Of the transforms analyzed, the Fourier, Hadamard, and Karhunen-Loeve transforms have proven to possess the desired property of image energy compaction in the transform domain.

The energy compaction property of these three transforms has been exploited to achieve a sample reduction by two means: zonal sampling and threshold sampling. In zonal sampling a sampling mask corresponds to the positional ordering of the largest eigenvalues of the covariance matrix of the class of images to be coded. For the Fourier and Hadamard transforms the best sampling mask has a hyperbolic shape in the transform domain. Examples of the sample reduction achievable by zonal sampling with the three transforms are shown in Figures 8-1a to 8-1c. The transforms were taken in blocks of 16 by 16 elements. The other technique of sample deletion, called threshold sampling, simply entails the coding of each transform domain sample that exceeds a magnitude threshold level. By this technique the reconstruction of a particular image will suffer the least degradation from the standpoint of energy loss. Figures 8-1d to 8-1f illustrate the performance of threshold coding. The important conclusions to be drawn from Figure 8-1 and the supporting experimental results

a. Fourier, hyperbolic zonal sampling
4:1 sample reduction

b. Fourier, threshold sampling
5:1 sample reduction

c. Hadamard, hyperbolic zonal sampling
4:1 sample reduction

d. Hadamard, threshold sampling
5:1 sample reduction

e. Karhunen-Loeve, zonal sampling
4:1 sample reduction

f. Karhunen-Loeve, threshold sampling
5:1 sample reduction

Figure 8-1 Summary of Fourier, Hadamard, and Karhunen-Loeve Transform Image Coding in 16 by 16 Element Blocks

of Section 3 are that:

a. significant sample reduction factors can be obtained by zonal and threshold sampling in the transform domain for the Karhunen-Loeve, Fourier, and Hadamard transforms.

b. threshold sampling provides better performance (higher sample reduction factors for the same degree of image quality) than zonal sampling.

c. the Karhunen-Loeve transform exhibits somewhat better performance than the other two transforms, which in turn, exhibit about the same degree of performance.

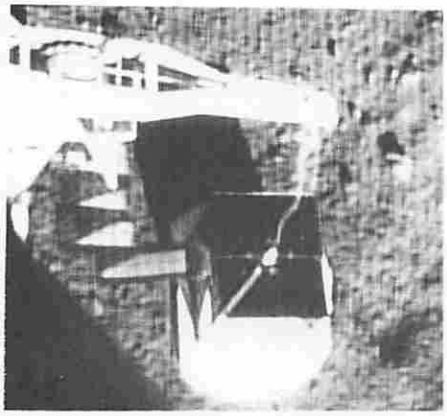
d. the sample reductions achieved were obtained by transform coding in blocks of only 16 by 16 elements. Image transformation in such small blocks can be implemented quite simply.

Fast computational algorithms exist for the Fourier and Hadamard transforms. Computation of these transforms on a general purpose computer in blocks of up to 1024 by 1024 elements appears feasible from a computational standpoint. There is no fast computation algorithm for the Karhunen-Loeve transform. This fact coupled with the realization that the Karhunen-Loeve does not perform appreciably better than the Fourier and Hadamard transforms seems to limit the practical utility of the Karhunen-Loeve transform. For these reasons the detailed analysis of the report has been limited primarily to the Fourier and Hadamard transforms.

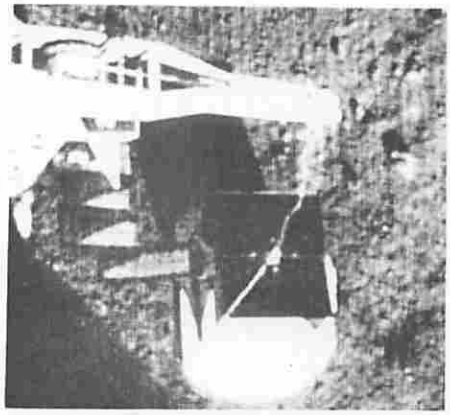
An analysis has been performed to determine the optimum means of transform domain sample quantization. The results of this analysis indicate that for a quantization strategy in which each sample is coded to the same number of levels, the optimum quantizer places the quantization levels along a nonlinear scale both in sample amplitude and position in the transform domain. Unfortunately, the optimum is difficult to implement. Therefore, several nonlinear scales that could be deterministically computed were analyzed. The best performance has been obtained with a Gaussian error function quantizer. With this quantizer, good quality reconstructions have been obtained with 64 quantization levels (6 bits) per transform sample component for both the Fourier and Hadamard transforms.

