pacity for noisy filter channels," in *Conf. Rec., 1972 IEEE Int. Conf. Communications*, Philadelphia, Pa., sec. 22, pp. 12–16.
[10] P. A. Bello, J. W. Graham, and L. Ehrman, "Troposcatter multichannel digital systems study," Rome Air Develop. Cent., Rome, N. Y., RADC TR-67-218, May 1967.
[11] ——, "Troposcatter multimode FSK/PSK modem," Rome Air Develop. Cent., Rome, N. Y., RADC Final Rep. F30602-68-C-047, July 1970.
[12] P. A. Bello, L. Ehrman, and D. Arnstein, "Modeling and data analysis—short and medium range digital troposcatter tests," Rome Air Develop. Cent., Rome, N. Y., RADC-TR-69-233, Oct. 1969.
[13] P. E. Green, Jr., "Radar measurements of target scattering properties," *Radar Astronomy*, J. V. Evans and T. Hagfors, Ed. New York: McGraw-Hill, 1968, ch. 1, p. 1.
[14] P. A. Bello, "A study of the relationship between multipath distortion and wave number spectrum of refractive index in radio links," *Proc. IEEE*, vol. 59, pp. 47–75, Jan. 1971.
[15] P. Monsen and S. H. Richman, "A comparison of digital techniques for troposcatter applications," in *Conf. Rec., 1973 Nat. Telecommunications Conf.*, Atlanta, Ga., Nov. 1973.
[16] E. R. Shamash, "Communication in the presence of intersymbol interference," Ph.D. dissertation, Univ. of California, Los Angeles, 1973.

Peter Monsen (S'60–M'64) was born in Quincy, Mass., on March 24, 1940. He received the B.S. degree in electrical engineering from Northeastern University, Boston, Mass., in 1962, the M.S. degree in operation research from the Massachusetts Institute of Technology, Cambridge, in 1963, and the Eng.Sc.D. degree in electrical engineering from Columbia University, New York, N. Y., in 1970.

During a two year period beginning in January 1964, he served as a lieutenant in the U. S. Army at the Defense Communication Agency, Arlington, Va., where he was concerned with communication systems engineering. From 1966 to 1972 he was with the Bell Telephone Laboratories, Holmdel, N. J., where he was Supervisor of a Transmission Studies Group whose work included fading-channel characterization and adaptive equalization. He is presently with Signatron, Inc., Lexington, Mass., where he is responsible for the systems engineering of a 13-Mbps troposcatter modem being developed for USA ECOM.

Dr. Monsen is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

Slant Transform Image Coding

WILLIAM K. PRATT, MEMBER, IEEE, WEN-HSIUNG CHEN, AND LLOYD R. WELCH

***Abstract*—A new unitary transform called the slant transform, specifically designed for image coding, has been developed. The transformation possesses a discrete sawtoothlike basis vector which efficiently represents linear brightness variations along an image line. A fast computational algorithm has been found for the transformation.**

The slant transformation has been utilized in several transform image-coding systems for monochrome and color images. Computer simulation results indicate that good quality coding can be accomplished with about 1 to 2 bits/pixel for monochrome images and 2 to 3 bits/pixel for color images.

I. INTRODUCTION

DURING the past twenty years the applications of electronic imagery have grown enormously. This growth has placed severe demands on the capabilities of communication systems, since conventional television transmission requires exceptionally wide bandwidths. One means of bandwidth reduction that has shown particular promise is the transform image-coding process [1].

In 1968, the concept of coding and transmitting the two-dimensional Fourier transform of an image, computed by a fast computational algorithm, rather than the image itself, was introduced [2], [3]. This was followed shortly thereafter by the discovery that the Hadamard transform could be utilized in place of the Fourier transform with a considerable decrease in computational requirements [4], [5]. Investigations then began into the application of the Karhunen–Loeve [6], [7] and Haar [8] transforms for image coding. The Karhunen–Loeve transform provides minimum mean-square error coding performance but, unfortunately, does not possess a fast computational algorithm. On the other hand, the Haar transform has the attribute of an extremely efficient computational algorithm, but results in a larger coding error. None of the transforms mentioned above, however, has been expressly tailored to the characteristics of an image.

A major attribute of an image transform is that the transform compact the image energy to a few of the transform domain samples. A high degree of energy compaction will result if the basis vectors of the transform matrix "resemble" typical horizontal or vertical lines of an image. If the lines of a typical monochrome image are examined, it will be found that a large number of the lines are of constant grey level over a considerable length. The Fourier, Hadamard, and Haar transforms possess

Paper approved by the Associate Editor for Communication Theory of the IEEE Communications Society for publication after presentation at the 1973 Symposium on Applications of Walsh Functions. Manuscript received October 15, 1973; revised March 1, 1974. This work was supported in part by the Advanced Research Projects Agency of the Department of Defense and monitored by the Air Force Eastern Test Range under Contract F08606-72-C-0008.

W. K. Pratt and L. R. Welch are with the Electrical Engineering Department, Image Processing Institute, University of Southern California, Los Angeles, Calif. 90007.

W.-H. Chen was with the Electrical Engineering Department, Image Processing Institute, University of Southern California, Los Angeles, Calif. 90007. He is now with the Philco-Ford Corporation, Palo Alto, Calif.

a constant valued basis vector that provides an efficient representation for constant grey level image lines, while the Karhunen–Loeve transform has a nearly constant basis vector suitable for this representation. Another typical image line is one which increases or decreases in brightness over the length in a linear fashion. None of the transforms previously mentioned possess a basis vector that efficiently represents such image lines.

Shibata and Enomoto have introduced orthogonal transforms containing a "slant" basis vector for data of vector lengths of four and eight [9]. The slant vector is a discrete sawtooth waveform decreasing in uniform steps over its length, which is suitable for efficiently representing gradual brightness changes in an image line. Their work gives no indication of a construction for larger size data vectors, nor does it exhibit the use of a fast computational algorithm. In order to achieve a high degree of image-coding compression with transform coding techniques, it is necessary to perform the two-dimensional transform over block sizes 16×16 picture elements or greater. For large block sizes, computation is usually not feasible unless a fast algorithm is employed.

With this background, an investigation was undertaken to develop an image-coding slant-transform matrix possessing the following properties: 1) orthonormal set of basis vectors; 2) one constant basis vector; 3) one slant basis vector; 4) sequency property; 5) variable size transformation; 6) fast computational algorithm; and 7) high energy compaction. The following sections describe the construction of the slant-transformation matrix, present a fast computational algorithm, discuss its image-coding performance, and provide examples of its use for coding monochrome and color images.

Original images employed in the simulation of the coding systems studied are shown in Fig. 1. The monochrome images contain 256×256 pixels with each calibrated intensity value linearly quantized to 256 levels. The red, green, and blue tristimulus values of the color pictures have each been linearly quantized to 256 levels.

II. SLANT TRANSFORM

A slant-transform matrix possessing all of the above properties has been developed [10]. A description of its construction, fast algorithm, and application to image transformation follows.

A. Two-Dimensional Transform

Let $[F]$ be an $N \times N$ matrix of the picture element (pixel) brightness values of an image and let $[f_i]$ be an $N \times 1$ vector representing the ith column of $[F]$. The one-dimensional transform of the ith image line is then

$$[\mathfrak{f}_i] = [S][f_i] \tag{1}$$

where $[S]$ is the $N \times N$ unitary slant matrix. A two-dimensional slant transform is performed by sequential row and column transformations on $[F]$, yielding

$$[\mathfrak{F}] = [S][F][S]^T. \tag{2}$$

The inverse transformation to recover $[F]$ from the transform components $[\mathfrak{F}]$ is given by

$$[F] = [S]^T[\mathfrak{F}][S]. \tag{3}$$

It is also convenient to establish a series representation of the transform operation. The two-dimensional forward and inverse transforms in series form can be expressed as

$$\mathfrak{F}(u,v) = \sum_{j=1}^{N}\sum_{k=1}^{N} F(j,k)\,S(u,j)\,S(k,v) \tag{4}$$

and

$$F(j,k) = \sum_{u=1}^{N}\sum_{v=1}^{N} \mathfrak{F}(u,v)\,S(j,u)\,S(v,k). \tag{5}$$

Fig. 2 contains full size, two-dimensional slant-transform displays of the three monochrome original images of Fig. 1. The logarithm of the absolute value of each transform sample is displayed in Fig. 2(a)–(c) rather than the absolute value itself in order to reduce the dynamic range of the display. Fig. 2(d) shows a threshold display of Fig. 2(a) in which all samples whose magnitude is below a specified threshold are set to zero and all samples whose magnitude is above the threshold are set to a constant brightness. The typical energy distribution of the slant transform of an image is apparent from these pictures.

B. Slant-Transform Matrix

The slant transform is a member of a class of transforms whose matrices are orthogonal, have a constant function for the first row, and have a second row which is a linear (slant) function of the column index. The matrices are formed by an iterative construction that exhibits the matrices as products of sparse matrices, which in turn leads to a fast transform algorithm.