Zonal and threshold sampling of quantized Fourier and Hadamard transforms of images has been investigated in detail for a variety of images. The transforms have been taken in blocks of up to 256 by 256 elements. A position coding technique for threshold sampling employing run length coding has been implemented and evaluated. Figure 8-2 illustrates the effects of threshold coded quantized Fourier and Hadamard transforms of images over a full frame of 256 by 256 elements.

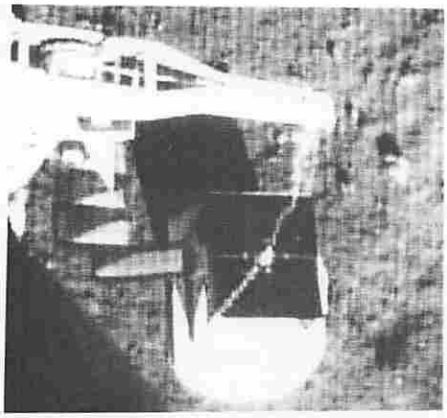
The conclusions to be drawn from these experiments are that:



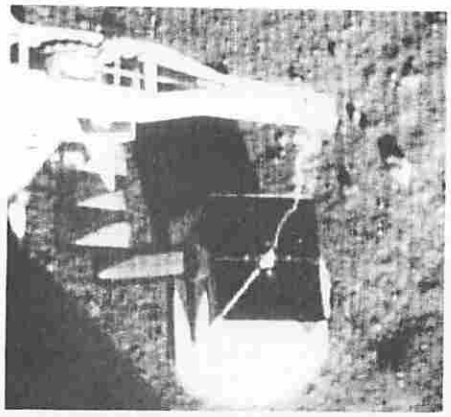
a. Fourier, hyperbolic zonal sampling
4:1 sample reduction



b. Fourier, threshold sampling
5:1 sample reduction



c. Hadamard, hyperbolic zonal sampling
4:1 sample reduction



d. Hadamard, threshold sampling
5:1 sample reduction

Figure 8-2 Summary of Fourier and Hadamard Transform Full Frame Image Coding -- Quantized and Coded Images

a. for any size block,threshold sampling provides better performance than zonal sampling.

b. performance is better for larger size blocks, but the difference in performance between blocks of 16 by 16 elements and blocks of 256 by 256 elements is not great.

c. for threshold sampling, simple run length coding can be employed to code the position of significant samples; the run length coding does not affect image quality, and can be accomplished with a relatively few number of bits per image element.

The effect of channel errors on transform coded images has been studied. It has been found that channel errors in the transform domain tend to cause a small overall loss in resolution; there are no discrete effects like the "salt and pepper" errors that appear in normal spatial domain coding. Experiments verify that errors in the position bits coding the position of significant samples in threshold coding are not serious. Errors in the lowest spatial frequencies (sequencies) have been found to degrade an image the most. By applying channel error correction to a relatively small number of these transform domain samples, a relatively large improvement in the tolerance to channel errors can be obtained. The equivalent amount of error correction in the spatial domain would provide no worthwhile improvement.

In final summary, it can be said that Fourier and Hadamard transform image coding techniques are a feasible means of obtaining significant bandwidth compressions for digital image transmission. Side benefits of transmitting the Fourier or Hadamard transform of an image rather than the image itself are an improved tolerance to channel errors and the fact that image enhancement methods can be readily performed in the transform domain.

9. Recommendations

The general concept of transform coding has now been studied and evaluated rather thoroughly in this research study and by other investigators. There remain three areas, listed below, that merit further study.

Transform Domain Coding

Zonal sampling in the transform domain has the advantage of simplicity, but achievable performance is not as great as can be obtained by threshold sampling. However, threshold sampling requires position coding of significant samples. It appears that advantages of both techniques might be obtained by a hybrid scheme of zonal sampling a set of the low spatial frequencies (sequencies) and threshold sampling the remainder of the transform domain. Schemes for performing this type of sampling should be investigated in conjunction with a study of the best means of position coding significant samples.

The quantization technique presented in this report adopted the strategy of assigning the same number of bits per transform domain sample and then determining the optimum scaling of quantization levels. Another technique that has been reported [37] utilizes a linear quantization scale for each sample, but the number of bits per sample is optimally selected to minimize the total number of image code bits for

a given error criterion. It appears that it would be advantageous to combine both strategies: assign the number of bits per sample on the basis of the sample variance, and select the quantization levels according to a nonlinear scale based upon the variance. This quantization method should be studied further.

Implementation

A number of companies have available equipment to perform a fast Fourier transform in one dimension for up to about 1024 points. A few companies have built fast Hadamard transform devices for one dimensional transforms. There are presently no two dimensional transform processors on the market.

In view of the great potential for image transform coding it would seem worthwhile to implement prototype Fourier and Hadamard transform processors. As a first step a 16 by 16 element processor should be built and evaluated.

Color Image Coding

Conventional color images are represented by three overlapping intensity planes corresponding to three primary colors--red, green, and blue. In normal television practice a linear color transition is made into three planes which represent the luminance (the monochromatic representation of an image) and the two chrominance variations of the image. The spatial frequency response of the eye to chrominance

information is very poor. Therefore, a great deal of spatial low pass filtering on the chrominance planes can be tolerated. Fourier and Hadamard zonal low pass filtering appear ideal for this application. Studies are needed to determine the effects of Fourier and Hadamard filtering on the chrominance planes and to determine the best color transitions for the subsequent filtering operation.

APPENDIX A

Image Covariance Function*

Measurements have been made of the covariance function of an image to determine the fit of the Gauss-Markov process model. Figure A-1 shows plots of the correlation between elements along a line, between elements along a column of the image, and between elements along the diagonal of an image. All measurements have been made on the Surveyor spacecraft scene. The data points have been fit by functions of the form A^n where A is the correlation between adjacent elements and n is the separation between elements. The fit along the rows and columns of the image appears to be reasonably good. As shown in the figure there is a small deviation between the Gauss-Markov process model for diagonal elements and actual measurements.

---

*Measurements have been performed by Professor Lee D. Davisson of the University of Southern California.