The slant-transform matrix of order two consisting of a constant and a slant basis vector is given by

$$S_2 = \frac{1}{2^{1/2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{6}$$

The slant matrix of order four is obtained by the operation

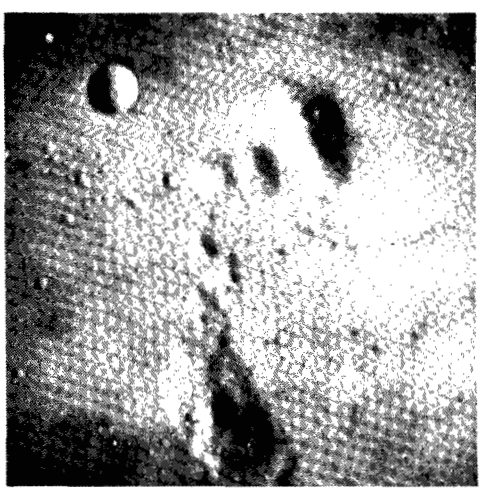
$$S_4 = \frac{1}{2^{1/2}}\left[\begin{array}{cc:cc} 1 & 0 & 1 & 0 \\ a_4 & b_4 & -a_4 & b_4 \\ \hdashline 0 & 1 & 0 & -1 \\ -b_4 & a_4 & b_4 & a_4 \end{array}\right]\left[\begin{array}{c:c} S_2 & 0 \\ \hdashline 0 & S_2 \end{array}\right] \tag{7a}$$

or

$$S_4 = \frac{1}{2^{1/2}}\begin{bmatrix} 1 & 1 & 1 & 1 \\ a_4 + b_4 & a_4 - b_4 & -a_4 + b_4 & -a_4 - b_4 \\ 1 & -1 & -1 & 1 \\ a_4 - b_4 & -a_4 - b_4 & a_4 + b_4 & -a_4 + b_4 \end{bmatrix} \tag{7b}$$

(a) monochrome girl

(b) monochrome couple



(c) monochrome moonscape

(d) color girl

(e) color couple

Fig. 1. Original monochrome and color images.

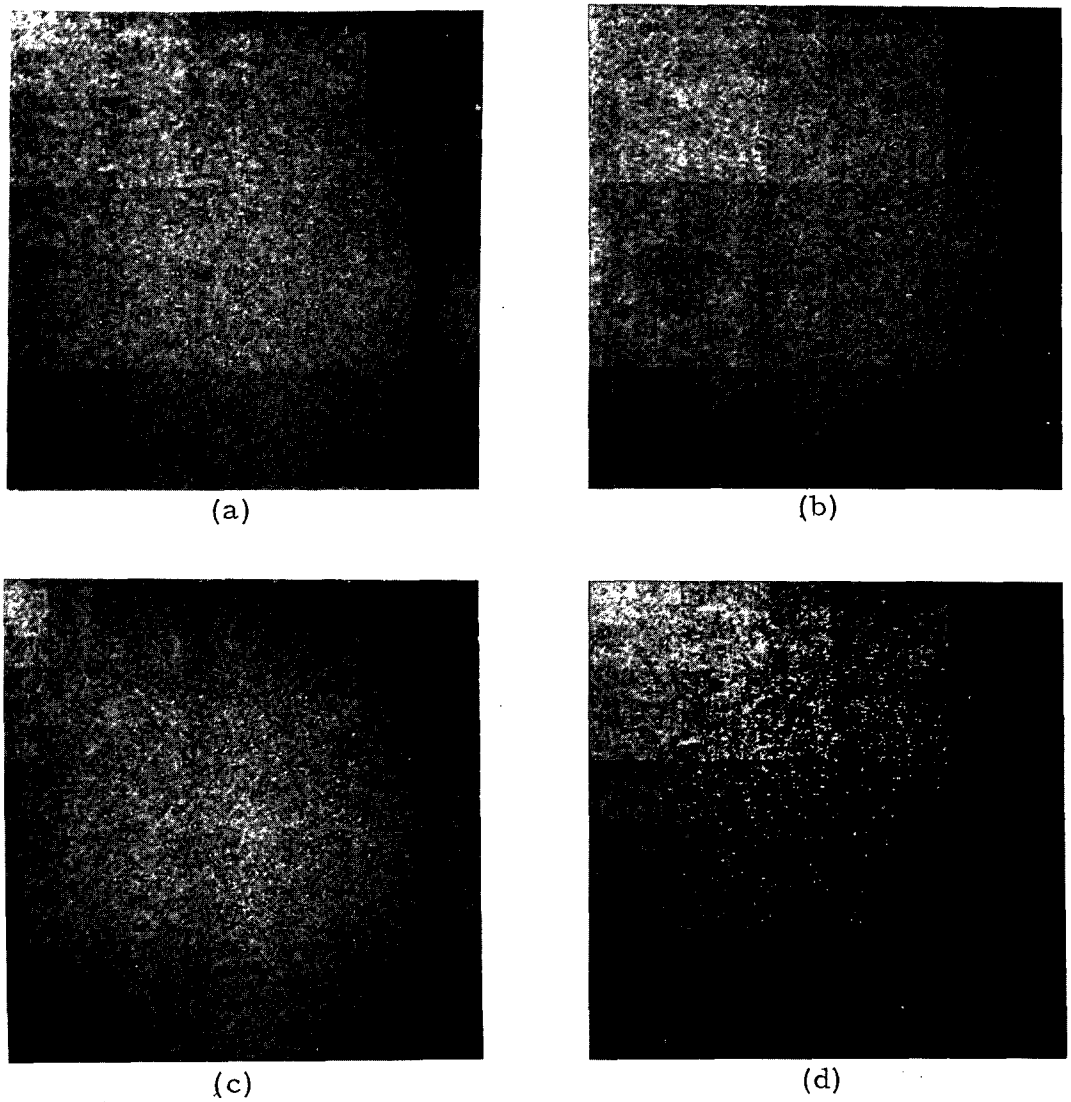


Fig. 2. Slant-transform domain logarithm of magnitude display. (a) Girl. (b) Couple. (c) Moonscape. (d) Transform threshold girl.

where a_4 and b_4 are scaling constants. The first two rows of the left factor of (7a) are uniquely determined by the orthogonality condition and the requirement that the first row of S_4 must be a positive constant and the second row must be linear with negative slope. The step sizes between adjacent elements of the slant vector of (7b) are $2b_1$, $2a_4 - 2b_4$, and $2b_4$. By setting these steps sizes equal, one finds that

$$a_4 = 2b_4.$$

The orthonormality condition $[S_4][S_4]^T = [I]$ leads to

$$b_4 = \frac{1}{5^{1/2}}.$$

The third and fourth rows of the left factor of (7a) form an orthonormal basis for vectors orthogonal to the first two rows and have the property that among all such bases, this basis requires the smallest number of nontrivial multiplications. In considering computational requirements, the scalar multiplier $1/\sqrt{2}$ is ignored; multiplication by ± 1 is considered trivial). When the values of a_4 and b_4 are substituted into (7b), the slant matrix of order 4 becomes

$$S_4 = \frac{1}{4^{1/2}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{3}{5^{1/2}} & \frac{1}{5^{1/2}} & \frac{-1}{5^{1/2}} & \frac{-3}{5^{1/2}} \\ 1 & -1 & -1 & 1 \\ \frac{1}{5^{1/2}} & \frac{-3}{5^{1/2}} & \frac{3}{5^{1/2}} & \frac{-1}{5^{1/2}} \end{bmatrix}. \tag{8}$$

It is easily seen that the rows of S_4 form an orthonormal set. Furthermore, S_4 possesses the sequency property; each row has a distinct number of sign reversals and each

integer from 0 to 3 is the number of sign reversals of some row.

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

$$S_8 = \frac{1}{2^{1/2}} \left[\begin{array}{cccc:cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ a_8 & b_8 & 0 & 0 & -a_8 & b_8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hdashline 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ -b_8 & a_8 & 0 & 0 & b_8 & a_8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{array}\right] \left[\begin{array}{c:c} S_4 & 0 \\ \hdashline 0 & S_4 \end{array}\right]. \tag{9}$$

The first two rows of the left factor are determined by orthonormality and the requirement that the first row of S_8 be constant and the second row be linear with negative slope. Rows five and six of S_8 are also linear combinations of the first two rows of S_4. The particular combinations are chosen to minimize nontrivial multiplications while maintaining orthonormality. The remaining rows form a basis for the orthogonal complement of the space spanned by the four rows already treated. It is a basis which requires no nontrivial multiplications. It can be shown that S_8 has the sequency property.

Equation (9) can be generalized to give the slant matrix of order $N (n = 2^n, n = 3, \cdots)$ in terms of the slant matrix of order $N/2$ by the following recursive relation:

$$a_2 = 1 \tag{11a}$$

$$b_N = [1 + 4(a_{N/2})^2]^{-1/2} \tag{11b}$$

$$a_N = 2b_N a_{N/2} \tag{11c}$$

or by the formulas

$$a_{2N} = \left(\frac{3N^2}{4N^2 - 1}\right)^{1/2}, \qquad b_{2N} = \left(\frac{N^2 - 1}{4N^2 - 1}\right)^{1/2}. \tag{11d}$$

Fig. 3 contains a plot of the slant-transform basis vectors of S_{16} represented as waveforms.

C. Slant-Transform Fast Computational Algorithm

The fast computational algorithm of the slant transform is based on the matrix factorization corresponding to (10). A column vector multiplication by S_N can be accomplished by multiplying by the left factor of (10). Letting A_N denote the number of additions and subtractions for

$$S_N = \frac{1}{2^{1/2}} \left[\begin{array}{cc:c:cc:c} 1 & 0 & & 1 & 0 & \\ a_N & b_N & 0 & -a_N & b_N & 0 \\ \hdashline 0 & & I_{(N/2)-2} & 0 & & I_{(N/2)-2} \\ \hdashline 0 & 1 & & 0 & -1 & \\ -b_N & a_N & 0 & b_N & a_N & 0 \\ \hdashline 0 & & I_{(N/2)-2} & 0 & & -I_{(N/2)-2} \end{array}\right] \left[\begin{array}{c:c} S_{N/2} & 0 \\ \hdashline 0 & S_{N/2} \end{array}\right]. \tag{10}$$

The matrix $I_{(N/2)-2}$ is the identity matrix of dimension $(N/2) - 2$ and the various partition blocks are determined by the same considerations as described above for S_8. The constants a_N, b_N may be computed from the recursive relation [7]:

S_N, and M_N the number of nontrivial multiplications in the above scheme, it can be shown that A_N and M_N satisfy the recursions $A_N = 2A_{N/2} + N + 4$ and $M_N = 2M_{N/2} + 8$. However, a more economical algorithm is obtained by a further factorization of (10) into

$$S_N = \frac{1}{2^{1/2}} \left[\begin{array}{cc|c|cc|c} 1 & 0 & & 1 & 0 & \\ 0 & b_N & 0 & a_N & 0 & 0 \\ \hline & 0 & I_{(N/2)-2} & 0 & & 0 \\ \hline 0 & 0 & & 0 & 1 & \\ 0 & a_N & 0 & -b_N & 0 & 0 \\ \hline & 0 & 0 & 0 & & I_{(N/2)-2} \end{array}\right] \left[\begin{array}{c|c} I_{N/2} & I_{N/2} \\ \hline I_{N/2} & -I_{N/2} \end{array}\right] \left[\begin{array}{c|c} S_{N/2} & 0 \\ \hline 0 & S_{N/2} \end{array}\right]. \tag{12}$$

Multiplication by the middle factor involves only N additions or subtractions, while multiplication by the left factor involves only 2 additions and 4 multiplications. The recursions for the number of operations for S_N are now $A_N = 2A_{N/2} + N + 2$ and $M_N = 2M_{N/2} + 4$.

A special, more efficient factorization of S_4 has been discovered. With this factorization

$$S_4 = \frac{1}{4^{1/2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & \frac{3}{5^{1/2}} & \frac{1}{5^{1/2}} \\ 1 & -1 & 0 & 0 \\ 0 & 0 & \frac{1}{5^{1/2}} & -\frac{3}{5^{1/2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}. \tag{13}$$

If S_4 is postmultiplied by a column data vector, the first computational pass requires 4 additions and subtractions while the second pass requires a similar number and 4 multiplications. Therefore $A_4 = 8$, $M_4 = 4$.

With some rearrangement of rows and columns, (13) can be applied to S_8 to give

$$S_8 = \frac{1}{2^{1/2}} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & & & & \\ 0 & b_8 & a_8 & 0 & & & & \\ 0 & 0 & 0 & 1 & & & 0 & \\ 0 & a_8 & -b_8 & 0 & & & & \\ \hline & & & & 1 & 0 & 0 & 0 \\ & & & & 0 & 1 & 0 & 0 \\ & & 0 & & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c|c} S_4 & 0 \\ \hline 0 & S_4 \end{array}\right]. \tag{14}$$

This rearrangement orders the rows in increasing order of sequency. If S_8 is postmultiplied by a column data vector, the first and second computational passes execute the multiplication by S_4 on the first and last halves. The third pass requires 8 additions or subtractions and the fourth pass requires 2 additions and 4 multiplications. The total computation is 26 additions or subtractions and 12 multiplications. A flow chart of this computation is shown in Fig. 4.

The decomposition of the general S_N follows from repeated application of (12) [11]. A total of $N \log_2 N + (N/2) - 2$ additions and subtractions together with $2N - 4$ multiplications are required to compute the slant transform of an N-dimensional data vector. For purposes of comparison the $N \times N$ Hadamard transform requires $N \log_2 N$ additions or subtractions.

III. STATISTICAL ANALYSIS OF SLANT TRANSFORM

The development of efficient quantization and coding requires an understanding of the statistical properties of the transform domain samples. This section presents a derivation of the statistical mean and covariance of the slant transform samples and develops models for their probability densities.

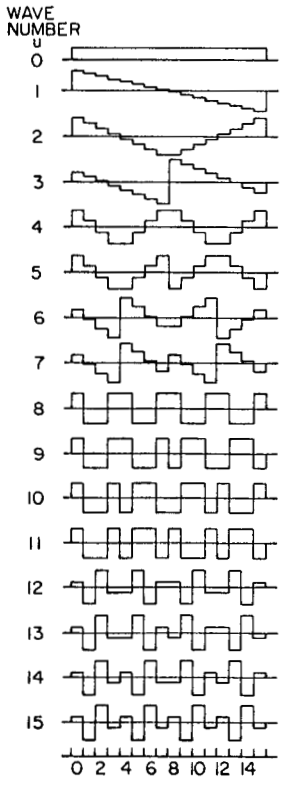


Fig. 3. Slant-transform basis waveforms.

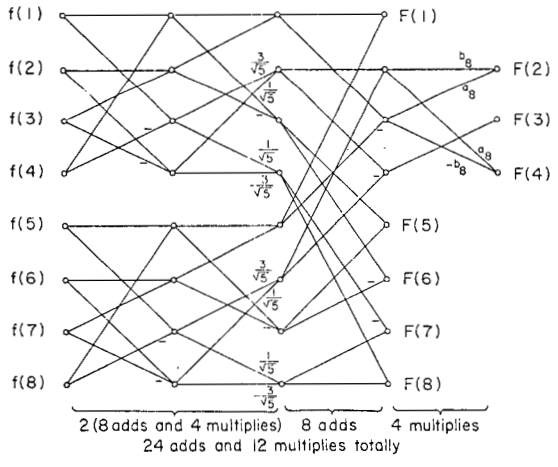


Fig. 4. Slant-transform of order 8 computational flowchart.

A. Mean and Covariance

For the statistical analysis it is assumed that the image array $F(j,k)$ is a sample of a stochastic process with mean

$$E\{F(j,k)\} \equiv \overline{F(j,k)} \tag{15a}$$

and covariance

$$E\{[F(j_1,k_1) - \overline{F(j_1,k_1)}][F(j_2,k_2) - \overline{F(j_2,k_2)}]\} \equiv C_F\{j_1,k_1,j_2,k_2\}. \tag{15b}$$

A convenient, and reasonably accurate, covariance model for an image is the first-order Markov process model for which

$$C_F\{j_1,k_1,j_2,k_2\} = \sigma_R^2\sigma_C^2\rho_R^{|j_1-j_2|}\rho_C^{|k_1-k_2|}$$

where σ_R^2 and σ_C^2 are the row and column variances, and ρ_R and ρ_C are the adjacent pixel correlation factors. From (4), the mean and covariance of the slant transform samples can then be expressed as[1]

$$E\{\mathfrak{F}(u,v)\} \equiv \overline{\mathfrak{F}(u,v)} = \sum_j \sum_k \overline{F(j,k)}\, S(u,j)\, S(k,v) \tag{16}$$

and

$$C_{\mathfrak{F}}\{u_1,u_2,v_1,v_2\} = \sum_{j_1} \sum_{j_2} \sum_{k_1} \sum_{k_2} C_F\{j_1,j_2,k_1,k_2\} \cdot S(u_1,j_1)\, S(k_1,v_1)\, S(u_2,j_2)\, S(k_2,v_2). \tag{17}$$

If the image array is stationary, then as a result of the orthogonality of the kernel,

$$\overline{\mathfrak{F}(0,0)} = \overline{NF(j,k)} \tag{18a}$$

$$\overline{\mathfrak{F}(u,v)} = 0,\ u,v \neq 0. \tag{18b}$$

No closed form expression has been found for the covariance of the slant transform samples, but the covariance may be computed as the two-dimensional slant transform of the function $C\{j_1 - j_2, k_1 - k_2\}$ if the original image field is stationary.