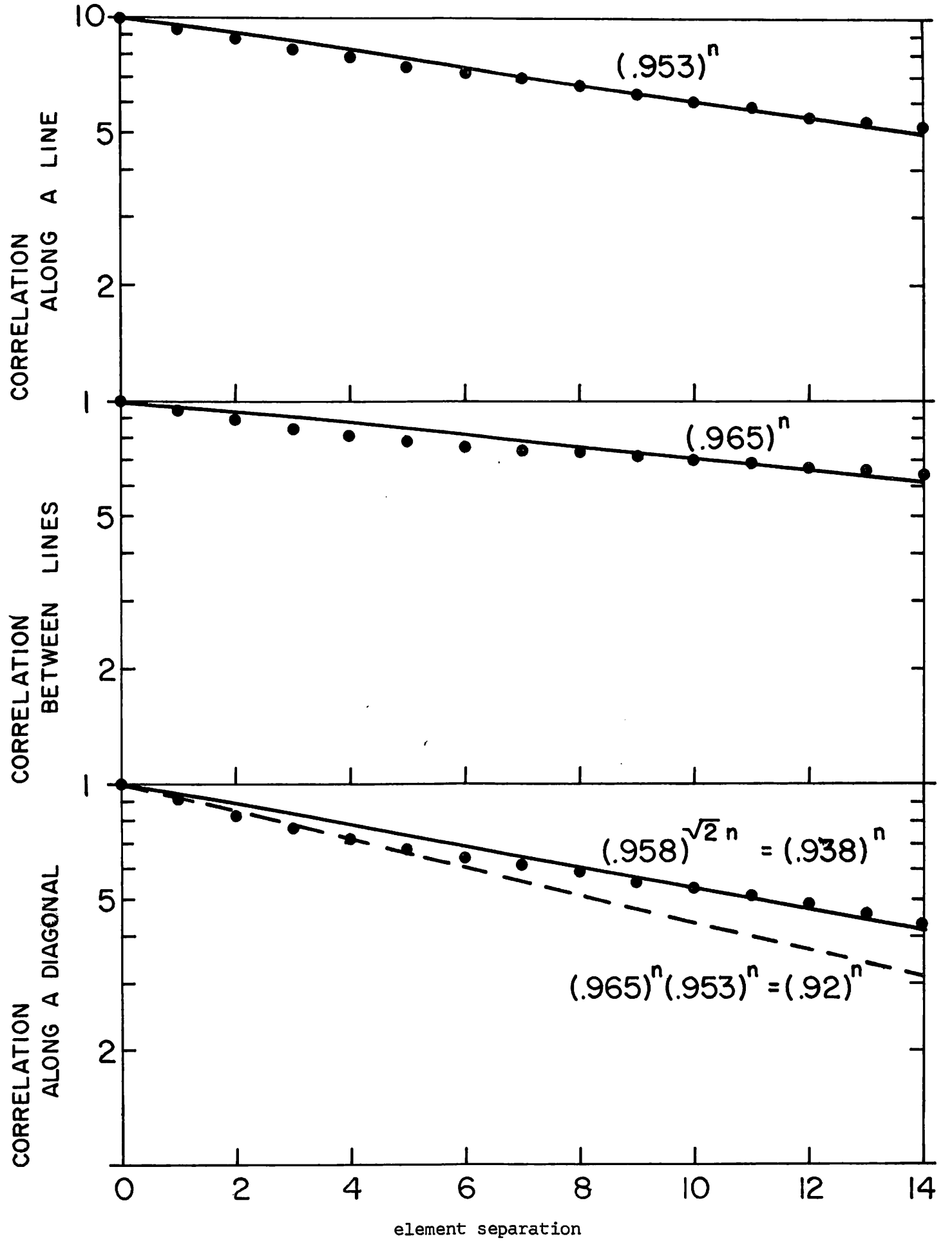


Figure A-1 Measurements of Surveyor Spacecraft Scene

References

1. W.K. Pratt "A Bibliography on Television Bandwidth Reduction Studies", IEEE Transactions on Information Theory, Vol. IT-13, No. 1 (January, 1967), pp. 114-115.

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

3. Special issue on redundancy reduction Proceedings IEEE, Vol. 55, No. 3 (March, 1967).

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

5. H.C. Andrews and W.K. Pratt "Television Bandwidth Reduction by Fourier Image Coding", Society of Motion Picture and Television Engineers, 103rd Technical Conference, (May, 1968).

6. H.C. Andrews and W.K. Pratt "Television Bandwidth Reduction by Encoding Spatial Frequencies", Journal Society of Motion Picture and Television Engineers, Vol. 77 (December, 1968), pp. 1279-1281.

7. W.K. Pratt, J. Kane, and H.C. Andrews "Hadamard Transform Image Coding", Proceedings IEEE, Vol. 57, No. 1 (January, 1969). pp, 58-68,

8. W.K. Pratt and H.C. Andrews "Application of Fourier-Hadamard Transformation to Bandwidth Compression", MIT Symposium on Picture Bandwidth Compression, (April, 1969).

9. W.K. Pratt and H.C. Andrews "Two Dimensional Transform Coding of Images", 1969 International Symposium of Information Theory Institute of Electrical and Electronic Engineers, (November, 1968).

10. H.C. Andrews and W.K. Pratt "Transformation Coding for Noise Immunity and Bandwidth Reduction", Second Annual Hawaii International Conference on System Sciences, (January, 1969).

11. H.C. Andrews and W.K. Pratt "Transform Image Coding", PIB International Symposium on Computer Processing in Communications, (April, 1969).

12. J. W. Cooley, P. A. W. Lewis, and P. D. Welch "Historical Notes on the Fast Fourier Transform", Proceedings IEEE, Vol. 55, October, 1967, pp. 1675-1677.