B. Probability Density Models

The probability density of slant transform samples is difficult to obtain since the probability density of the original image array is not usually well defined, and also, the slant transform representation is mathematically complex. However, since the transform operation forms a well behaved weighted sum over all of the pixels in the original image, one can evoke qualitative arguments based upon the central limit theorem [12] to determine reasonable probability density models for the transform domain samples.

The transform domain sample $\mathfrak{F}(0,0)$ is a nonnegative weighted sum of pixel values. Its histogram will generally follow the histogram of pixel values which is often modeled by a Rayleigh density

$$p_{\mathfrak{F}(0,0)}(x) = \frac{x}{\alpha^2} \exp\left\{\frac{-x^2}{2\alpha^2}\right\}, \qquad x \geq 0. \tag{19}$$

All other transform components are bipolar and possess a zero mean. These components generally can be modeled by a Gaussian density

$$p_{\mathfrak{F}(u,v)}(x) = [2\pi\sigma^2(u,v)]^{-1/2} \exp\left\{-\frac{x^2}{2\sigma^2(u,v)}\right\}, \qquad x \geq 0;\ (u,v) \neq (0,0) \tag{20}$$

where $\sigma^2(u,v) = C_{\mathfrak{F}}(u,u,v,v)$ is the variance of the transform samples as obtained from (17).

IV. SLANT-TRANSFORM MONOCHROME IMAGE CODING

The basic premise of a monochrome image-transform coding system is that the two-dimensional transform of an image has an energy distribution more suitable to coding

[1] Unless otherwise noted, the summation indices are 1 to N.

than the spatial domain representation. As a result of the inherent pixel-to-pixel correlation of natural monochrome images, the energy in the transform domain tends to be clustered into a relatively small number of transform samples. Low magnitude transform samples can be discarded in an analog transmission system, or grossly quantized in a digital transmission system, without introducing serious image degradation in order to achieve a bandwidth reduction.

Fig. 5 contains a block diagram of the slant-transform coding system for monochrome images. In operation, a two-dimensional slant transform is taken of the image pixels over the entire image, or repeatedly over subsections of the image, called blocks. The transform domain samples are then operated upon by a sample selector that decides which samples are to be transmitted. For an analog communication system, the selected samples are distributed uniformly in time and transmitted by analog modulation, while for a digital communication link, the selected samples are quantized, coded, and transmitted in binary form. At the receiver the incoming data are decoded, and an inverse slant transform is performed to reconstruct the original image.

There are two basic strategies of sample selection: zonal sampling and threshold sampling [1]. In zonal sampling the reconstruction is made with a subset of transform samples lying in certain pre-specified geometric zones—usually the low frequency coefficients. For analog transmission the amplitude of each component in the zone is transmitted. For digital transmission each component in a zone is quantized and assigned a binary code word. The number of quantization levels is usually made proportional to the expected variance of the component, and the number of code bits made proportional to its expected probability of occurrence. In threshold sampling the image reconstruction is made with a subset of the samples which are larger than a specified threshold. Since the locations of the significant samples must be communicated, threshold sampling is usually employed only in digital links.

The following subsections contain an analysis of the slant-transform image-coding process for a zonal sample reduction system utilizing analog transmission and for zonal and threshold coding systems over a digital link. In all instances, a mean-square error performance criterion

$$\mathcal{E} = \frac{1}{N^2} \sum_j \sum_k E\{[F(j,k) - \hat{F}(j,k)]^2\} \qquad (21)$$

is utilized. While it is known that this measure results in some anomalies, it has proven reliable as a performance measure between different transforms and for variants of sampling and coding strategies.

A. Zonal Sampling

The sample selection process for two zones can be analyzed conveniently by defining a transform domain sampling function $T(u,v)$ which takes on the value one for samples to be transmitted and zero for samples to be discarded. The reconstructed image then becomes

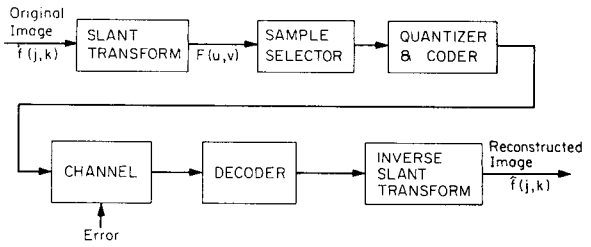


Fig. 5. Slant-transform monochrome-image coding system.

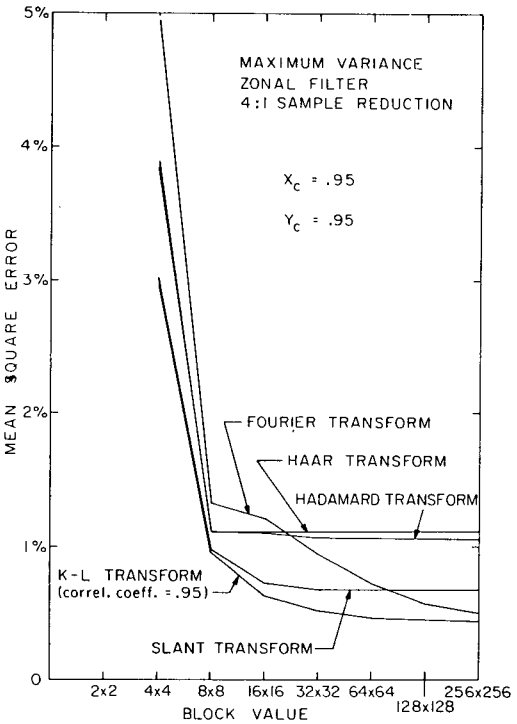


Fig. 6. Zonal sampling mean-square error performance of image transforms as a function of block size.

$$\hat{F}(j,k) = \sum_u \sum_v \mathfrak{F}(u,v) T(u,v) S(u,j) S(k,v). \qquad (22)$$

It is then easily shown that the mean-square error expression can be written as

$$\mathcal{E} = \frac{1}{N^2} \sum_u \sum_v E\{(\mathfrak{F}(u,v)[1 - T(u,v)])^2\}. \qquad (23)$$

There are a number of zones that could logically be employed for zonal sampling; for example, a rectangular, elliptical, or triangular zone. Both analytic and experimental studies [11] have indicated that the optimum zone is the so-called maximum variance zone in which $T(u,v)$ is chosen to be unity for those samples having the largest variance as computed by (17) for a given covariance model of the original image. Fig. 6 contains a plot of the mean-square error of an image with a Markov process covariance as a function of block size for various transformations. In this plot the 25 percent of the coefficients with the largest variances were selected, and the remainder discarded. From the figure it is seen that the Karhunen–Loeve transform provides the best mean-square error, while the slant transform results in only a slightly

greater error. Also note that the rate of decrease in mean-square error for larger block sizes becomes quite small after a block size of about 16 × 16.

Fig. 7 shows reconstructions of images for maximum variance zonal sampling with the slant transform in 16 × 16 blocks. For purposes of comparison, a series of experiments were performed with the Hadamard, Haar, Fourier, and Karhunen–Loeve transforms, as shown in Fig. 8. It can be seen that the slant transform provides a better subjective quality reconstruction and smaller mean-square error than any other transform except the Karhunen–Loeve transform which does not possess a fast algorithm.

B. Zonal Coding

In the zonal coding system a set of zones is established in each transform block. Transform samples in each zone are then quantized with the same number of quantization levels which is set proportional to the expected variance of the transform coefficients. For a constant word length code, $N_B(u,v)$ code bits are assigned to each coefficient, resulting in

$$L_C(u,v) = 2^{N_B(u,v)} \tag{24}$$

quantization levels. A total of

$$N_B = \sum_u \sum_v N_B(u,v) \tag{25}$$

bits are then required to code the picture. The bit assignment $N_B(u,v)$ for each coefficient is based upon a relation of rate distortion theory [13], [14]. The number of bits is given by

$$N_B(u,v) = \ln [V_F(u,v)] - \ln [D] \tag{26}$$

where $V_F(u,v)$ is the variance of a transform coefficient and D is proportional to the mean-square quantization error. Fig. 9 illustrates a typical bit assignment for coding in 16 × 16 blocks. Quantization decision and reconstruction levels are selected to minimize the mean-square quantization error for the probability density models of Section III using the Max quantization algorithm [15]. With this quantization and coding strategy, it is possible to predict the mean-square coding error for transform coding. Fig. 10 contains a plot of mean-square error as a function of block size for several transforms for coding with an average of 1.5 bits/pixel. The figure indicates that the performance of the slant transform is quite close to the optimal Karhunen–Loeve transform. It is possible to achieve a slightly lower mean-square error for a given channel rate by employing Huffman coding of the quantized coefficients rather than constant word-length coding, but the coder will be much more complex to implement.

Fig. 11 contains simulation results of zonal coding with the slant transform in blocks of 16 × 16 pixels. A comparison with other transforms is shown in Fig. 12.

C. Threshold Coding

In a threshold coding system, each sample whose magnitude is greater than a given threshold level is quantized with a fixed number of levels and its amplitude is coded. It is necessary to code the position of each significant sample in the transform plane. A simple, but quite efficient, technique for position coding is to code the number of nonsignificant samples between significant samples. This scheme, called run length coding, can be implemented as follows.

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

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

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

The advantage of including a line synchronization code group is that it becomes unnecessary to code the line number and, also, it prevents the propagation of channel errors over more than one line.

Fig. 13 contains simulation results for slant-transform threshold coding. As expected, since the coding process is adaptive, its performance is somewhat better than the simpler zonal coding process.

V. SLANT-TRANSFORM COLOR-IMAGE CODING

Fig. 14 contains a block diagram of a typical slant-transform color-image coding system. In the system, the color image is represented by three source tristimulus signals $R(j,k)$, $G(j,k)$, $B(j,k)$ that specify the red, green, and blue content of a pixel at coordinate (j,k), according to the National Television System Commission (NTSC) receiver phosphor primary system [16]. The source tristimulus signals are then converted to a new three-dimensional space, $Y(j,k)$, $I(j,k)$, and $Q(j,k)$, which specify the luminance and the chrominance information of the image pixel according to the NTSC television transmission primary system. The conversion is defined by

$$\begin{bmatrix} Y(j,k) \\ I(j,k) \\ Q(j,k) \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & 0.312 \end{bmatrix} \begin{bmatrix} R(j,k) \\ G(j,k) \\ B(j,k) \end{bmatrix}. \tag{27}$$

The reason for transform coding the $Y\ I\ Q$ signals rather than the $R\ G\ B$ signals is that the $Y\ I\ Q$ signals are reasonably well uncorrelated, and most of the color-image energy is compacted into the Y plane. This permits a more efficient design of the quantizers. Table I compares the energy distribution of the $R\ G\ B$ and $Y\ I\ Q$ color planes. The converted signals then individually undergo a two-dimensional slant transform. This results in three transform domain planes, $\mathfrak{F}_Y(u,v)$, $\mathfrak{F}_I(u,v)$, $\mathfrak{F}_Q(u,v)$, obtained

4:1 sample reduction

6:1 sample reduction

4:1 sample reduction

6:1 sample reduction

4:1 sample reduction

6:1 sample reduction

Fig. 7. Slant-transform zonal sampling in 16 × 16 pixel blocks, unquantized transform.

Hadamard transform
4:1 sample reduction

Fourier transform
4:1 sample reduction

Haar transform
4:1 sample reduction

Karhunen-Loeve transform
4:1 sample reduction

Fig. 8. Hadamard, Fourier, Haar, and Karhunen–Loeve transform zonal sampling in 16 × 16 pixel blocks, unquantized transform.

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

Fig. 9. Typical bit assignments for slant-transform zonal coding in 16 × 16 pixel blocks.

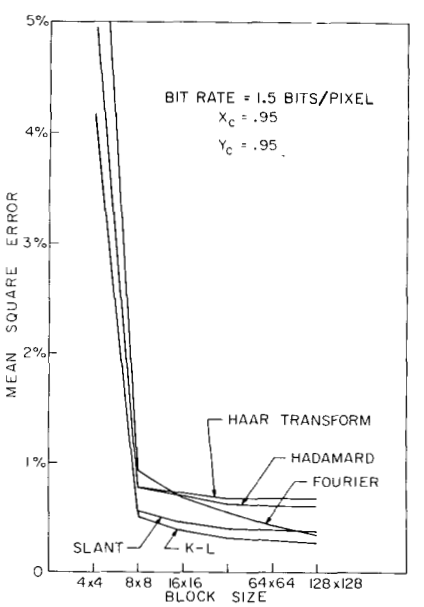


Fig. 10. Zonal coding mean-square error performance of image transforms as a function of block size.

1.5 bits/pixel

1.0 bit/pixel

1.5 bits/pixel

1.0 bit/pixel

1.5 bits/pixel

1.0 bit/pixel

Fig. 11. Slant-transform zonal coding in 16 × 16 pixel blocks, quantized transform.

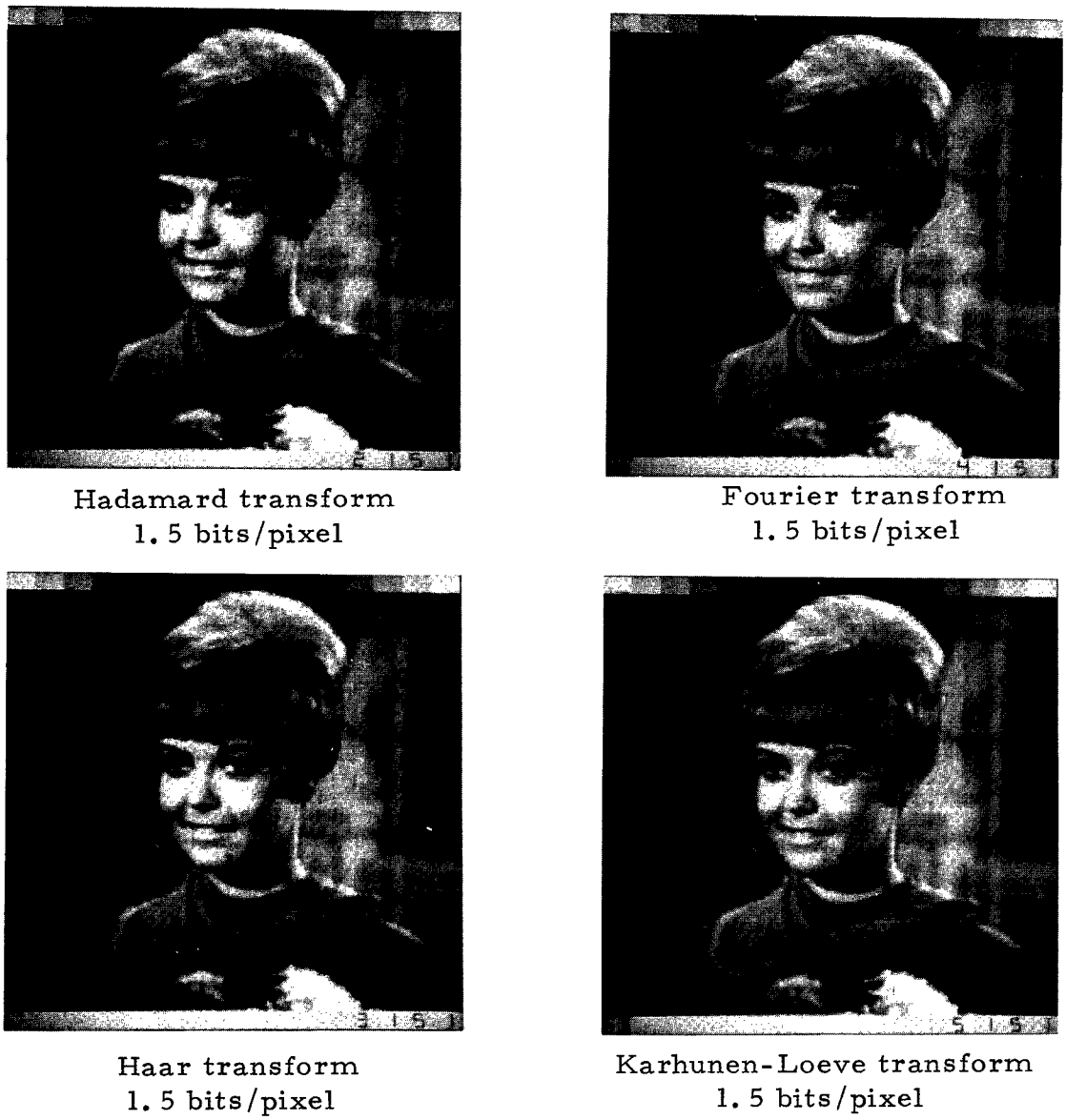


Fig. 12. Hadamard, Fourier, Haar, and Karhunen–Loeve transform zonal coding in 16 × 16 pixel blocks, quantized transform.

from

$$[\mathfrak{F}_Y] = [S][Y][S]^T \tag{28a}$$

$$[\mathfrak{F}_I] = [S][I][S]^T \tag{28b}$$

$$[\mathfrak{F}_Q] = [S][Q][S]^T \tag{28c}$$

where $[S]$ is the slant-transform matrix. Note that, since the coordinate conversion and spatial transformation are linear operations, their order may be reversed. Next, the transform samples are quantized with the number of quantum levels made proportional to the expected variance of each pixel, and with the quantization level spacing allowed to be variable to minimize the mean-square quantization error. The quantized samples $\hat{\mathfrak{F}}_Y(j,k)$, $\hat{\mathfrak{F}}_I(j,k)$, and $\hat{\mathfrak{F}}_Q(j,k)$ are then coded and transmitted over a possibly noisy channel. At the receiver the channel output is decoded, and inverse slant transforms are taken to obtain

$$[\hat{Y}] = [S]^T[\hat{\mathfrak{F}}_Y][S] \tag{29a}$$

$$[\hat{I}] = [S]^T[\hat{\mathfrak{F}}_I][S] \tag{29b}$$

$$[\hat{Q}] = [S]^T[\hat{\mathfrak{F}}_Q][S]. \tag{29c}$$

Finally, an inverse coordinate conversion results in the reconstructed tristimulus signals

$$\begin{bmatrix} \hat{R}(j,k) \\ \hat{G}(j,k) \\ \hat{B}(j,k) \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.621 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.106 & 1.703 \end{bmatrix} \begin{bmatrix} \hat{Y}(j,k) \\ \hat{I}(j,k) \\ \hat{Q}(j,k) \end{bmatrix}. \tag{30}$$