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

14. J. Hadamard "Resolution d'une Question Relative aux Determinants", Bulletin des Sciences Mathematiques , (2), Vol. 17, part 1, (1893), pp. 240-246.

15. H. J. Ryser, Combinatorial Mathematics, John Wiley, New York, (1963).

16. S. W. Golomb, et al. Digital Communications, Prentice-Hall, (1964).

17. H. F. Harmuth "A Generalized Concept of Frequency and Some Applications", IEEE Transactions on Information Theory, Vol. IT-14, No. 3, (May, 1968), pp. 375-382.

18. J. L. Walsh "A Closed Set of Orthogonal Functions", American Journal Mathematics, Vol. 45, (1923), pp. 5-24.

19. N. J. Fine "On the Walsh Functions", Transactions American Mathematical Society, Vol. 65, (1949), pp. 372-414.

20. N. J. Fine "The Generalized Walsh Functions", Transactions American Mathematical Society, Vol. 69, (1950), pp. 66-77.

21. G. W. Morgenthaler "On Walsh-Fourier Series", Transactions American Mathematical Society, Vol. 84, (1957), pp. 472-507.

22. K. W. Henderson "Some Notes on the Walsh Functions", IEEE Transactions on Electronic Computers, Vol. EC-13, (February, 1964), pp. 50-52.

23. H. Rademacher "Einige Satze von Allgemeinen Orthogonal-Funktionen", Mathematics Annals, Vol. 87, (1922), pp. 122-138.

24. H. C. Andrews and J. Kane "Kronecker Matrices, Computer Implementation, and Generalized Spectra", Journal of the Association of Computer Machinery, (April, 1970).

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

26. H. C. Andrews and W. K. Pratt "Transform Data Coding", PIB Symposium on Computer Processing in Communications, April, 1969.

27. W. K. Pratt and H. C. Andrews "Transform Processing and Coding of Images", University of Southern California, Electronic Sciences Laboratory, USCEE Report No. 341 (March, 1969), Chapter 2.

28. 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.

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

30. 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.

31. J. E. Whelchel, Jr., and D. F. Guinn "The Fast Fourier-Hadamard Transform and Its Use in Signal Representation and Classification", EASCON 1968 Convention Record, (1968), pp. 561-573.

32. J. J. Y. Huang and P. M. Schutheiss "Block Quantization of Correlated Gaussian Random Variables", IEEE Transactions on Communication Systems, Vol. CS-11, No. 3 (September, 1963), pp. 289-296.

33. T. Y. Young and W. H. Huggins "On the Representation of Electrocardiographs", IEEE Transactions on Bio-Medical Electronics, Vol. BME-10, No. 3 (July, 1963), pp. 86-95.

34. L. M. Goodman "A Binary Linear Transformation for Redundancy Reduction" , Proceedings IEEE Letters, Vol. 55, No. 3 (March, 1967), pp. 467-468.

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

36. C.J. Palermo, R.V. Palermo, and H. Horowitz "The Use of Data Omission for Redundancy Removal", Record International Space Electronics and Telemetry Symposium, (1965), pp. (11)D1-(11)D16.

37. A. Habibi and P. Wintz "Optimum Linear Transformations for Encoding 2-Dimensional Data".

38. L.A. Pipes Matrix Methods in Engineering, Prentice-Hall, Englewood Cliffs, New Jersey (1963).

39. A. Papoulis Probability, Random Variables, and Stochastic Processes, McGraw-Hill Book Company (1965).

40. P.F. Panter and W. Dite "Quantization Distortion in Pulse Count Modulation with Nonuniform Spacing of Levels", Proceedings IRE, Vol. 39, No. (January, 1951), pp. 44-48.

41. W.W. Peterson Error Correcting Codes, The MIT Press, Cambridge, Massachusetts (1961).

42. "Tables of the Binomial Probability Distribution", Departments of Commerce, National Bureau of Standards, Applied Mathematics Series No. 6 (January, 1950).

43. W.K. Pratt "Stop Scan Edge Detection Systems of Television Bandwidth Reduction", USCEE Report 13IT, (June, 1965).