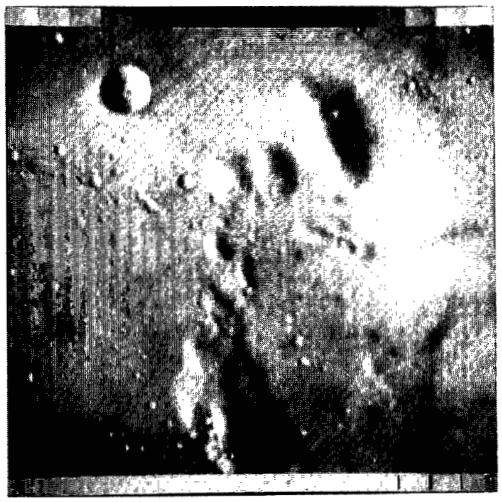
The coordinate conversion from the $R\,G\,B$ color space to the $Y\,I\,Q$ color space can be considered along with the spatial slant transforms as a three-dimensional transformation. The coordinate conversion provides an energy compaction between color planes and the spatial slant

2.0 bits/pixel

1.15 bits/pixel

2.0 bits/pixel

1.15 bits/pixel



2.0 bits/pixel

1.15 bits/pixel

Fig. 13. Slant-transform threshold coding in 16 × 16 pixel blocks, quantized transform.

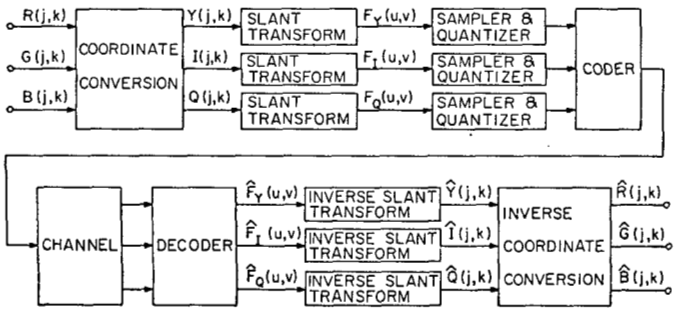


Fig. 14. Slant-transform color-image coding system.

TABLE I
ENERGY COMPACTION OF COORDINATE CONVERSIONS

Test Image	Coordinate System	Percentage of σ_1^2	Percentage of σ_1^2	Percentage of σ_3^2
GIRL	RGB	45.14	35.41	19.45
	YIQ	78.32	17.54	4.14
	$K_1K_2K_3$	85.84	12.10	2.06
COUPLE	RGB	51.55	31.09	17.36
	YIQ	84.84	13.81	1.35
	$K_1K_2K_3$	92.75	6.46	0.79

transforms provide an energy compaction within the color planes. Adopting this philosophy, the optimal three-dimensional transform would be a Karhunen–Loeve transformation which completely decorrelates the $3N^2$ color-image components. It has been shown [17] that the $Y\,I\,Q$ coordinate conversion provides almost as high an energy compaction for color images as does a Karhunen–Loeve color coordinate conversion. This result is verified by the color-image energy distribution of two color images described in Table I.

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

$$\mathcal{E} = \frac{1}{3N^2} E\{[Y(j,k) - \hat{Y}(j,k)]^2 + [I(j,k) - \hat{I}(j,k)]^2 + [Q(j,k) - \hat{Q}(j,k)]^2\}. \quad (31)$$

Several variations of the quantization procedure have been investigated [18]. The following has been found to be a highly effective implementation.

1) Model the row and column variance matrices of $R\,G\,B$ as first-order Markov processes and compute the variances of the elements of $Y\,I\,Q$.

2) Model the probability density of the constant basis vector component of F_Y by a Rayleigh density, and the probability densities of the other basis vector components of $F_YF_IF_Q$ by Gaussian densities with variances computed in 1).

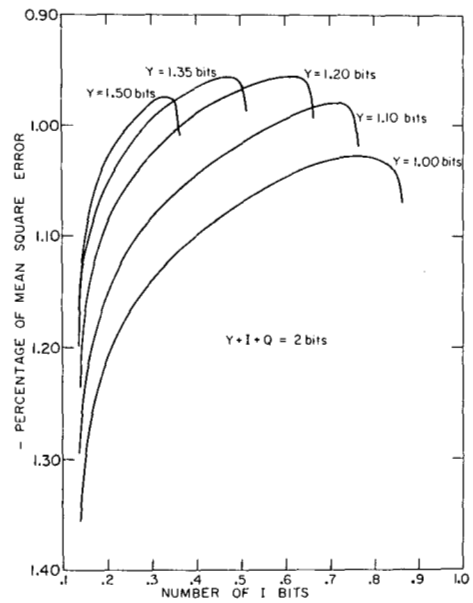


Fig. 15. Quantization error for various color plane bit assignments.

3) Assign

$$L_Y(u,v) = 2^{N_Y(u,v)} \quad (32a)$$

$$L_I(u,v) = 2^{N_I(u,v)} \quad (32b)$$

$$L_Q(u,v) = 2^{N_Q(u,v)} \quad (32c)$$

quantization levels to each transform component where the number of bits allotted to each component is made proportional to the logarithm of its variance computed in 1).

4) The total number of bits allotted to a color image is set at

$$N_B = N_{BY} + N_{BI} + N_{BQ} \quad (33a)$$

where

$$N_{BY} = \sum_{u=1}^{N}\sum_{v=1}^{N} N_Y(u,v) \quad (33b)$$

$$N_{BI} = \sum_{u=1}^{N}\sum_{v=1}^{N} N_I(u,v) \quad (33c)$$

$$N_{BQ} = \sum_{u=1}^{N}\sum_{v=1}^{N} N_Q(u,v). \quad (33d)$$

5) For a given value of N_B, select trial values of N_{BY}, N_{BI}, N_{BQ}, then compute quantization levels from 3), and with the probability density models of 2), perform a variable spacing quantization of each transform component using the Max [15] quantization rule.

6) For representative color images, compute the mean-square error, and, by iterative search techniques, determine optimum bit allocations for transform planes.

It should be noted that the above procedure need not be performed dynamically for every color image to be coded. The optimization need only be performed for typical color images to be coded to obtain a quantization scale which can be designed into the quantizer hardware.

Fig. 15 contains a plot of the mean-square error versus

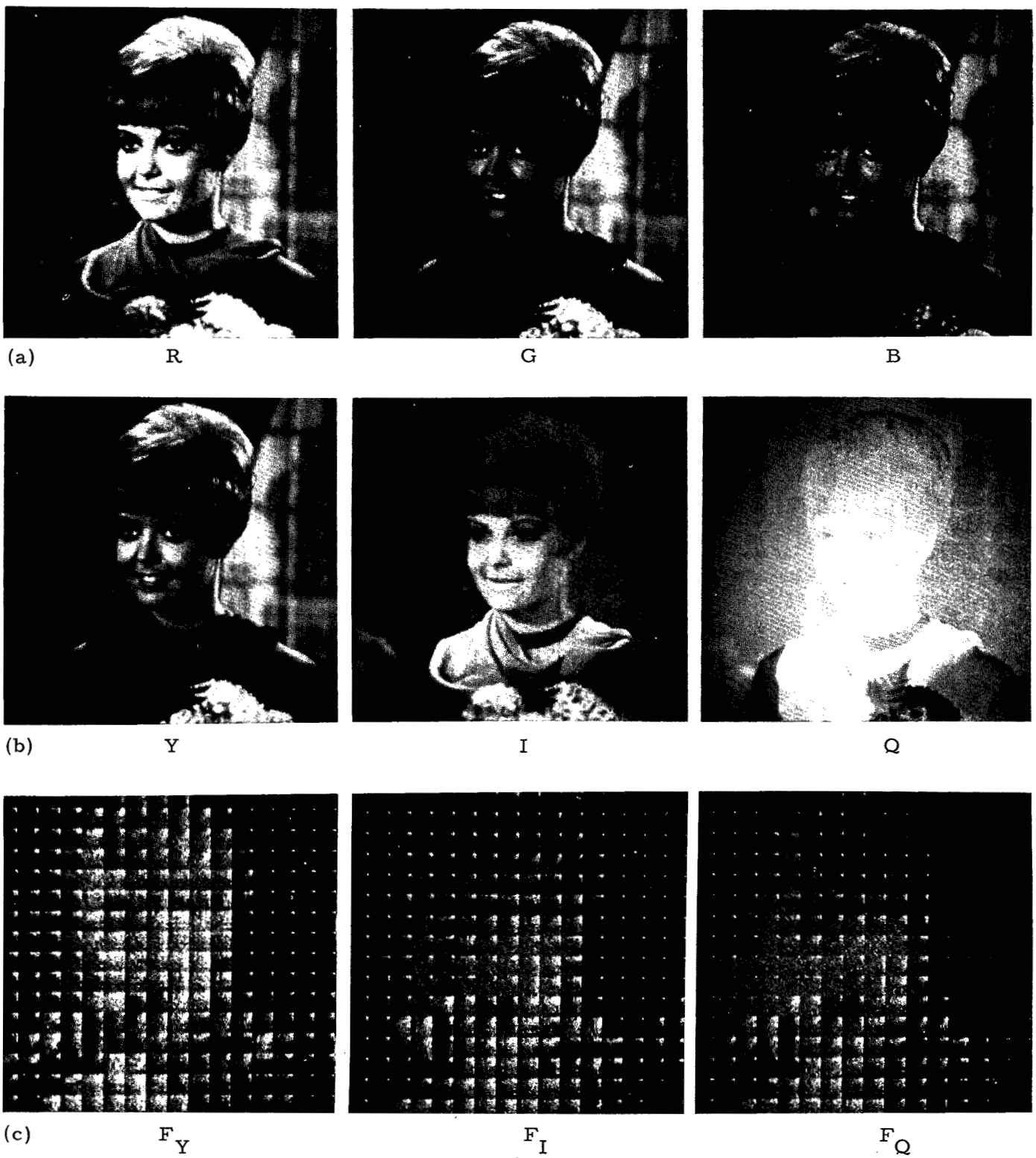


Fig. 16. Slant-transform color-image coding example.

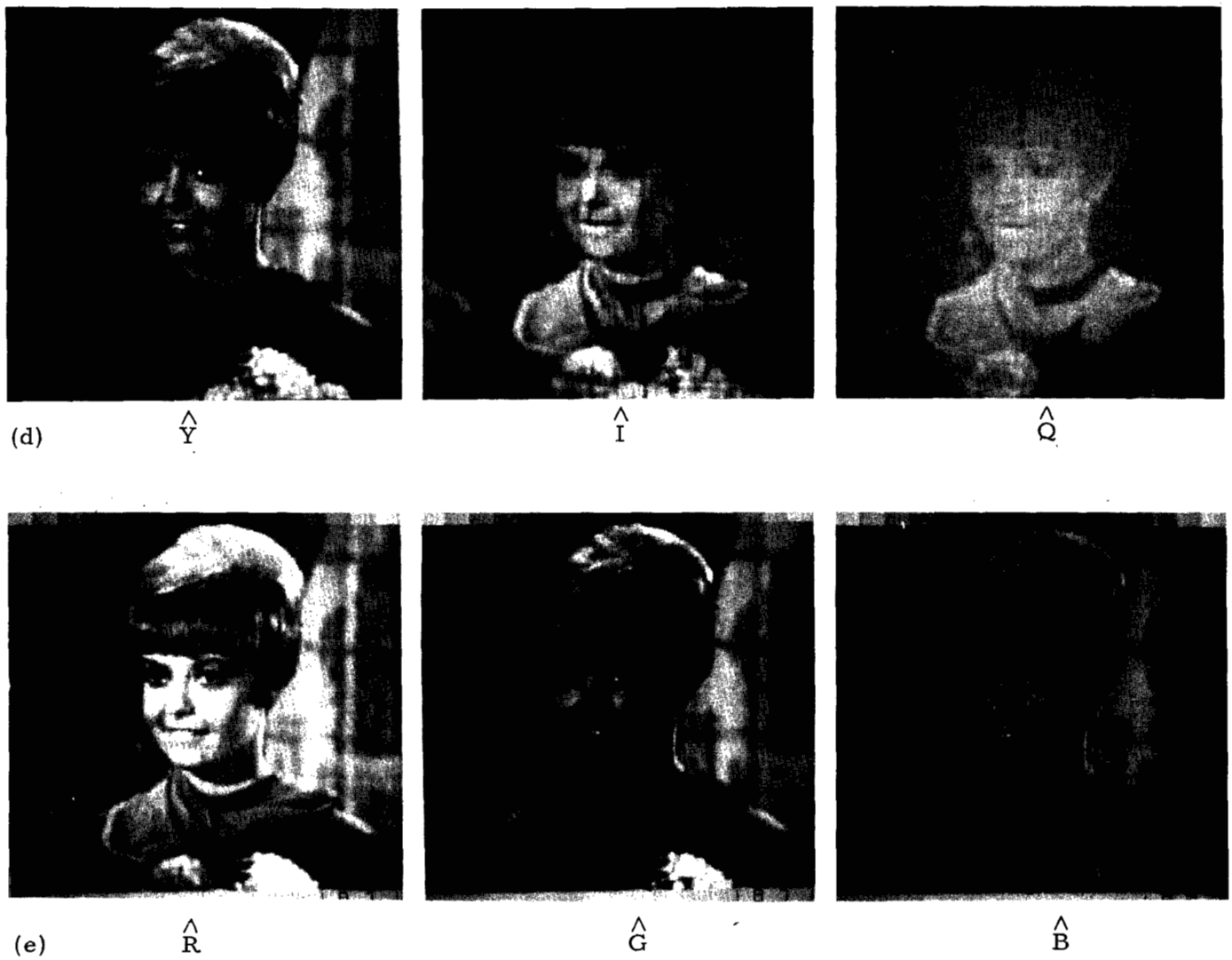


Fig. 16. Cont'd.

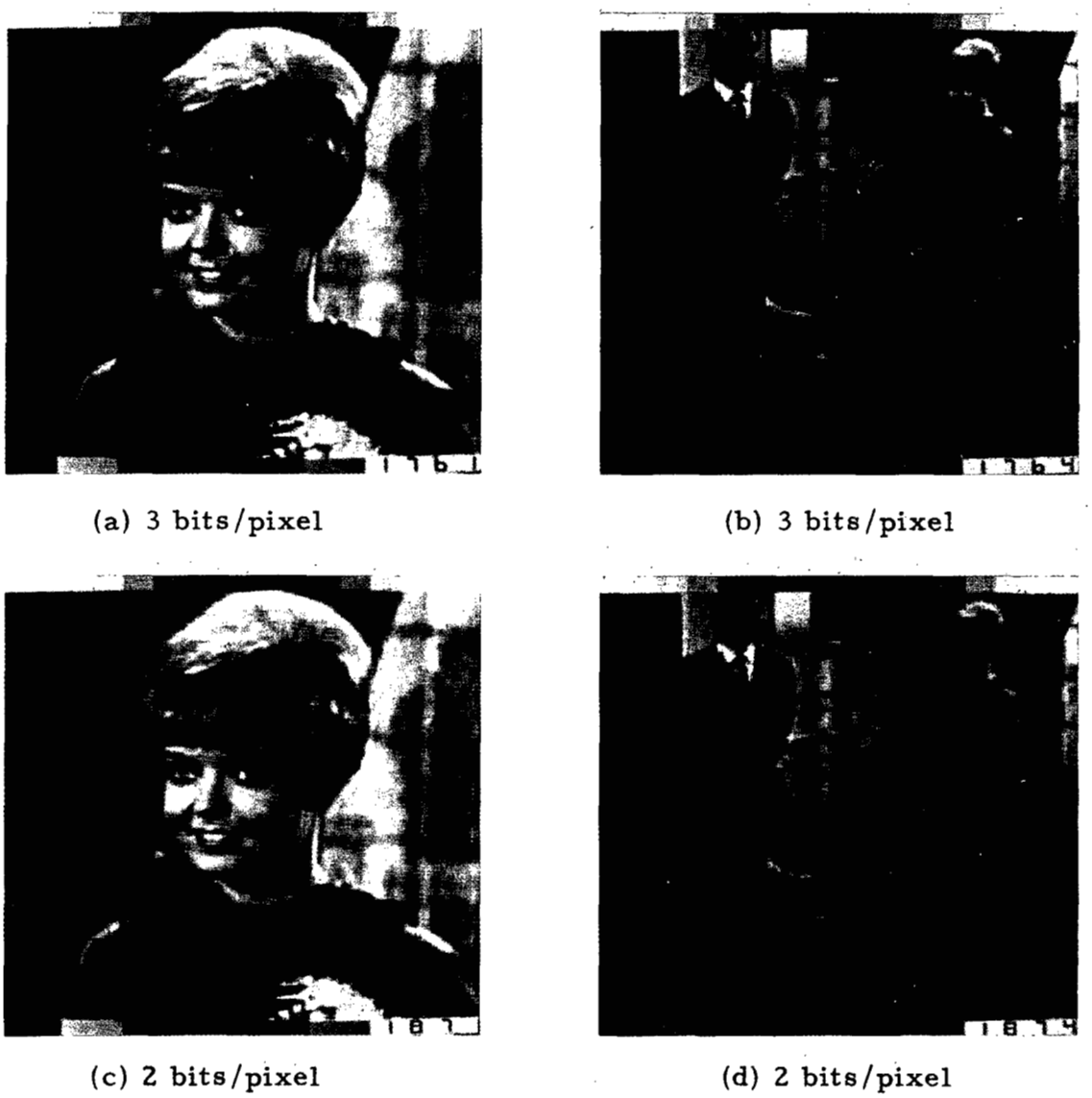
N_{BI} for several values of N_{BY} and a fixed value N_B for the girl image. The optimum average bit allocation for this test image is found to be: $N_{BY} = 1.25$, $N_{BI} = 0.55$, $N_{BQ} = 0.20$. The optimum scale does not change appreciably for other images or total bit allotments.

A computer simulation has been performed to subjectively evaluate the performance of the slant-transform color-image coding system. Fig. 16(a) contains monochrome pictures of the red, green, and blue components of an original image of 256 × 256 pixels. Each component of the original is quantized to 255 levels. It should be noted that visually the $R\,G\,B$ components are highly correlated. The corresponding $Y\,I\,Q$ components in Fig. 16(b) appear much less correlated. Fig. 16(c) contains illustrations of the logarithm of the magnitude of each slant-transform plane of the color image for transformation in 16 × 16 pixel blocks to illustrate the spatial energy compaction. In one of the simulation experiments the transform coefficients, $\mathfrak{F}_Y$, $\mathfrak{F}_I$, $\mathfrak{F}_Q$ were assigned code bits such that Y, I, Q were coded with an average of 1.2, 0.54, and 0.26 bits/pixel, respectively. The corresponding reproductions of Y, I, Q and R, G, B are shown in Fig. 16(d) and (e). In this experiment the coding has been reduced from 24 bits/pixel to 2 bits/pixel. The $R\,G\,B$ reconstructions exhibit some degradation as a result of the coding process, but the visual effect of the degradation is much less visible in the color reconstruction because of the spatial frequency limitations of the human visual system. Fig. 17 contains color images slant transform coded with 3 and 2 bits/pixel.

VI. ADDITIONAL TOPICS

The effect of channel errors on slant-transform coded images has been investigated quite extensively [11]. Studies indicate that the coding technique is relatively tolerant to channel errors. With zonal coding, for example, binary symmetric channel errors do not become subjectively noticeable until the error rate is reduced to about 10^{-3}.

Adaptive and semiadaptive quantization and coding



(a) 3 bits/pixel (b) 3 bits/pixel

(c) 2 bits/pixel (d) 2 bits/pixel

Fig. 17. Example of slant-transform color-image coding.

techniques have been developed and tested by simulation for the slant transform. These techniques, which are somewhat more complex than the fixed quantization and coding methods described in this paper, provide additional bandwidth reductions of about 2:1 for the same degree of image quality.

VII. SUMMARY

Slant-transform coding of monochrome and color images has proven to provide a substantial bandwidth reduction compared to conventional PCM coding with only a minimal amount of image degradation. It has been found that the slant transform results in a lower mean-square error for moderate size image blocks as compared to other unitary transforms with fast computational algorithms. Subjectively, it has been observed that the image quality of slant-transform coded images is somewhat higher than for images coded with other fast computational unitary transforms.

REFERENCES

[1] P. A. Wintz, "Transform picture coding," *Proc. IEEE*, vol. 60, pp. 809–820, July 1972.
[2] H. C. Andrews and W. K. Pratt, "Fourier transform coding of images," in *Conf. Rec. Hawaii Int. Conf. System Science*, Jan. 1968, pp. 677–679.
[3] G. B. Anderson and T. S. Huang, "Piecewise Fourier transformation for picture bandwidth compression," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 133–140, Apr. 1971.
——, "Correction to 'Piecewise Fourier transformation for picture bandwidth compression,'" *IEEE Trans. Commun.* (Corresp.), vol. COM-20, pp. 488–492, June 1972.
[4] W. K. Pratt, J. Kane, and H. C. Andrews, "Hadamard transform image coding," *Proc. IEEE*, vol. 57, pp. 58–68, Jan. 1969.
[5] J. W. Woods and T. S. Huang, "Picture bandwidth compression by linear transformation and block quantization," in *Picture Bandwidth Compression*, T. S. Huang and O. J. Tretiak, Ed. New York: Gordon and Breach, 1972, pp. 555–573.
[6] A. Habibi and P. A. Wintz, "Image coding by linear transformation and block quantization," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 50–62, Feb. 1971.
[7] M. Tasto and P. A. Wintz, "Image coding by adaptive block quantization," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 957–972, Dec. 1971.
[8] H. C. Andrews, *Computer Techniques in Image Processing*. New York: Academic, 1970.
[9] H. Enomoto and K. Shibata, "Orthogonal transform coding

system for television signals," *IEEE Trans. Electromagn. Compat.*, vol. EMC-13, pp. 11–17, Aug. 1971.
[10] W. K. Pratt, L. R. Welch, and W. H. Chen, "Slant transform for image coding," in *Proc. Symp. Application of Walsh Functions*, Mar. 1972.
[11] W. H. Chen, "Slant transform image coding," University of Southern California, Image Processing Institute, Los Angeles, USCEE Rep. 441, May 1973.
[12] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1965.
[13] J. Pearl, H. C. Andrews, and W. K. Pratt, "Performance measures for transform data coding," *IEEE Trans. Commun.* (Concise Papers), vol. COM-20, pp. 411–415, June 1972.
[14] L. D. Davisson, "Rate-distortion theory and application," *Proc. IEEE*, vol. 60, pp. 800–808, July 1972.
[15] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 7–12, Mar. 1960.
[16] D. G. Fink, Ed., *Television Engineering Handbook*. New York: McGraw-Hill, 1957.
[17] W. K. Pratt, "Spatial transform coding of color images," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 980–992, Dec. 1971.
[18] W. H. Chen and W. K. Pratt, "Color image coding with the slant transform," in *Proc. Symp. Application of Walsh Functions*, Apr. 1, 1973.

★

William K. Pratt (S'57–M'61) was born in Kankakee, Ill., on June 24, 1937. He received the B.S.E.E. degree from Bradley University, Peoria, Ill. in 1959, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1961 and 1965, respectively.

He received Masters and Doctoral fellowships from Hughes Aircraft Company and was employed there from 1959 to 1965. He became an Assistant Professor of Electrical Engineering at the University of Southern California in 1965, and an Associate Professor in 1969. In this capacity he is presently concerned with teaching and research in the areas of image processing and laser communications. He is director of the University of Southern California Image Processing Institute and also director of the Engineering Computer Laboratory.

Dr. Pratt is a member of Sigma Tau, Omicron Delta Kappa, Sigma Xi, and the Optical Society of America.

Wen-Hsiung Chen was born in Hsin-chu, Taiwan, on December 8, 1938. He received the B.S. degree from the National Taiwan University, Taipei, Taiwan, in 1962, the M.S. degree from the Kansas State University, Manhattan, in 1966, and the Ph.D. degree from the University of Southern California, Los Angeles, in 1973, all in electrical engineering.

He joined the Electrical Engineering Department of the National Taiwan University as a Teaching Assistant in 1963. From 1966 to 1968 he was employed with the Allis-Chalmers Company, Chicago, Ill., in the capacity of Electrical Applications Engineer. He was a Teaching and Research Assistant at the Electrical Engineering Department of the University of Southern California from 1968 to 1973, and was a Research Associate at the Image Processing Institute of USC in 1973. Since October 1973, he has been employed with Philco-Ford Corporation, Palo Alto, Calif., as a Senior Engineering Specialist working on communication system analysis and digital image processing.

★

Lloyd R. Welch was born in Detroit, Mich., on September 28, 1927. He received the B.S. degree from the University of Illinois, Urbana, in 1951, and the Ph.D. degree from the California Institute of Technology, Pasadena, in 1958, both in mathematics.

From 1956 to 1959 he worked as a Mathematician for the Jet Propulsion Laboratory, California Institute of Technology. From 1959 to 1965 he was a Mathematician with the Institute for Defense Analyses. From 1965 to 1968 he was a Visiting Associate Professor with the Department of Electrical Engineering, University of Southern California, Los Angeles, where he is presently a Professor of Electrical Engineering.

Dr. Welch is a member of Phi Beta Kappa, Sigma Xi, and Eta Kappa Nu.

(d) color girl

(e) color couple

(a) 3 bits/pixel

(b) 3 bits/pixel

(c) 2 bits/pixel

(d) 2 bits/pixel

Fig. 17. Example of slant-transform color-image coding.